



УКРАЇНА

(19) **UA** (11) **110018** (13) **C2**
(51) МПК (2015.01)
C12N 15/82 (2006.01)
A01H 5/00

ДЕРЖАВНА СЛУЖБА
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ
УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВІНАХІД

(21) Номер заявки:	а 2012 03925	(72) Винахідник(и):	Кун Йозеф Мартін (DE), Лоялл Лінда Патріція (DE), Зібберт Мальте (DE), Дувеніг Ельке (DE)
(22) Дата подання заявки:	11.08.2010	(73) Власник(и):	БАСФ ПЛАНТ САЕНС КОМПАНІ ГМБХ, D-67056 Ludwigshafen, Germany (DE)
(24) Дата, з якої є чинними права на винахід:	10.11.2015	(74) Представник:	Шамріна Олена Олексіївна, реєстр. №141
(31) Номер попередньої заявки відповідно до Паризької конвенції:	61/238233, 09169017.2	(56) Перелік документів, взятих до уваги експертизою:	WO 2007112326 A1, 04.10.2007 WO 9320216 A1, 14.10.1993 CHEN Z-L ET AL: "A DNA SEQUENCE ELEMENT THAT CONFERS SEED-SPECIFIC ENHANCEMENT TO A CONSTITUTIVE PROMOTER" EMBO JOURNAL, vol. 7, no. 2, 1988, P. 297-302.
(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:	11.06.2012, Бюл.№ 11		
(46) Публікація відомостей про видачу патенту:	10.11.2015, Бюл.№ 21		
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ	PCT/EP2010/061661, 11.08.2010		

(54) РЕГУЛЯТОРНА МОЛЕКУЛА НУКЛЕІНОВОЇ КИСЛОТИ ДЛЯ ПОСИЛЕННЯ НАСІННО-СПЕЦИФІЧНОЇ ТА/АБО НАСІННО-СЕЛЕКТИВНОЇ ЕКСПРЕСІЇ ГЕНІВ У РОСЛИНАХ

(57) Реферат:

Винахід належить до молекулярної біології рослин і забезпечує спосіб одержання вискоекспресійного насінно-специфічного та/або насінно-селективного промотору та одержання рослини з посиленою насінно-специфічною та/або насінно-селективною експресією нуклеїнових кислот, причому нуклеїнові кислоти, які посилюють експресію нуклеїнових кислот (NEENA), є гетерологічними і функціонально зв'язаними з промоторами і/або включеними у рослини.

UA 110018 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, Callis 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 (Schünmann et al., 2004). Інтрони або NEENA, які посилюють насінно-специфічну 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 премРНК. 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 і з 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 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 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 якими 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 зміну у температурному профілі, кількості циклів та/або складу або концентрації буфера для одержання відповідної молекули NEENA. 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 індукцйбельних 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, 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, індукцйбельні, специфічні для розвитку або конститутивні промотори. 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, відмінних 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. 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 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 застосовуватись 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 застосовуватись, є, наприклад, SBP-промотор з *Vicia 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 maize*, промотор глобуліну з *Zea maize*, промотор pKG86 з *Zea maize*, як описано нижче у Прикладі 6, і т. ін.

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, соя, канولا, 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 застосовуватись 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 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 однієї або кільком 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. Молекула нуклеїнової кислоти може бути кодуною білок молекулою нуклеїнової кислоти або некодуною молекулою, такою, як антисмислова РНК, рРНК, тРНК, мікроРНК, та-міРНК, міРНК, длРНК, мяРНК, мноРНК або будь-яка некодуюча РНК, відома спеціалістам у даній галузі.

5 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

10 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 одній або кільком NEENA, перебуваючи
15 під контролем зазначеного промотора, в оптимальному варіанті – насінно-специфічного та/або насінно-селективного промотора, та

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

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

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.

25 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
30 включати експресійну послідовність, яка включає вискоекспресійний промотор, функціонально зв'язаний з NEENA, у прийнятну геномну ділянку. Прийнятними геномними ділянками є, 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.

35 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 сайту початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти.

Наближення до сайту 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 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 сайту початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти. Слід розуміти, що NEENA може бути включена на
45 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 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
50 включаються після сайту початку транскрипції відповідного промотора, в оптимальному варіанті – насінно-специфічного та/або насінно-селективного промотора. Сайт включення після сайту початку транскрипції може бути в 5' UTR, 3' UTR, екзоні або інtronі, або може замінювати інtron 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'
55 UTR або інtron, або NEENA замінює інtron або, частково або повністю, 5'UTR, у найкращому варіанті – 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
60 вище у пунктах з і) по vi).

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.

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.

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.

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.

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.

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 *Agrobacterium*, 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.

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

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 Мурасіге-Скуга; NAA – 1-нафталіноцтова кислота; MES, 2-(N-морфоліно-етансульфонова кислота, IAA – індолицтова кислота; Kan – сульфат канаміцину; GA3 – гіберелова кислота; Timentin™: тикарцилін динатрію / клавуланат 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 енхансери, 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 "посилувати" 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, в 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 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)

Analyst Biochem 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, ендогенного гена або гетерологічного гена, у клітині. Наприклад, у разі структурного гена, експресія включає транскрипцію структурного гена на мРНК та, необов'язково, наступну трансляцію мРНК на один або кілька поліпептидів. В інших випадках експресія може стосуватися 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 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 або будь-яку іншу некодуючу регуляторну РНК, у смисловому або антисмисловому напрямку. Експресійна послідовність, яка включає потрібну нуклеотидну послідовність, може бути химерною, тобто, один або кілька її компонентів є гетерологічними одному або кільком іншим її компонентам. Експресійна послідовність також може бути такою, що трапляється у природі, але 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 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 (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 (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 зазначеної нуклеїновокислотної послідовності. Результат експресії може залежати від розташування нуклеїновокислотних 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. Оптимальними варіантами розташування є ті, в яких нуклеїновокислотна 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 нуклеїновокислотною послідовністю, яка підлягає рекомбінантній експресії, в оптимальному варіанті складає менше, ніж 200 пар основ, у ще кращому варіанті – менше, ніж 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

5 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

10 "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.

20

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

25 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

30 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 ендегенною або природно пов'язаною з клітиною, до якої вона включена, а була одержана з іншої клітини або синтезована. Гетерологічна ДНК також включає ендегенну послідовність ДНК, яка містить певну модифікацію, неприродні, повторювані копії ендегенної 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.

35 40 45 50

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

55 60

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 нуклеїновою 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 Fasta, яка у даний час застосовується у версії Fasta 3 (W. R. Pearson and D. J. Lipman, 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 Biomat (Biomax, 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 query-anchored, 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; -I показ GI у defines [T/F]; за замовчуванням = F; -q штраф за невідповідність нуклеотидів (лише blastn) [ціле число]; за замовчуванням = -3; -r 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 SeqAlign [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; -w вивід T Produce HTML [T/F]; за замовчуванням = F; -I 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; -u 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; -n 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 blastn 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 здійснюють 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 нуклеїнових кислотних 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 кодуючі 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 одноланцюговій або дволанцюговій формі. Якщо виділена нуклеїновокислотна послідовність має бути застосована для експресії білка, нуклеїновокислотна послідовність міститиме принаймні частину смислового або кодуючого ланцюга (тобто, нуклеїновокислотна послідовність може бути одноланцюговою). 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".

Некодуючий: термін "некодуючий" стосується послідовностей молекул нуклеїнових кислот, які не кодують частину експресованого білка або весь експресований білок. До некодуючих 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 нуклеїновоокислотна 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 заміненям 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.

Нуклеїновоокислотна послідовність: фраза "нуклеїновоокислотна послідовність" стосується одно- або дволанцюгового полімеру дезоксирибонуклеотидних або рибонуклеотидних основ, які читаються з 5'- до 3'-кінця. Вона включає хромосомну ДНК, самореplikовані плазмідні, інфекційні полімери ДНК або РНК та ДНК або РНК, які виконують здебільшого структурну роль. "Нуклеїновоокислотна 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 нуклеїновоокислотний 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 транслюється 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, фосфодіестерами) 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.

5 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
10 будь-якій потрібній стадії розвитку після стадії сіянців. Сіянець означає молоду 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
15 (печінкові мохи) та Musci; птеридофіти, такі як папороть, хвощі та плауни; голонасінні, такі як хвойні, цикадові, гінкго та Gnetatae; водорості, такі як Chlorophyceae, Phaeophyceae, Rhodophyceae, Muxophyceae, 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
20 видам, як Graveolens dulce (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 Tastie (капуста), oleracea cv Snowball Y
30 (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.

45 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>,
50 <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
55 транскрипції для започаткування транскрипції. Зазначений промотор включає, наприклад, 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
60 100 п. о. У ще одному оптимальному варіанті втілення промотор включає принаймні 50 п. о.

поблизу від сайта 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 – 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.

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 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.

"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 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 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, 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" 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 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 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 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 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 відрізку 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 GAP-аналіз 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 хромосому 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" живаються взаємозамінно, 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, мутований 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). 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*

- 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:

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 Pфam-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 / антиквітин	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 застібки (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	aataatccatggtttgacctacaataacaaagcagtc	21	
NEENAss2_for	tttttggtaccagttctttgctttcgaagttgc	22	2
NEENAss2_rev	tttttccatggtactacgtactgttttcaattct	23	
NEENAss3_for	aaaaaagggtaccatttccacacgctttctatcatcttc	24	8
NEENAss3_rev	aaaaaacatggttatctctctctaaaaataaaaaacgaatc	25	
NEENAss4_for	aataaagggtaccgtccagaattttctccattga	26	6

Праймерні послідовності

Назва праймера	Послідовність	SEQ ID NO	ПЛР, що забезпечує SEQ ID NO
NEENAss4_rev	aataaaccatggtcttactatccaaagctctca	27	
NEENAss5_for	tttttggtaccgtctactttcattacagtgactctg	28	9
NEENAss5_rev	tttttccatggttatattttacctgcaacacaattcaa	29	
NEENAss6_for	ttttatggtaccactcgaatactgcatgcaa	30	16
NEENAss6_rev	ttttatccatggttatgtagcctttacacagaaaacaa	31	
NEENAss7_for	tatataggtaccaacaactatggcctgagggt	32	18
NEENAss7_rev	tatataccatggttatcttactgttttaaccaaataaaat	33	
NEENAss8_for	tttttaggtaccatcttagggtttcgcgagatctca	34	11
NEENAss8_rev	tttttccatggtgctaagctatctctgtaataataaaattg	35	
NEENAss9_for	tttttggtaccattttgttggtgaaaggtaga	36	13
NEENAss9_rev	ttttaccatggttacgttttgtctctgcttctct	37	
NEENAss10_for	tatattggtacctctgggaatatcgattttgatct	38	15
NEENAss10_rev	tatataccatggtctcaccacatcccaaagctc	39	
NEENAss11_for	ttttatggtaccgcacaactcttagcttaccttgaa	40	10
NEENAss11_rev	ttttatccatggttatattaaccacaagccttgccctc	41	
NEENAss12_for	tttttaggtacctgtcggagaagtgggacg	42	17
NEENAsskor-rev	ttttaccatggagaagtgggacgacg	43	
NEENAss13_for	ttttatggtacctagcttaactcagattcgatcgt	44	7
NEENAss13_rev	ttttatccatggttagtatctacataccaatcatacaaatg	45	
NEENAss14_for	tttttggtacctttcacgatttgaattga	46	5
NEENAss14_rev	tttttccatggtctacaacattaaaacgaccatta	47	
NEENAss15_for	tatataggtaccagggttctgtttgtttca	48	3
NEENAss15_rev	tatataccatggttatctctgctcaaagaaacca	49	
NEENAss16_for	tttataggtaccagaagctcatttctcgatac	50	12
NEENAss16_rev	tttataccatggtctctgcgcaaaaattcacc	51	
NEENAss17_for	tatatggtacctctaaaatacagggcacc	52	19
NEENAss17_rev	tatattccatggttactctcgttcgagaagccta	53	
NEENAss18_for	tatataggtaccactgtttaagcttactgtct	54	4
NEENAss18_rev	tatataccatggttcttctaaagctgaaagt	55	
NEENAss20_for	tatataggtaccttaagctttaagaatctctactcaca	56	14
NEENAss20(2)_rev	atatatccatggttaaatttacctgtcatcaaaaacaaca	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 HF буфера

5 4,00 мкл дНТФ (2,5 мМ)

2,50 мкл прямого праймера (10 мМ)

2,50 мкл зворотного праймера (10 мМ)

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

10 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 комплексу 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 комплексу Qiagen Gel Extraction Kit (Qiagen, Hilden, Німеччина).

1.3 Побудова вектора

20 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 р-, 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.

- 5 Вектор рENTR/A, який містить промотор р-AtPXR, клонували через сайт-специфічну рекомбінацію (BP-реакцію) між вектором рDONR/A та продуктами ампліфікації р-AtPXR з праймерами р-AtPXR-for та р-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 рENTR/A піддавали аналізу послідовності для перевірки
- 10 правильності промотора р-AtPXR.

Таблиця 3

Праймерні послідовності (р-AtPXR)

Назва праймера	Послідовність	SEQ ID NO.
р-AtPXR-for	ggggacaacttgtatagaaaagttggccacatcatgttagacttacc	58
р-AtPXR-rev	ggggactgctttttgtacaaactgtttaccttttatattatataatag	59

- 15 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 рENTR/B показано у Таблиці 4, з промоторними молекулами, які мають префікс р-, кодуючими послідовностями, які мають префікс с-, та термінуючими молекулами, які мають префікс t-.

20

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

- 25 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, кодуючою

- послідовністю люциферази 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 кодуючою послідовністю люциферази 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 Мурасіке-Скуга 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'ядоль визначали у розчині, який містив 0,1 mM 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 pENTR/C. Шляхом виконання сайт-специфічної рекомбінації (LR-реакції), як було описано вище (див. вище, 1.3), pENTR/A, pENTR/B та pENTR/C, які містили касету селектованого маркера, комбінували з вектором призначення pSUN згідно з інструкціями Multisite Gateway від виробника (Invitrogen, Carlsbad, CA, США). Реакції забезпечували один бінарний вектор з промотором p-AtPXR, кодууючою 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 кодууючою 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 Мурасіге-Скуга (Murashige and Skoog, 1962, Physiol. Plant 15, 473), доповненому 3 % сахарозою (3MS-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 3 MS. 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 бензиламінопурина (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 10 000 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/хв).

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 планшет-ридері (Thermo Electron Corporation, Multiskan Ascent 354). Загальну кількість білка у зразках розраховували за попередньо побудованою стандартною кривою концентрації. Значення за результатами співвідношення RLU/хв 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 насінно-специфічний 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. 2а). 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 насінно-специфічного 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	ggggacaactttgtatagaaaagttcacgggcaggacatagggactactac	63
p-LuPXR-rev	ggggactgctttttgtacaaactggatttatgataaaaatgtcggtttc	64
p-VfUSP-for	ggggacaactttgtatagaaaagttctgcagcaaatttacacattgccac	65
p-VfUSP-rev	ggggactgctttttgtacaaactgactggctatgaagaaattataatc	66

15 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

20 промотором p-LuPXR, кодуючою послідовністю люцифери 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 люцифери 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

25 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	

30 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 насінно-специфічними 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).

Експресія LUC у листі стійко трансформованих рослин олійного рапсу

Рослинний вектор експресії	Склад експресійної касети Промотор:SEQ ID NO:ген-репортер:термінатор	Експресія LUC у листі олійного 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 родів 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 експлантів 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 ANAS 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-KG86 °F однодольних, що не містить NEENA, з *Z. mais*, комбінують з кодуючою послідовністю гена бета-глюкуронідази (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:

5

Таблиця 11

Праймерні послідовності

Назва праймера	Послідовність	SEQ ID NO	ПЛР, що забезпечує SEQ ID NO
NEENAss1_forII	aataatggcgcgccttggtgcttaaacactctggtgagt	68	1
NEENAss1_revII	aataatggcgcgcctttgacctacaaaatcaaagcagtca	69	
NEENAss2_forII	tttttggcgcgccagttcttgccttcgaagtgc	70	2
NEENAss2_revII	tttttggcgcgcctactacgtactgtttcaattct	71	
NEENAss4_forII	aataaaggcgcgcgcctccagaattttctccattga	72	6
NEENAss4_revII	aataaaggcgcgcgcctcttactatccaaagctctca	73	
NEENAss13_forII	ttttatggcgcgcctagcttaatctcagattcgaatcgt	74	7
NEENAss13_revII	ttttatggcgcgcctagctatctacataccaatcatacaaatg	75	
NEENAss14_forII	tttttggcgcgccttcacgatttgaatttga	76	5
NEENAss14_revII	tttttggcgcgcctctacaacattaaaacgaccatta	77	
NEENAss15_forII	tatataggcgcgcgcaggttctgtttgtttca	78	3
NEENAss15_revII	tatataggcgcgcgccttatctcctgctcaaagaaacca	79	
NEENAss18_forII	tatataggcgcgcgcactgttaagcttactgtct	80	4
NEENAss18_revII	tatataggcgcgccttcttctaaagctgaaagt	81	

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

10

ПЛР-фрагменти NEENA (див. вище) клонують окремо перед кодуючою послідовністю бета-глюкуронідази з застосуванням сайтів рестрикції *AscI*. Реакція забезпечує один бінарний вектор з промотором p-KG86, кодуючою послідовністю бета-глюкуронідази c-GUS та термінатором t-nos 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 p-, кодуючими послідовностями, які мають префікс c-, та термінуючими молекулами, які мають префікс t-.

15

20

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, двотижнєве 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 Muracige-Скуга (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. Остаточно 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 двотижнєве 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 мл/л), 0,2 М маніту, 0,2 М сорбіту, pH 5,8]. Ці шматки інкубують протягом 1-2 годин.

У разі культивованих клітин BMS однотижнєву суспензію клітин центрифугують при 1000 g у центрифугі 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-зabarвлення, 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 мМ NaHPO₄, 10 мМ EDTA, 0,05% Triton X100, 0,025% розчину X-Gluc (5-бromo-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 насінно-специфічний 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 трансформації культурних рослин (Gelvin and Schilperoort, 1995, Plant Molecular Biology Manual, 2nd Edition, Dordrecht: Kluwer Academic Publ. ISBN 0-7923-2731-4; Glick and Thompson (1993) Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, ISBN 0-8493-5164-2). 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 мМ NaHPO₄, 10 мМ EDTA, 0,05% Triton X100, 0,025% розчину X-Gluc (5-бromo-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 кодуючою 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	atatacgcggtggtgcttaaacactctggtgagt	83	1
NEENAss1_revIII	atatggcgcgcccttgacctaataaagcagtc	84	
NEENAss2_forIII	atatacgcgtagttcttgcgttcgaagttgc	85	2
NEENAss2_revIII	atatggcgcgccctactacgtactgtttcaattct	86	

Ампліфікацію під час ПЛР та очищення продуктів ампліфікації здійснювали, як детально

описано вище (приклад 1.2). Після рестрикції ДНК за допомогою рестрикційних ендонуклеаз MluI (10 О/мкл) та AclI (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 люцифери світляка з застосуванням сайтів рестрикції AclI. Реакція забезпечувала один бінарний вектор з промотором p-KG86, кодуючою послідовністю люцифери світляка c-LUC та термінатором t-nos і два вектори, які включали SEQ ID NO1 та NO2, безпосередньо перед кодуючою послідовністю люцифери 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 Промотор:SEQ ID NO:ген-репортер:термінатор	SEQ ID NO
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 AHAS 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. 4a). Натомість у листі кукурудзи не виявляли значних змін активності люцифери, 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 pENTR/B LJK1, LJK19 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 люцифери світляка c-LUC та термінатором t-nos, а також 2 вектори, які включали SEQ ID NO1 та NO2 безпосередньо перед кодуючою послідовністю люцифери 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 p-, кодуючими послідовностями, які мають префікс c-, та термінуючими молекулами, які мають префікс t-.

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).

5 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).

40 8.3 Послідовності NEENA опосередковують сильне тканинно-специфічне підвищення експресії генів у рослинах рису

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 NO 1 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:

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).

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

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.

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.

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 – CD30972) 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 Саєнс Компані ГМБХ

<120> Регуляторні молекули нуклеинових 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 tgcgcatct cattaccatt      60
tcttaaattt ctctccctaa atattccgag ttcttgattt ttgataactt cagggttttct      120
ctttttgata aatctggtct ttccattttt ttttttttgt ggttaattta gtttcctatg      180
ttcttcgatt gtattatgca tgcctctgtt ttggattctg ttagattatg tattggtgaa      240
tatgtatgtg tttttgcatg tctggttttg gtcttaaaaa tgttcaaact tgatgatttg      300
attgaagctt ttttagtggt gggttgattc ttctcaaaac tactgttaat ttactatcat      360
gttttccaac ttgattcat gatgacctt ttgttctgct ttgttataaa attttggttg      420
gtttgatttt gtaattatag tgtaattttg ttaggaatga acatgtttta atactctgtt      480
ttcgatttgt cacacattcg aattattaat cgataattta actgaaaatt catggttcta      540
gatcttggtg tcatcagatt atttgtttcg ataattcatc aaatatgtag tccttttgc      600
gatttgcgac tgtttcattt tttctcaaaa ttgttttttg ttaagtttat ctaacagtta      660
tcgttgtaaa 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 ccttgctcct tgcctatgta aaattactag gttttcatca gttacactga      120
ttaagttcgt tatagtggaa gataaaatgc cctcaaagca ttttgagga tatctttgat      180
ttttcaaaga tatggaactg tagagtttga tagtggtctt gaatgtggtt gcatgaagtt      240

```

tttttgggtct gcatgttatt ttttcctcga aatatgtttt gagtccaaca agtgattcac	300
ttgggattca gaaagttggt ttctcaatat gtaacagttt ttttctatgg agaaaaatca	360
tagggaccgt tggttttggc ttctttaatt ttgagctcag attaaacca ttttaccggy	420
tgttcttggc agaattgaaa acagtacgta gtacc	455

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

<400> 3	
agggtttcgt ttttgtttca tcgataaaact caaagggtgat gattttagggt tcttgtgagt	60
gtgctttttt gtttgattct actgtagggt ttatgttctt tagctcatag gttttgtgta	120
tttcttagaa atgtggcttc tttaatctct gggtttgtga ctttttgtgt ggtttctgtg	180
tttttcatat caaaaacctt ttttttccga gttttttttt acaaattctt actctcaagc	240
ttgaatactt cacatgcagt gttcttttgt agattttaga gttaatgtgt taaaaagttt	300
ggatttttct tgcttataga gcttcttcac tttgattttg tgggtttttt tgttttaaag	360
gtgagatttt tgatgagggt tttgcttcaa agatgtcacc tttctgggtt tgtcttttga	420
ataaagctat gaactgtcac atggctgacg caattttgtt actatgtcat gaaagctgac	480
gtttttccgt gttatacatg tttgcttaca cttgcatgcy tcaaaaaaat tggggctttt	540
tagttttagt caaagatttt acttctcttt tgggatttat gaaggaaagt tgcaaaacttt	600
ctcaaatttt accatttttg ctttgatggt tgttttagatt ggcacagaac aaactcatat	660
atgttgaaat ttttgcttgg ttttgatatag gattgtgtct tttgcttata aatgttgaaa	720
tctgaacttt ttttttgttt ggtttctttg agcaggag	758

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

<400> 4	
actgtttaag cttcactgtc tctgaatcgg caaaggtaaa cgtatcaatt attctacaaa	60
cccttttatt tttcttttga attaccgtct tcattgggtta tatgataact tgataagtaa	120
agcttcaata attgaatttg atctgtgttt ttttggcctt aatactaaat ccttacataa	180
gctttgttgc ttctcctctt gtgagttgag tgttaagttg taataatggt tcactttcag	240
ctttagaaga aa	252

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

<400> 5
 tttcacgatt tggaatttga ttcctgcgat cacaggtagt acaggtaga ttttgttttg 60
 tatagtgtga tacatacttc tttgtgatgt tttgtttact taatcgaatt tttggagtgt 120
 ttttaaggtct ctcgtttaga aatcgtggaa aatatcactg tgtgtgtgtt cttatgattc 180
 acagtgttta tgggtttcat gttctttgtt ttatcattga atgggaagaa atttcgttgg 240
 gatacaaatt tctcatgttc ttactgatcg ttattaggag tttggggaaa aaggaagagt 300
 ttttttggtt gggtcgagtg attatgaggt tatttctgta tttgatttat gagttaatgg 360
 tcgttttaat gttgtag 377

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

<400> 6
 gtccagaatt ttctccattg aagctggatt ctaaggtagt ttcttacttc tttatctcaa 60
 tctgatgatt ccatatcgaa agtcttactt tttcacttca atttcaatct gatgattcta 120
 agatctttga ttcgaggtcg atctctgata gttactacat gtttctgggt ttatttattt 180
 ttaatccata tagtaattaa aaactcttat gaggtttaat tatggttact tgagaatttg 240
 caatcgtcat ctttctttga ctctatcca ttttttggtt tttcctttgt ttaatttctg 300
 tttcataatt gtaattgtaa attaaccaaa acaaattgat cagaaacctt tttcctatgg 360
 aatatttatc acacgcaagc ctgtgagttg tgactctgta atcacttctt tgttctggta 420
 atttcagtgg ttaaggctct ctttttttct gatgttgtca gcaaaagtta gtttttcttc 480
 ttctttaatg ggtaattac acctaaatct ctggttatta aacaatccag aaagaaaaaa 540
 agtttattcc ttctctatg tatatagttt cacatgcaag catcacttgt ttgttctgac 600
 aaattgcaga gttttgagtt ctgttttttt ttttttctaa tgttttgtct ttaagaaagt 660
 tctgtttttt tttctgcagg aaagtatca aaagttttga gagctttgga tagtgaag 718

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

<400> 7
 cttagcttaat ctcagattcg aatcgttcca tagtggtgag cttcgtgttc ttctttcgtc 60
 tcttactcct gattctcgat tttagggttt tcagtaattg cgtcggcggc gaaagtcttt 120
 atcgccgac gatcttctt atctagaaat tattgatcag aaactgttgg gttttgtttg 180
 attcttgtca agttttgatt tttcatgca aattgctcaa tcccaattca aagttacgat 240
 ttttattgaa aaccctagat tggtttcttc aagtttgtca ctttgattca atctaatagc 300
 tttagcttaat cgtaagtct cttttttgggt tttagggttc atttgcgatt taaaggttct 360

tgttttggtta tttgttttgc tttggtcctt taagtttgag aggcttatgt agattataag	420
agagaagagt attgctttgc atgtttaaag gaagaacttt taactgaaca tttgtatgat	480
tggtatgtag atact	495

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

<400> 8	
atttccacac gctttctatc atttccaccc aaaaggtaac gcgcttttta tttcctttcc	60
tgcattcata aatttgtctc ctgcatgttg aaaaaaaaaa atttaccatcg agattcgttt	120
ttatttttta gagagagat	139

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

<400> 9	
gtctactttc attacagtga ctctgcatgc ttcagggtctc gtctaattct tgaattctct	60
tcttttctgt tccgtaattt actttctagg gtctctagat ttgtgtctcc tctaacaaaa	120
gaccttatct ttcgacaaat ttaatttcat cattgacctt tgtcgattcc attctctctc	180
tatctctctg tttcttcgaa aacctagagg ttttgaattt aatgattcct ttttatgtca	240
ataaatttgc aatcaatggg agctttttta aatcatcggt atatctataa acaaaaaaac	300
agtaattact cttcttagat ctaaaacaat taataaatct ttcccttttt tctcatcata	360
attttttcgt atttaactct tgtaaaaatt tgcttagccg tttcgctttc tcaggcccca	420
ggtgattcgt gtcttctagg tcagcttggt aaacctgaga gaagccatct tttgtttgcg	480
gttacaaact ttgccgcttc aatatttcat tgctgttttc tgggaaaacc tttttctagt	540
tttttcggct tattatgcct ttttaacttt tgtgcattta acattttattg ttagtgcttt	600
gcttagtgta aagtagtagt tctctttgta atattaccat aagggtcaga agtaaathtt	660
tctaaaattg ttttcttggt ggaaattcag actgatttca gcaacatgca tgggcttaaa	720
atcagcttct aagactgaga tttagtgacc agtggtggtg tgtcttggtc tctgttcttg	780
ggagaacaca aaggcagtggt ggagtctggt gagttttctg attcttgaaa agatttataa	840
attttcttgc aaaattagtc tttatgttga attgtgttgc aggtaaaat	889

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

<400> 10

```
gcacaatctt agcttacctt gaatcacaaac ttcaggtata tgtaactgat tctaaattga      60
agattgtgtg caaatcttat atccattttt tattattaaa tttattgaaa aagctagcgg      120
tgtaaattaa tgtcacaaaa tcagtatatatt gttagttttt gttttttttg aagttttatg      180
caaatcttca aaaagtatat tcagtgttgt aattgacaaa tagagactct agttcttttt      240
ttttttttct tttttttaac atctgactct tatagagact ctagttcatg tacacttttt      300
ttaatggaaa aacaaatttg aaactgaata tcttatttcc acgtagattg tatattagtt      360
taatttgatt gttatatttg taaatgtcta ctaaacagga attggatggg gaggaggcaa      420
ggcttgtgga tta                                                                433
```

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

```
<400> 11
atcttagggg ttcgcgagat ctcaactctca ctggatgtgc tgtgtttctt cttccatttt      60
ctgtttctat tggaaacttc tctctccaat ttcgttttct tcacttcttt gatccttttag      120
ctttgacaaa accgtagtaa aggatcaaaa gttatcatct ttggtccatg ttgtgaatcg      180
tgctctgctt gggctgtgac tcccaaattc ggatttgaaa ccagcatatc tgagcttaat      240
tcgagcatgc atgcgcttct ttttttctga ttttttttag actttggttc taaatccctt      300
aactttggat taactgtcaa tctacaattt tatattaaca gagatagctt agca              354
```

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

```
<400> 12
cagaagctca tttcttcgat acgatcaacc attaggtgat ttttttctct gatcttcgag      60
ttctgataat tgctcttttt tctctggctt tgttatcgat aatttctctg gattttcttt      120
ctgggggtgaa tttttgcgca gag                                                                143
```

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

```
<400> 13
atttttgttg gtgaaaggta gaattcgtaa atttcttctg ctcactttat tgtttcgact      60
cataccgat aatctcttct atgtttggta gagatatctt ctcaaagtct tatctttcct      120
taccgtgttc tgtgtttttt gatgatttag gtgaagaaga agaagcagag acaaaaacga      180
tt                                                                182
```

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

<400> 14
 ttaagctttt aagaatctct actcacattt tctctgtgag tgttctttta tacttctttg 60
 ttattttcaa tttttctttc tttcctctaa aaatttttagg aactattgaa tcatttaatt 120
 tctgtttggt gataaaaattt cgatcaactg ttctcggctt accgatgcat tttttgtaaa 180
 accgtctttt tttggtgaat aaaattttta attcatataa aaaaaaaca tatttgatac 240
 tatttttagct ccattgtatc tgaatcttca tttgttaatt tttttgtttc ctctgttctc 300
 acttgaattt tggaatattt tctctagggt ttacctata ttcttcactt taagaactat 360
 atgaagattt gattggaagt aataatattc ggtgatagaa tctgagtttg tttgattctg 420
 gtgtggggct tatactaac ttttttcttt gtaccaatac attttcaatt ttacattttt 480
 gattagctta aaatgtgaag gataccttgt aaataactat tacactattg cttgtcttag 540
 tctaatagtc ttactaata ttttgtgcag tagaagtaaa tattataaag agttgttggt 600
 tgattataga gagttgttgt ctattcttta acttgatgtg atgttgtttt tgatgacagg 660
 taaaa 665

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

<400> 15
 tctgggaaat atcgattttg atctattaag agctggtgag agccaaagtt tcctttttgt 60
 ttgtttgttt gtttgtttgt tgtttgtatt tttgtatctc tgtgatcgct tctacgtgtt 120
 gggatcatgca 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 ctctgcccac 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 tccaggtact tatgtttttg ataattttat 60
 tgatagactc tttacaatta tacttaagct tgttactttt tattgttacc aacaaaagct 120
 aatgtatagt tcataactca caggtcctgc gtctttcggg cgcaccactt ctcctacaga 180
 ttcgcaaact tctgatcctg aagggtactcg cgaacttttt actgcaactt ctagttctaa 240
 ctccaaaaca ttttgttcag aatttgtttc taaaagattt tcgggtttgt tgacgtcaca 300
 taactcgcag ggtctgcttc tttccgtccg cccacttttc cgaca 345

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

 <400> 18
 aacaactatg gcctgagggt aacaagagta tcaggtatat gtgaaaactc tacttttgaa 60
 gtttaccaaa aaaaatactc tacttttgga aagacattgc tcctaaaatc ttattagttg 120
 tatataattt actaaaacac atagtctctg aattcttggt aatgagcatg ttaccttgga 180
 caagtgaccc tttttctaca ttttgttttt ctatcacacg tcatgcggtt tgattgtttc 240
 cttacgagtt ttaattttat tttttggtta aaaacagtaa gataa 285

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

 <400> 19
 tctaaaaata cagggcaccg aaccaataa aggtgagaat gatgagaagc cgtttcttac 60
 ttttcattgt tttcttctct ctatccctct tcatttcttc tctgatcgcc agtgatttag 120
 gcttctgcaa cgaagag 137

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

 <220>
 <223> синтетична послідовність

 <400> 20
 aataatggta 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 ccagttcttt 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 acgctttcta tcatttc 37

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

<220>

<223> синтетична послідовність

<400> 25
aaaaaacccat ggttatctct ctctaataaaa taataaacgaa tc 42

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

<220>
 <223> синтетична послідовність

 <400> 26
 aataaaggta ccgtccagaa tttctccat tga 33

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

<220>
 <223> синтетична послідовність

 <400> 27
 aataaaccat ggtcttcact atccaaagct ctca 34

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

<220>
 <223> синтетична послідовність

 <400> 28
 ttttttgga ccgtctactt tcattacagt gactctg 37

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

<220>
 <223> синтетична послідовність

 <400> 29
 ttttttccat gggtatattt tacctgcaac acaattcaa 39

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

<220>
 <223> синтетична послідовність

 <400> 30
 ttttatgga cccactcgaa tactgcatgc aa 32

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

<220>
 <223> синтетична послідовність

 <400> 31

ttttatccat ggttatgtag 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 ggttatctta ctgtttttaa ccaaaaaata aaat 44

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

<220>
<223> синтетична послідовність

<400> 34
tttttaggta ccactcttagg gtttcgag atctca 36

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

<220>
<223> синтетична послідовність

<400> 35
ttttttccat 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 catcccaaag ctс 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 ggttatttaa tccacaagcc ttgcctc 37

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

<220>
 <223> синтетична послідовність

<400> 42
 tttttaccat ggtgtcggag aagtgggcg 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
 ttttatccat ggtagtatct acataccaat catacaaatg 40

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

<220>
 <223> синтетична послідовність

<400> 46
 ttttttggtta cctttcacga tttggaattt ga 32

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

<220>
 <223> синтетична послідовність

<400> 47

ttttttccat ggtctacaac attaaaacga ccatta 36

<210> 48
<211> 32
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 48
tatataggta ccagggtttc gtttttggtt ca 32

<210> 49
<211> 35
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 49
tatataccat gggtatctcc tgctcaaaga aacca 35

<210> 50
<211> 33
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 50
tttataggta ccagaagctc atttcttcga tac 33

<210> 51
<211> 32
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 51
tttataccat ggtctctgcg caaaaattca cc 32

<210> 52
<211> 31
<212> ДНК
<213> штучна

<220>
<223> синтетична послідовність

<400> 52
tatattggtta cctctaaaaa tacagggcac c 31

<210> 53

<211> 35
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 53
 tatattccat ggttactctt 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 ggtttcttct 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 ggttaaattt tacctgtcat caaaaacaac a 41

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

<220>
<223> синтетична послідовність

<400> 58
ggggacaact ttgtatagaa aagttggcca catcatgttt agacttatc 49

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

<220>
<223> синтетична послідовність

<400> 59
ggggactgct tttttgtaca aacttggtta ccttttatat ttatatatag 50

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

<220>
<223> синтетична послідовність

<400> 60
ggtaccgagg gatcctctag catatgctcg aggcggccgc agatatacaga tctgggtcgac 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 gcgggaaacg acaatcagat ctagtaggaa acagctatga ccatgattac 180
gccaaagctat cgattacgcc aagctatcaa ctttgtatag aaaagttggc cacatcatgt 240
ttagacttat ctccataaag aaaaccactc atcaaagcta atacaaaagc tctagtgtga 300
cacttcaacg tcacttcac aggatcagca ggtaaattcc gaaaattctc acgcagccac 360
gctaaggaca catgagaacc atgaagatcc ttgggacctg gcctatgacc aagcaaattc 420
tcacataaat cagcccagtt gtattttgta ctaccagtta ctgcaggctc atcgacacgt 480
agacccaaca aaatattcac atcttgtaaa gtcacagtga tctctccagc aggaagatga 540
aaagtatgcg tttcgggtct ccatctctcc accaaagctg ttatcagagc ataataaagt 600
tgtataaagg caaccttgta aactccatat agaccaaaact ctatcaactt ttgacacacg 660

agaggatcca gaggccaatc tcgcatcccc aataacttgt gccgacatgt cagttcacga	720
ggaggaacct gaatgtgaag tataacggta aaaaggaaat aattaaaaca acggaagcaa	780
aacaagaaac aagatgaaat gagaaactag 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 ttcataatth gaacacaaga aatcctaaac tggtagaaa gaaaaattgc	1080
tcaacaaaaa aatcaagcta attactcgta taaaagaca cgaagaacta atacaagaaa	1140
caagaacaa caaaccacaa agagattgaa ccaatccaaa ttcgacaaca taaaccaagt	1200
gtgtgggtga ttggaatcag aggacgtacc aaagaaaagc gtccgtgatc aacaaaaacc	1260
aaaaagagac gtttgaaata aaccagagga agacgaagaa taattaagca aagagaagcg	1320
ttaagcggga gcgagaaagg aaacgagaga aagagagagc ttccagatcc gacagaagtt	1380
ttcggcttct tctttttctg ttaagaactt ctgatcctcc taggtctgtc cgaagaacta	1440
atctttttga ggtaacgacg ccgtttttct caaaacatgg gcccatthac catagtctcg	1500
goccaaacga aacttaatac gacaatgttt ggggtgtaaac gcaaagattt tgtcgttat	1560
cacaagtaaa aaaataaata caaacacttg agtctctcta gacatcgtgc atgccttag	1620
ctttaagttt tttctcgaac caaaagagtt attttatttg aactttgaag attatacgaa	1680
gacacgtggc gtgaacccaa ttcataacaa cgccacgcta tactcttttg catgcacctc	1740
aatttgaaca tcatcaagtc tctctctctt tttctgactt tgatccacga acctaacag	1800
cttgcatct ctatttaatc ggtcctcgac gcaacttcaa cttctactac atccattcac	1860
atcaaataca tacagaaagt tttttctata tataaatata aaaggtaaac aagtttgtac	1920
aaaaaagcag gctggtacct ggtgcttaaa cactctgggt agttctagta cttctgctat	1980
gatcgatctc attaccattt cttaaatttc tctccctaaa tattccgagt tcttgatttt	2040
tgataacttc aggttttctc tttttgataa atctgggtct tccatttttt tttttttgtg	2100
gttaatttag tttcctatgt tcttcgattg tattatgcat gatctgtgtt tggattctgt	2160
tagattatgt attggtgaat atgtatgtgt ttttgcatgt ctgggttttg tcttaaaaat	2220
gttcaaatac gatgatttga ttgaagcttt tttagtgttg gtttgattct tctcaaaact	2280
actgttaatt tactatcatg ttttccaact ttgattcatg atgacacttt tgttctgctt	2340
tgttataaaa ttttggtttg tttgattttg taattatagt gtaattttgt taggaatgaa	2400
catgttttaa tactctgttt tcgatttgc acacattcga attattaatc gataatttaa	2460
ctgaaaattc atggttctag atcttggtgt catcagatta tttgtttcga taattcatca	2520
aatatgtagt ccttttgctg atttgcgact gtttcatttt ttctcaaaat tgttttttgt	2580

taagttttatc taacagttat cgttgtcaaa agtctctttc attttgcaaa atcttctttt	2640
tttttttggt tgtaactttg ttttttaagc tacacattta gtctgtaaaa tagcatcgag	2700
gaacagttgt cttagtagac ttgcatgttc ttgtaacttc tatttgtttc agtttggtga	2760
tgactgcttt gatttttag gtcaaaccat ggaagacgcc aaaaacataa agaaaggccc	2820
ggcgccattc tatccgctgg aagatggaac cgctggagag caactgcata aggctatgaa	2880
gagatacgcc ctggttcctg gaacaattgc ttttacagat gcacatatcg aggtggacat	2940
cacttacgct gagtacttcg aaatgtccgt tcggttgga gaagctatga aacgatatgg	3000
gctgaataca aatcacagaa tcgtcgtatg cagtgaaaac tctcttcaat tctttatgcc	3060
ggtgttgggc gcgttattta tcggagttgc agttgcgcc gcgaacgaca tttataatga	3120
acgtgaattg ctcaacagta tgggcatttc gcagcctacc gtggtgttcg tttccaaaaa	3180
ggggttgcaa aaaattttga acgtgcaaaa aaagctccca atcatccaaa aaattattat	3240
catggattct aaaacggatt accagggatt tcagtcgatg tacacgttcg tcacatctca	3300
tctacctccc ggttttaatg aatacgattt tgtgccagag tccttcgata gggacaagac	3360
aattgcactg atcatgaact cctctggatc tactggtctg cctaaagggtg tcgctctgcc	3420
tcatagaact gcctgcgtga gattctcgca tgccagagat cctatttttg gcaatcaa	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 gatttatcta atttacacga	3720
aattgcttct ggtggcgctc ccctctctaa ggaagtcggg gaagcggttg ccaagagggt	3780
ccatctgcca ggtatcaggc aaggatatgg gctcactgag actacatcag ctattctgat	3840
tacaccogag ggggatgata aaccgggcgc ggtcggtaaa gttgttccat tttttgaagc	3900
gaaggttggt gatctggata ccgggaaaac gctgggcgtt aatcaaagag gcgaactgtg	3960
tgtgagaggt cctatgatta tgtccggtta tgtaaacaaat ccggaagcga ccaacgcctt	4020
gattgacaag gatggatggc tacattcttg agacatagct tactgggacg aagacgaaca	4080
cttcttcacg gttgaccgcc tgaagtctct gattaagtac aaaggctatc aggtggctcc	4140
cgctgaattg gaatccatct tgctccaaca cccaacatc ttcgacgcag gtgtcgcagg	4200
tcttcccgac gatgacgcc gtgaacttcc cgccgcggtt gttgttttg agcacggaaa	4260
gacgatgacg gaaaaagaga tcgtggatta cgtcgcagc caagtaacaa ccgcgaaaaa	4320
gttgcgcgga ggagttgtgt ttgtggacga agtaccgaaa ggtcttaccg gaaaactcga	4380
cgcaagaaaa atcagagaga tcctcataaa ggccaagaag ggcggaaaga tcgccgtgta	4440

actcgagcat atgggctcga atttccccga tcgttcaaac atttggcaat aaagtttctt	4500
aagattgaat cctgttgccg gtcttgcgat gattatcata taatttctgt tgaattacgt	4560
taagcatgta ataattaaca tgtaatgcat gacgttatTT atgagatggg tttttatgat	4620
tagagtcccg caattataca tttaatacgc gatagaaaac aaaatatagc gcgcaaacta	4680
ggataaatta tcgcgcgcgg tgatcatctat gttactagat cgggaattca agcttggcgt	4740
aatcatggac ccagctttct tgtacaaagt ggggtaccg gggatcctct agcatatgct	4800
cgaggcgcc gcagatatca gatctggtcg acggcatgca agcttggcgt aatcatggca	4860
actttattat acatagtTga taattcactg gccggataat tTactggccg tcgttttaca	4920
acgactcagg atcctgtcaa acactgatag tttaaactga aggcgggaaa cgacaatctg	4980
atcatgagcg gagaattaag ggagtcacgt tatgaccccc gccgatgacg cgggacaagc	5040
cgttttacgt ttggaactga cagaaccgca acgttgaagg agccactcag ccgcgggttt	5100
ctggagttta atgagctaag cacatacgtc agaaaccatt attgcgcgtt caaaagtcgc	5160
ctaaggTcac tatcagctag caaatatttc ttgtcaaaaa tgctccactg acgttccata	5220
aattccccctc ggtatccaat tagagtctca tattcactct caatccaaat aatctgcacc	5280
ggatctggat cgtttcgcat gattgaacaa gatggattgc acgcaggttc tccggccgct	5340
tgggtggaga ggctattcgg ctatgactgg gcacaacaga caatcggtg ctctgatgcc	5400
gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt ttgtcaagac cgacctgtcc	5460
ggtgcctga atgaactgca ggacgaggca gcgcggctat cgtggctggc cagcagggc	5520
gttccttgcg cagctgtgct cgacgttgtc actgaagcgg gaagggactg gctgctattg	5580
ggcgaagtgc cggggcagga tctcctgtca tctcaccttg ctctgccga gaaagtatcc	5640
atcatggctg atgcaatgcg gcggctgcat acgcttgatc cggctacctg cccattcgac	5700
caccaagcga aacatcgcat cgagcgagca cgtactcgga tggaagccgg tcttgtcgat	5760
caggatgatc tggacgaaga gcatcagggg ctcgcgccag ccgaactgtt cgccaggctc	5820
aaggcgcgca tgcccgcgg cgaggatctc gtcgtgaccc atggcgatgc ctgcttgccg	5880
aatatcatgg tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg	5940
gcggaccgct atcaggacat agcgttggct acccgtgata ttgctgaaga gcttggcggc	6000
gaatgggctg accgcttct cgtgctttac ggtatcgccg ctcccgattc gcagcgcatc	6060
gccttctatc gccttcttga cgagtcttct tgagcgggac ccaagctcta gatcttgcgtg	6120
cgttcggata ttttcgtgga gttcccgcca cagaccgga tgatccccga tcgttcaaac	6180
atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata	6240
taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttatTT	6300
atgagatggg tttttatgat tagagtcccg caattataca tttaatacgc gatagaaaac	6360

aaaatatagc gcgcaaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat	6420
cgggcctcct gtcaagctct gcttggaat aattgtcatt agattgtttt tatgcataga	6480
tgactcga atcagccaat tttagacaag tatcaaacgg atgttaattc agtacattaa	6540
agacgtccgc aatgtgttat taagttgtct aagcgtcaat ttgtttacac cacaatatat	6600
cctgccacca gccagccaac agtccccga cgggcagctc ggcacaaaat caccacgcgt	6660
taccaccacg ccggccggcc gcatggtgtt gaccgtgttc gccggcattg ccgagttcga	6720
gcgttcccta atcatcgacc gcaccggag cgggcgcgag gccgccaagg ccgagggcgt	6780
gaagtttggc ccccgcccta ccctacccc ggcacagatc gcgcacgccc gcgagctgat	6840
cgaccaggaa ggccgcaccg tgaaagaggc ggctgcactg cttggcgtgc atcgctcgac	6900
cctgtaccgc gcacttgagc gcagcgagga agtgacgccc accgaggcca ggcggcgcgg	6960
tgcttccgt gaggacgcat tgaccgaggc cgacgccctg gcggccgccc agaataaacg	7020
ccaagaggaa caagcatgaa accgcaccag gacggccagg acgaaccgtt tttcattacc	7080
gaagagatcg aggcggagat gatcgcgccc gggtagtgt tcgagccgcc cgcgcacgtc	7140
tcaaccgtgc ggctgcatga aatcctggcc ggtttgtctg atgccaaagt ggcggcctgg	7200
ccggccagct tggccgctga agaaaccgag cgccgccgtc taaaaaggtg atgtgtat	7260
gagtaaaaca gcttgctca tgcggtcgtc gcgtatatga tgcgatgagt aaataaaca	7320
atacgcaagg ggaacgcatg aaggttatcg ctgtacttaa ccagaaaggc gggtcaggca	7380
agacgaccat cgcaaccat ctagcccgcg ccctgcaact cgccggggcc gatgttctgt	7440
tagtcgattc cgatccccag ggcagtgcc gcgattgggc ggccgtgcgg gaagatcaac	7500
cgctaaccgt tgcggcatc gaccgcccga cgattgaccg cgacgtgaag gccatcgccc	7560
ggcgcgactt cgtagtgatc gacggagcgc ccaggcggc ggacttggct gtgtccgcga	7620
tcaaggcagc cgacttcgtg ctgattccgg tgcagccaag cccttacgac atatgggcca	7680
ccgccgacct ggtggagctg gttaagcagc gcattgaggt cacggtgga aggtacaag	7740
cggcctttgt cgtgtcgcgg gcgatcaaag gcacgcgcat cggcggtagg gttgccgagg	7800
cgctggccgg gtacgagctg ccattcttg agtcccgat cacgcagcgc gtgagctacc	7860
caggcactgc cgccggcgcc acaaccgttc ttgaatcaga acccgagggc gacgtgccc	7920
gcgaggcca ggcgtggcc gctgaaatta aatcaaaact catttgagtt aatgaggtaa	7980
agagaaaatg agcaaaagca caaacacgct aagtgccggc cgtccgagcg cacgcagcag	8040
caaggctgca acgttggcca gcctggcaga cacgccagcc atgaagcggg tcaactttca	8100
gttgccggcg gaggatcaca ccaagctgaa gatgtacgcg gtacgccaag gcaagaccat	8160
taccgagctg ctatctgaat acatcgcgca gctaccagag taaatgagca aatgaataaa	8220
tgagtagatg aatttttagcg gctaaaggag gcggcatgga aaatcaagaa caaccaggca	8280

ccgacgccgt ggaatgcccc atgtgtggag gaacggggcg ttggccaggc gtaagcggct	8340
gggttgtctg ccggccctgc aatggcactg gaacccccaa gcccaggaa tcggcgtgag	8400
cggtcgcaaa ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga	8460
gaagttgaag gccgcgcagg ccgcccagcg gcaacgcata gaggcagaag cacgccccgg	8520
tgaatcgtgg caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc	8580
cggtgcgcg tcgattagga agccgccccaa gggcgacgag caaccagatt ttttcgttcc	8640
gatgctctat gacgtgggca cccgcgatag tcgcagcata atggacgtgg cgtttttccg	8700
tctgtcgaag cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca	8760
cgtagagggt tccgcagggc cgcccgccat ggccagtgtg tgggattacg acctggtact	8820
gatggcgggt tcccatactaa ccgaatccat gaaccgatac cggaaggga agggagacaa	8880
gcccgccgc gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga	8940
tggcggaag cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt	9000
tgccatgcag cgtacgaaga aggccaagaa cgccgcctg gtgacggtat ccgagggtga	9060
agccttgatt agccgtaca agatcgtaaa gagcgaaacc gggcgcccg agtacatcga	9120
gatcgagcta gctgattgga tgtaccgca gatcacagaa ggcaagaacc cggacgtgct	9180
gacggttcac cccgattact ttttgatcga tcccggcata gccggttttc tctaccgcct	9240
ggcagccgc gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg	9300
cagtggcagc gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc	9360
aaatgacctg ccggagtacg atttgaagga ggaggcggg caggctggcc cgatocctagt	9420
catgcgctac cgcaacctga tcgaggcga agcatccgcc ggttcctaata gtacggagca	9480
gatgctaggg caaattgccc tagcagggga aaaaggctcga aaaggctctct ttctgtgga	9540
tagcacgtac attgggaacc caaagccgta cattgggaac cggaaccgt acattgggaa	9600
cccaaagccg tacattggga accggtcaca catgtaagt actgatataa aagagaaaaa	9660
aggcgatttt tccgcctaaa actctttaaa acttattaaa actcttaaaa ccgcctggc	9720
ctgtgcataa ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg	9780
gtcgtgcgc tccctacgcc ccgcgccttc gcgtcggcct atcgcggccg ctggccgctc	9840
aaaaatggct ggctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc	9900
actcgaccgc cggcgcccac atcaaggcac cctgcctcgc gcgtttcggg gatgacgggtg	9960
aaaacctctg acacatgcag ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg	10020
ggagcagaca agcccgtcag ggcgctcag cgggtgttgg cgggtgtcgg ggcgcagcca	10080
tgaccagtc acgtacgat agcggagtgt atactggctt aactatgcgg catcagagca	10140
gattgtactg agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa	10200

ataccgcatc	aggcgctctt	ccgcttcctc	gtcactgac	tcgctgcgct	cggtcgttcg	10260
gctgcggcga	gcggtatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	10320
ggataacgca	ggaaagaaca	tgtgagcaaa	agggcagcaa	aaggccagga	accgtaaaaa	10380
ggccgcgttg	ctggcgtttt	tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	10440
acgctcaagt	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttcccc	10500
tggaagctcc	ctcgtgcgct	ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	10560
ctttctccct	tcgggaagcg	tggcgctttc	tcatagctca	cgctgtaggt	atctcagttc	10620
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	cccccggtc	agcccgaccg	10680
ctgcgcctta	tccggtaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	10740
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	10800
gttcttgaag	tggtggccta	actacggcta	cactagaagg	acagtatttg	gtatctgcgc	10860
tctgctgaag	ccagttacct	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaaac	10920
caccgctggt	agcgggtggt	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	10980
atctcaagaa	gaccttttga	tcttttctac	ggggtctgac	gtcagtgga	acgaaaactc	11040
acgttaaggg	attttgggtca	tgcatgatat	atctcccaat	ttgtgtaggg	cttattatgc	11100
acgcttaaaa	ataataaaaag	cagacttgac	ctgatagttt	ggctgtgagc	aattatgtgc	11160
ttagtgcatc	taacgcttga	gttaagccgc	gccgcgaagc	ggcgtcggct	tgaacgaatt	11220
tctagctaga	cattatttgc	cgactacctt	ggtgatctcg	cctttcacgt	agtggacaaa	11280
ttcttccaac	tgatctgcgc	gcgaggccaa	gcgatcttct	tcttgtccaa	gataagcctg	11340
tctagcttca	agtatgacgg	gctgatactg	ggccggcagg	cgctccattg	cccagtcggc	11400
agcgacatcc	ttcgggcgca	ttttgccggt	tactgcgctg	taccaaagtc	gggacaacgt	11460
aagcactaca	tttcgctcat	cgccagccca	gtcggggcggc	gagttccata	gcgttaaggt	11520
ttcatttagc	gcctcaaata	gacctgttgc	aggaaccgga	tcaaagagtt	cctccgccgc	11580
tggacctacc	aaggcaacgc	tatgttctct	tgcttttgtc	agcaagatag	ccagatcaat	11640
gtcgatcgtg	gctggctcga	agatacctgc	aagaatgtca	ttgcgctgcc	attctccaaa	11700
ttgcagttcg	cgcttagctg	gataacgcca	cggaatgatg	tcgtcgtgca	caacaatggt	11760
gacttctaca	gcgcggagaa	tctcgctctc	tccaggggaa	gccgaagttt	ccaaaaggtc	11820
gttgatcaaa	gctcgccgcg	ttgtttcatc	aagccttacg	gtcaccgtaa	ccagcaaadc	11880
aatatcactg	tgtggcttca	ggccgccatc	cactgcggag	ccgtacaaat	gtacggccag	11940
caacgtcggg	tcgagatggc	gctcgatgac	gccaaactacc	tctgatagtt	gagtcgatac	12000
ttcggcgatc	accgcttccc	ccatgatgtt	taactttgtt	ttagggcgac	tgccctgctg	12060
cgtaacatcg	ttgctgctcc	ataacatcaa	acatcgaccc	acggcgtaac	gcgcttgctg	12120

cttggatgcc cgaggcatag actgtacccc aaaaaaacag tcataacaag ccatgaaaac 12180
cgccactgcg ttccatg 12197

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

<220>
<223> векторна послідовність

<400> 62
gacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatac cgattattct 60
aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt 120
taaaactgaag gcgggaaacg acaatcagat ctagtaggaa acagctatga ccatgattac 180
gccaaagctat cgattacgcc aagctatcaa ctttgtatag aaaagttggc cacatcatgt 240
ttagacttat ctccataaag aaaaccactc atcaaagcta atacaaaagc tctagtgtga 300
cacttcaacg tcacttcac aggatcagca ggtaaattcc gaaaattctc acgcagccac 360
gctaaggaca catgagaacc atgaagatcc ttgggacctg gcctatgacc aagcaaattcc 420
tcacataaat cagcccagtt gtattttgta ctaccagtta ctgcagggtc atcgacacgt 480
agacccaaca aaatattcac atcttgtaaa gtcacagtga tctctccagc aggaagatga 540
aaagtatgcg ttctgggtct ccatctctcc accaaagctg ttatcagagc ataataaagt 600
tgtataaagg caaccttgta aactccatat agaccaaact ctatcaactt ttgacacacg 660
agaggatcca gaggccaatc tcgcatcccc aataacttgt gccgacatgt cagttcacga 720
ggaggaacct gaatgtgaag tataacggta aaaaggaaat aattaaaaca acggaagcaa 780
aacaagaaac aagatgaaat gagaaactag taacacacct catcttcca tatagcagct 840
gatctatgct catgttgcca caccaatata gattgatcaa ctggaccagg atccaaatca 900
aagtttaata gactttgcac ctccatctat ataatatatc acaggacaat aaacacaatg 960
atcagtgatt atacaacatc aaagaaaact tgtaattctg ggaatataac tgagaaatga 1020
gaattaaaga ttcataatth gaacacaaga aatcctaaac tggtagcga gaaaaattgc 1080
tcaacaaaaa aatcaagcta attactcgta tacaagaca cgaagaacta atacaagaaa 1140
caagaaacaa caaaccacaa agagattgaa ccaatccaaa ttcgacaaca taaaccaagt 1200
gtgtgggtga ttggaatcag aggacgtacc aaagaaaagc gtccgtgatc 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 caaaacatgg gccattaac catagtctcg 1500

gccccaaacga aacttaatac gacaatgttt ggggtgtaaac gcaaagattt tgtcgattat	1560
cacaagtaaa aaaataaata caaacacttg agtctctcta gacatcgtgc atcgccttag	1620
ctttaagttt tttctcgaaa caaaagagtt attttatttg aactttgaag attatacgaa	1680
gacacgtggc gtgaacccaa ttcataacaa cgccacgcta tactcttttg catgcacctc	1740
aatttgaaca tcatcaagtc tctctctctt tttctgactt tgatccacga acctaaccag	1800
cttgcgatct ctattttaatc ggtcctcgac gcaacttcaa cttctactac atccattcac	1860
atcaaatacaa tacagaaagt tttttctata tataaatata aaaggtaaac aagtttgtac	1920
aaaaaagcag gctggtacct ggtgcttaaa cactctgggtg agttctagta cttctgctat	1980
gatcgatctc attaccattt cttaaatttc tctccctaaa tattccgagt tcttgatttt	2040
tgataacttc aggttttctc tttttgataa atctgggtctt tccatttttt tttttttgtg	2100
gttaatttag tttcctatgt tcttcgattg tattatgcat gatctgtgtt tggattctgt	2160
tagattatgt attggtgaat atgtatgtgt ttttgcatgt ctggtttttg tcttaaaaat	2220
gttcaaactc gatgatttga ttgaagcttt tttagtgttg gtttgattct tctcaaaact	2280
actgttaatt tactatcatg ttttccaact ttgattcatg atgacacttt tgttctgctt	2340
tgttataaaa ttttggttg tttgattttg taattatagt gtaattttgt taggaatgaa	2400
catgttttaa tactctgttt tcgatttgtc acacattcga attattaatc gataatttaa	2460
ctgaaaattc atggttctag atcttggtgt catcagatta tttgtttcga taattcatca	2520
aatatgtagt ctttttgctg atttgcgact gtttcatttt ttctcaaaat tgttttttgt	2580
taagtttatc taacagttat cgttgtcaaa agtctctttc attttgcaaa atcttctttt	2640
tttttttgtt tgtaactttg ttttttaagc tacacattta gtctgtaaaa tagcatcgag	2700
gaacagttgt cttagtagac ttgcatgttc ttgtaacttc tatttgtttc agtttgttga	2760
tgactgcttt gatttttag gtcaaaccat ggaagacgcc aaaaacataa agaaaggccc	2820
ggcgccattc tatccgctgg aagatggaac cgctggagag caactgcata aggctatgaa	2880
gagatacgcc ctggttcctg gaacaattgc ttttacagat gcacatatcg aggtggacat	2940
cacttacgct gagtacttcg aaatgtccgt tcggttgga gaagctatga aacgatatgg	3000
gctgaataca aatcacagaa tcgtcgtatg cagtgaaaac tctcttcaat tctttatgcc	3060
ggtgttgggc gcgttattta tcggagttgc agttgcgccc gcgaacgaca tttataatga	3120
acgtgaattg ctcaacagta tgggcatttc gcagcctacc gtggtgttcg tttccaaaaa	3180
ggggttgcaa aaaattttga acgtgcaaaa aaagctccca atcatccaaa aaattattat	3240
catggattct aaaacggatt accagggatt tcagtcgatg tacacgttcg tcacatctca	3300
tctacctccc ggttttaatg aatacgattt tgtgccagag tccttcgata gggacaagac	3360
aattgcactg atcatgaact cctctggatc tactgggtctg cctaaagggtg tcgctctgcc	3420

tcatagaact	gcctgcgtga	gattctcgca	tgccagagat	cctatTTTTg	gcaatcaa	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	gatttatcta	atttacacga	3720
aattgcttct	ggtggcgctc	ccctctctaa	ggaagtcggg	gaagcgggtg	ccaagagggt	3780
ccatctgcca	ggtatcaggc	aaggatatgg	gctcactgag	actacatcag	ctattctgat	3840
tacacccgag	ggggatgata	aaccggggcg	ggtcggtaaa	gttgttccat	tttttgaagc	3900
gaaggtttgt	gatctggata	ccgggaaaa	gctgggcgtt	aatcaaagag	gcgaactgtg	3960
tgtgagaggt	cctatgatta	tgtccggtta	tgtaaacaat	ccggaagcga	ccaacgcctt	4020
gattgacaag	gatggatggc	tacattctgg	agacatagct	tactgggacg	aagacgaaca	4080
cttcttcctc	gttgaccgcc	tgaagtctct	gattaagtac	aaaggctatc	aggtggctcc	4140
cgctgaattg	gaatccatct	tgctccaaca	ccccaacatc	ttcgacgcag	gtgtcgcagg	4200
tcttcccgac	gatgacgccg	gtgaacttcc	cgccgccgtt	gttgTTTTg	agcacggaaa	4260
gacgatgacg	gaaaaagaga	tcgtggatta	cgctgccagt	caagtaacaa	ccgcgaaaaa	4320
gttgcgcgga	ggagtttgt	ttgtggacga	agtaccgaaa	ggtcttaccg	gaaaactcga	4380
cgcaagaaaa	atcagagaga	tcctcataaa	ggccaagaag	ggcggaaaga	tcgccgtgta	4440
actcgagcat	atgggctcga	atttccccga	tcgttcaaac	atttggaat	aaagtttctt	4500
aagattgaat	cctgttgccg	gtcttgccat	gattatcata	taatttctgt	tgaattacgt	4560
taagcatgta	ataattaaca	tgtaatgcat	gacgttattt	atgagatggg	tttttatgat	4620
tagagtcccg	caattataca	tttaatacgc	gatagaaaac	aaaatatagc	gcgcaaacta	4680
ggataaatta	tcgcgcgcgg	tgatcatctat	gttactagat	cggaattca	agcttggcgt	4740
aatcatggac	ccagctttct	tgtacaaagt	ggggtaccaa	ttcgaatcca	aaaattacgg	4800
atatgaatat	aggcatatcc	gtatccgaat	tatccgtttg	acagctagca	acgattgtac	4860
aattgcttct	ttaaaaaagg	aagaaagaaa	gaaagaaaag	aatcaacatc	agcgttaaca	4920
aacggccccc	ttacggccca	aacggtcata	tagagtaacg	gcgttaagcg	ttgaaagact	4980
cctatcgaaa	tacgtaaccg	caaacgtgtc	atagtcagat	cccctcttcc	ttcacgcctt	5040
caaacacaaa	aataatcttc	tacagcctat	atatacaacc	cccccttcta	tctctccttt	5100
ctcacaattc	atcatctttc	tttctctacc	cccaatttta	agaaatcctc	tcttctcctc	5160
ttcattttca	aggtaaactt	ctctctctct	ctctctctct	gttatctctt	gttttaatta	5220
ggtatgtatt	attgctagtt	tgtaaatctg	cttatcttat	gtatgcctta	tgtgaatatc	5280
tttatcttgt	tcatctcctc	cgtttagaag	ctataaaattt	gttgatttga	ctgtgtatct	5340

acacgtgggt	atgtttatat	ctaatacagat	atgaatttct	tcatattgtt	gcgtttgtgt	5400
gtaccaatcc	gaaatcggtg	atTTTTTTca	tttaatcgtg	tagctaattg	tacgtataca	5460
tatggatcta	cgtatcaatt	gttcatctgt	ttgtgtttgt	atgtatacag	atctgaaaac	5520
atcacttctc	tcatctgatt	gtgttggtac	atacatagat	atagatctgt	tatatcattt	5580
TTTTtattaa	ttgtgtatat	atatatgtgc	atagatctgg	attacatgat	tgtgattatt	5640
tacatgattt	tgttattttac	gtatgtatat	atgtagatct	ggactTTTTg	gagttgttga	5700
cttgattgta	tttgtgtgtg	tatatgtgtg	ttctgatctt	gatatgttat	gtatgtgcag	5760
ctgaaccatg	gcggcggcaa	caacaacaac	aacaacatct	tcttcgatct	ccttctccac	5820
caaaccatct	ccttctctct	ccaaatcacc	attaccaatc	tccagattct	ccctcccatt	5880
ctccctaaac	cccaacaaat	catcctcctc	ctcccgcgcg	cgcggtatca	aatccagctc	5940
tccctcctcc	atctccgccg	tgtcacaacac	aaccaccaat	gtcacaacca	ctccctctcc	6000
aaccaaacct	accaaaccgg	aaacattcat	ctcccgattc	gctccagatc	aaccccgcaa	6060
aggcgctgat	atcctcgtcg	aagcttttaga	acgtcaaggc	gtagaaaccg	tattcgctta	6120
ccctggaggt	acatcaatgg	agattcacca	agccttaacc	cgctcttcct	caatccgtaa	6180
cgctcttcct	cgtcacgaac	aaggaggtgt	attcgagca	gaaggatacg	ctcgatcctc	6240
aggtaaacca	ggtatctgta	tagccacttc	aggccccga	gctacaaatc	tcgttagcgg	6300
attagccgat	gcgttggttag	atagtgttcc	tcttgtagca	atcacaggac	aagtccctcg	6360
tcgtatgatt	ggtacagatg	cgtttcaaga	gactccgatt	gttgaggtaa	cgcgttcgat	6420
tacgaagcat	aactatcttg	tgatggatgt	tgaagatata	cctaggatta	ttgaggaagc	6480
tttcttttta	gctacttctg	gtagacctgg	acctgttttg	gttgatgttc	ctaaagatat	6540
tcaacaacag	cttgcgattc	ctaattggga	acaggctatg	agattacctg	gttatatgtc	6600
taggatgcct	aaacctccgg	aagattctca	tttgagcag	attgttaggt	tgatttctga	6660
gtctaagaag	cctgtgttgt	atgttggtgg	tggttgtttg	aattctagcg	atgaattggg	6720
taggtttgtt	gagcttacgg	ggatccctgt	tgcgagtacg	ttgatggggc	tgggatctta	6780
tccttgatgat	gatgagttgt	cgttacatat	gcttggaatg	catgggactg	tgtatgcaaa	6840
ttacgctgtg	gagcatagtg	atTTgttgtt	ggcgTTTTggg	gtaaggTTTTg	atgatcggtg	6900
cacgggtaag	cttgaggctt	ttgctagtag	ggctaagatt	gttcatattg	atattgactc	6960
ggctgagatt	gggaagaata	agactcctca	tgtgtctgtg	tgtgggtgatg	ttaagctggc	7020
tttgcaaggg	atgaataagg	ttcttgagaa	ccgagcggag	gagcttaagc	ttgatttttg	7080
agtttgaggg	aatgagttga	acgtacagaa	acagaagttt	ccgttgagct	ttaagacggt	7140
tggggaagct	attcctccac	agtatgcgat	taaggtcctt	gatgagttga	ctgatggaaa	7200
agccataata	agtactggtg	tcgggcaaca	tcaaatgtgg	gcggcgagct	tctacaatta	7260

caagaaacca aggcagtggc tatcatcagg aggccttggg gctatgggat ttggacttcc	7320
tgctgcgatt ggagcgtctg ttgctaaccg tgatgcgata gttgtggata ttgacggaga	7380
tggaagcttt ataatgaatg tgcaagagct agccactatt cgtgtagaga atcttccagt	7440
gaaggtactt ttattaaaca accagcatct tggcatgggt atgcaatggg aagatcgggt	7500
ctacaaagct aaccgagctc acacatttct cggggatccg gctcaggagg acgagatatt	7560
cccgaacatg ttgctgtttg cagcagcttg cgggattcca gcggcgaggg tgacaaagaa	7620
agcagatctc cgagaagcta ttcagacaat gctggataca ccaggacctt acctgttggg	7680
tgtgatttgt ccgcaccaag aacatgtgtt gccgatgatc ccgaatggtg gcactttcaa	7740
cgatgtcata acggaaggag atggccggat taaatactga gagatgaaac cggtgattat	7800
cagaaccttt tatggtcttt gtatgcatat ggtaaaaaaa cttagtttgc aatttcctgt	7860
ttgttttggg aatttgagtt tcttttagtt gttgatctgc ctgctttttg gtttacgtca	7920
gactactact gctgttggtt tttggtttcc tttctttcat tttataaata aataatccgg	7980
ttcggtttac tcttgtgac tggctcagtt tggttattgc gaaatgcgaa tggtaaattg	8040
agtaattgaa attcgttatt agggttctaa gctgttttaa cagtcactgg gttaatatct	8100
ctcgaatctt gcatggaaaa tgctcttacc attggttttt aattgaaatg tgctcatatg	8160
ggccgtgggt tccaaattaa ataaaactac gatgtcatcg agaagtaaaa tcaactgtgt	8220
ccacattatc agttttgtgt atacgatgaa atagggtaat tcaaaatcta gcttgatatg	8280
ccttttgggt cattttaacc ttctgtaaac attttttcag attttgaaca agtaaatcca	8340
aaaaaaaaa aaaaaaatct caactcaaca ctaaattatt ttaatgtata aaagatgctt	8400
aaaacatttg gcttaaaaga aagaagctaa aaacatagag aactcttgta aattgaagta	8460
tgaaaatata ctgaattggg tattatatga atttttctga tttaggattc acatgatcca	8520
aaaaggaaat ccagaagcac taatcagaca ttggaagtag gaatatttca aaaagttttt	8580
tttttttaag taagtgacaa aagcttttaa aaaatagaaa agaaactagt attaaagttg	8640
taaatttaat aaacaaaaga aattttttat attttttcat ttctttttcc agcatgaggt	8700
tatgatggca ggatgtggat ttcatttttt tctttttgat agccttttaa ttgatctatt	8760
ataattgacg aaaaaatatt agttaattat agatatattt taggtagtat tagcaattta	8820
cacttccaaa agactatgta agttgtaaat atgatgcgtt gatctcttca tcattcaatg	8880
gttagtcaaa aaaataaaag cttaactagt aaactaaagt agtcaaaaat tgtactttag	8940
tttaaaatat tacatgaata atccaaaacg acatttatgt gaaacaaaaa caatatctag	9000
agtcgactta attaaactag tggcgcgcca attgactagt aggcctatcg attaatgaag	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	caatttggtt	acaccacaat	9300
atatactgcc	accagccagc	caacagctcc	ccgaccggca	gctcggcaca	aaatcaccac	9360
gcgttaccac	cacgccggcc	ggccgcatgg	tggtgaccgt	gttcgccggc	attgccgagt	9420
tcgagcgctt	cctaatactc	gaccgcaccc	ggagcgggcg	cgaggccgcc	aaggcccag	9480
gcgtgaagtt	tggcccccg	cctaccctca	ccccggcaca	gatcgcgcac	gcccgcgagc	9540
tgatcgacca	ggaaggccgc	accgtgaaa	aggcggctgc	actgcttggc	gtgcatcgct	9600
cgaccctgta	ccgcgcactt	gagcgagcg	aggaagtgc	gccaccgag	gccaggcggc	9660
gcggtgcctt	ccgtgaggac	gcattgaccg	aggccgacgc	cctggcggcc	gccgagaatg	9720
aacgccaaga	ggaacaagca	tgaaccgca	ccaggacggc	caggacgaac	cgtttttcat	9780
taccgaagag	atcgaggcgg	agatgatcgc	ggccgggtac	gtgttcgagc	cgcccgcgca	9840
cgtctcaacc	gtgcggctgc	atgaaatcct	ggccggtttg	tctgatgcca	agctggcggc	9900
ctggccggcc	agcttgggcg	ctgaagaaac	cgagcgccgc	cgtctaaaaa	ggtgatgtgt	9960
atttgagtaa	aacagcttgc	gtcatgcggt	cgctgcgtat	atgatgcgat	gagtaaataa	10020
acaaatacgc	aaggggaacg	catgaagggt	atcgctgtac	ttaaccagaa	aggcgggtca	10080
ggcaagacga	ccatcgcaac	ccatctagcc	cgcgccctgc	aactcgccgg	ggccgatgtt	10140
ctgttagtcg	attccgatcc	ccagggcagt	gcccgcgatt	ggcgggccgt	gcgggaagat	10200
caaccgctaa	ccgttgctcg	catcgaccgc	ccgacgattg	accgcgacgt	gaaggccatc	10260
ggccggcgcg	acttcgtagt	gatcgacgga	gcgccccagg	cgccggactt	ggctgtgtcc	10320
gcgatcaagg	cagccgactt	cgtgctgatt	ccggtgcagc	caagccctta	cgacatatgg	10380
gccaccgccc	acctggtgga	gctgggttaag	cagcgcatgt	aggtcacgga	tggaaggcta	10440
caagcggcct	ttgtcgtgtc	gcgggcatc	aaaggcacgc	gcatcgggcg	tgaggttgcc	10500
gaggcgctgg	ccgggtacga	gctgcccatt	cttgagtccc	gtatcacgca	gcgcgtgagc	10560
taccagggca	ctgccgccgc	cggcacaacc	gttcttgaat	cagaaccgga	gggcgacgct	10620
gcccgcgagg	tccaggcgct	ggccgctgaa	attaaatcaa	aactcatttg	agttaatgag	10680
gtaaagagaa	aatgagcaaa	agcacaaaca	cgctaagtgc	cgcccgctcc	agcgcacgca	10740
gcagcaaggc	tgcaacgttg	gccagcctgg	cagacacgcc	agccatgaag	cggttcaact	10800
ttcagttgcc	ggcggaggat	cacaccaagc	tgaagatgta	cgcggtacgc	caaggcaaga	10860
ccattaccga	gctgctatct	gaatacatcg	cgcagctacc	agagtaaatg	agcaaatgaa	10920
taaatgagta	gatgaatttt	agcggctaaa	ggaggcggca	tggaaaatca	agaacaacca	10980
ggcaccgacg	ccgtggaatg	ccccatgtgt	ggaggaaacg	gcggttgccc	aggcgtaagc	11040
ggctgggttg	tctgccggcc	ctgcaatggc	actggaaccc	ccaagcccga	ggaatcggcg	11100

tgagcggctcg	caaaccatcc	ggcccgggtac	aaatcggcgc	ggcgtgggt	gatgacctgg	11160
tggagaagtt	gaaggccgcg	caggccgccc	agcggcaacg	catcgaggca	gaagcacgcc	11220
ccggtgaatc	gtggcaagcg	gccgtgatc	gaatccgcaa	agaatcccgg	caaccgccgg	11280
cagccgggtgc	gccgtcgatt	aggaagccgc	ccaagggcga	cgagcaacca	gattttttcg	11340
ttccgatgct	ctatgacgtg	ggcaccgcg	atagtcgcag	catcatggac	gtggccgttt	11400
tccgtctgtc	gaagcgtgac	cgacgagctg	gcgaggtgat	ccgtacgag	cttcacagcg	11460
ggcacgtaga	ggtttccgca	gggcccggcg	gcatggccag	tgtgtgggat	tacgacctgg	11520
tactgatggc	ggtttcccat	ctaaccgaat	ccatgaaccg	ataccgggaa	gggaagggag	11580
acaagcccgg	ccgctgttc	cgtccacacg	ttgcggacgt	actcaagttc	tgccggcgag	11640
ccgatggcgg	aaagcagaaa	gacgacctgg	tagaaacctg	cattcggtta	aacaccacgc	11700
acgttgccat	gcagcgtacg	aagaaggcca	agaacggccg	cctggtgacg	gtatccgagg	11760
gtgaagcctt	gattagccgc	tacaagatcg	taaagagcga	aaccgggccc	ccggagtaca	11820
tcgagatcga	gctagctgat	tggatgtacc	gcgagatcac	agaaggcaag	aaccgggacg	11880
tgctgacggc	tcaccccgat	tactttttga	tcgatcccgg	catcgggcgt	tttctctacc	11940
gcctggcacg	ccgcgccgca	ggcaaggcag	aagccagatg	gttggtcaag	acgatctacg	12000
aacgcagtgg	cagcgcggga	gagttcaaga	agttctgttt	caccgtgcgc	aagctgatcg	12060
ggtcaaatga	cctgccggag	tacgatttga	aggaggaggc	ggggcaggct	ggcccgatcc	12120
tagtcatgcg	ctaccgcaac	ctgatcgagg	gcgaagcatc	cgccgggttc	taatgtacgg	12180
agcagatgct	agggcaaatt	gccctagcag	gggaaaaagg	tcgaaaaggc	ctctttcctg	12240
tggatagcac	gtacattggg	aacccaaagc	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
ttcggtcgct	gcgtcccta	cgccccgcg	cttcgcgtcg	gcctatcgcg	gccgctggcc	12540
gctcaaaaat	ggctggccta	cggccaggca	atctaccagg	gcgcggacaa	gccgcgccgt	12600
cgccactcga	ccgcgggcgc	ccacatcaag	gcaccctgcc	tcgcgcgttt	cggatgatgac	12660
ggtgaaaacc	tctgacacat	gcagctcccg	gagacgggtca	cagcttgtct	gtaagcggat	12720
gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	ttggcgggtg	tcggggcgca	12780
gccatgaccc	agtcacgtag	cgatagcgga	gtgtatactg	gcttaactat	gcggcatcag	12840
agcagattgt	actgagagtg	caccatatgc	ggtgtgaaat	accgcacaga	tcgtaagga	12900
gaaaataccg	catcaggcgc	tcttccgctt	cctcgctcac	tgactcgctg	cgctcggtcg	12960
ttcggtcgcg	gcgagcggta	tcagctcact	caaaggcggc	aatacgggtta	tccacagaat	13020

caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	13080
aaaaggccgc gttgctggcg tttttccata ggctccgcc ccctgacgag catcacaaaa	13140
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	13200
cccctggaag ctccctcgtg cgctctcctg ttccgacct gccgcttacc ggatacctgt	13260
ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca	13320
gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccg	13380
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccgtaaga cagcacttat	13440
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	13500
cagagttctt gaagtggagg cctaactacg gctacactag aaggacagta tttggtatct	13560
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac	13620
aaaccaccgc tggtagcggg ggtttttttg tttgcaagca gcagattacg cgcagaaaaa	13680
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa	13740
actcagtta agggattttg gtcattgcag atatatctcc caatttggtg agggcttatt	13800
atgcacgctt aaaaataata aaagcagact tgacctgata gtttggtgtg gagcaattat	13860
gtgcttagtg catctaacgc ttgagttaag ccgcgccgcg aagcggcgtc ggcttgaacg	13920
aatttctagc tagacattat ttgccgacta ccttggtgat ctgcctttc acgtagtggg	13980
caaattcttc caactgatct gcgcgcgagg ccaagcgatc ttcttcttgt ccaagataag	14040
cctgtctagc ttcaagtatg acgggctgat actgggccgg caggcgctcc attgcccagt	14100
cggcagcgac atccttcggc gcgattttgc cggttactgc gctgtaccaa atgcgggaca	14160
acgtaagcac tacatttcgc tcatcgccag ccagtcggg cggcgagttc catagcgta	14220
aggtttcatt tagcgctca aatagatcct gttcaggaac cggatcaaag agttcctccg	14280
ccgctggacc taccaaggca acgctatggt ctcttgcttt tgtcagcaag atagccagat	14340
caatgtcgat cgtggctggc tcgaagatac ctgcaagaat gtcattgcgc tgccattctc	14400
caaattgcag ttgcgcctta gctggataac gccacggaat gatgtcgtcg tgcacaacaa	14460
tggtgacttc tacagcgcg agaatctcgc tctctccagg ggaagccgaa gtttccaaaa	14520
ggtcgttgat caaagctcgc cgcgttggtt catcaagcct tacggtcacc gtaaccagca	14580
aatcaatatc actgtgtggc ttcaggccgc catccactgc ggagccgtac aaatgtacgg	14640
ccagcaacgt cggttcgaga tggcgctcga tgacgccaac tacctctgat agttgagtcg	14700
atacttcggc gatcacgcgt tccccatga tgtttaactt tgttttaggg cgactgcctt	14760
gctgcgtaac atcggtgctg ctccataaca tcaaacatcg acccagggcg taacgcgctt	14820
gctgcttgga tgcccaggcg 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 ttcacgcctt tttaaataac cgattattct 60
 aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt 120
 taaactgaag 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 aaagtgcctag aggggcatag taatcaaact tgtcaaaacc	420
gtcatcatga tgagggatga cataatataa aaagttgact aaggtcttgg tagtactctt	480
tgattagtat tatatatattg tgagaacatg agtcaagagg agacaagaaa ccgaggaacc	540
atagtttagc aacaagatgg aagttgcaaa gttgagctag ccgctcgatt agttacatct	600
cctaagcagt actacaagga atggtctcta tacttttcacg tttagcacat ggtagtgagg	660
attgacaagt tagaaacagt gcttaggaga caaagagtca gtaaagggtat tgaaagagtg	720
aagttgatgc tcgacaggtc aggagaagtc cctccgccag atggtgacta ccaagggggt	780
ggtatcagct gagacccaaa taagattctt cggttgaaacc agtggttcga ccgagactct	840
taggggtggga tttcactgta agatttgtgc attttgttga atataaattg acaatttttt	900
ttatttaatt atagattatt tagaatgaat tacatatatta gtttctaaca aggatagcaa	960
tggatgggta tgggtacagg ttaaaccatat ctattacca cccatctagt cgtcgggttt	1020
tacacgtacc caccggttta 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
aggtatcgag gcggatacta ctaaattccat cccactcgct atccgataat cactgggttc	1440
gggtataccc attcccgtca 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 aaaatcaaaa tgcatggcag gctgtttgta accttggaat aagatgttgg	1860
ccaattctgg agccgccacg tacgcaagac tcagggccac gttctcttca tgcaaggata	1920
gtagaacacc actccacca cctcctatat tagacctttg cccaaccctc cccaactttc	1980
ccatcccatc cacaagaaa ccgacatttt tatcataaat cactagtccc gggtacccaa	2040
gtttgtacaa aaaagcaggc tggtagctgg tgcttaaaca ctctgggtgag ttctagtact	2100

tctgctatga tcgatctcat taccatttct taaatttctc tccctaaata ttccgagttc	2160
ttgatttttg ataacttcag gttttctctt tttgataaat ctggctcttc catttttttt	2220
tttttggtggt taatttagtt tcctatgttc ttcgattgta ttatgcatga tctgtgtttg	2280
gattctgtta gattatgtat tggatgaatat gtatgtgttt ttgcatgtct ggttttggtc	2340
ttaaaaatgt tcaaactctga tgatttgatt gaagcttttt tagtggtggt ttgattcttc	2400
tcaaaactac tgtaattta ctatcatgtt ttccaacttt gattcatgat gacacttttg	2460
ttctgctttg ttataaaatt ttggttggtt tgattttgta attatagtgt aattttgtta	2520
ggaatgaaca tgttttaata ctctgttttc gatttgtcac acattcgaat tattaatcga	2580
taatttaact gaaaattcat ggttctagat cttgttgtca tcagattatt tgtttcgata	2640
attcatcaaa tatgtagtcc ttttgctgat ttgcgactgt ttcatTTTTT ctcaaaattg	2700
ttttttgtta agtttatcta acagttatcg ttgtcaaaag tctctttcat ttgcaaaat	2760
cttctttttt tttttgtttg taactttgtt ttttaagcta cacatttagt ctgtaaaata	2820
gcatcgagga acagttgtct tagtagactt gcatgttctt gtaacttcta tttgtttcag	2880
tttgttgatg actgctttga tttttaggtt caaaccatgg aagacgcaa aaacataaag	2940
aaaggcccg cgccattcta tccgctggaa gatggaaccg ctggagagca actgcataag	3000
gctatgaaga gatacgccct ggttcctgga acaattgctt ttacagatgc acatatcgag	3060
gtggacatca cttacgctga gtacttcgaa atgtccgttc ggttggcaga agctatgaaa	3120
cgatatgggc tgaatacaaa tcacagaatc gtcgtatgca gtgaaaactc tcttcaattc	3180
tttatgccgg tgttggggcg gttatttatc ggagttgcag ttgcgcccgc gaacgacatt	3240
tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt ggtgttcgtt	3300
tccaaaaagg ggttgcaaaa aattttgaac gtgcaaaaaa agctcccaat catccaaaaa	3360
attattatca tggattctaa aacggattac cagggatttc agtcgatgta cacgttcgtc	3420
acatctcatc tacctcccgg ttttaatgaa tacgattttg tgccagagtc cttcgatagg	3480
gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc taaagggtgc	3540
gctctgcctc atagaactgc ctgcgtgaga ttctcgcatg ccagagatcc tatttttggc	3600
aatcaaatca ttccggatac tgcgatttta agtggtgttc cattccatca cggttttgga	3660
atgtttacta cactcgata tttgatatgt ggatttcgag tcgtcttaat gtatagattt	3720
gaagaagagc tgtttctgag gagccttcag gattacaaga ttcaaagtgc gctgctggtg	3780
ccaaccctat tctccttctt cgccaaaagc actctgattg acaaatacga tttatcta	3840
ttacacgaaa ttgcttctgg tggcgctccc ctctctaagg aagtcgggga agcggttgcc	3900
aagagggtcc atctgccagg tatcaggcaa ggatatgggc tactgagac tacatcagct	3960
attctgatta caccgaggg ggatgataaa ccgggcgcgg tcggtaaagt tgttccattt	4020

tttgaagcga aggttgtgga tctggatacc gggaaaacgc tgggcgttaa tcaaagaggc	4080
gaactgtgtg tgagaggtcc tatgattatg tccggttatg taaacaatcc ggaagcgacc	4140
aacgccttga ttgacaagga tggatggcta cattctggag acatagctta ctgggacgaa	4200
gacgaacact tcttcatcgt tgaccgcctg aagtctctga ttaagtacaa aggctatcag	4260
gtggctcccg ctgaattgga atccatcttg ctccaacacc ccaacatctt cgacgcaggt	4320
gtcgcaggtc ttcccgacga tgacgcgggt gaacttcccg ccgccgttgt tgttttggag	4380
cacggaaaga cgatgacgga aaaagagatc gtggattacg tcgccagtca agtaacaacc	4440
gcgaaaaagt tgcgcggagg agttgtgttt gtggacgaag taccgaaagg tcttaccgga	4500
aaactcgacg caagaaaaat cagagagatc ctcataaagg ccaagaaggg cggaaagatc	4560
gccgtgtaac tcgagcatat gggctcgaat ttccccgacg gttcaaacaat ttggcaataa	4620
agtttcttaa gattgaatcc tgttgccgggt cttgcgatga ttatcatata atttctgttg	4680
aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt	4740
tttatgatta gagtcccgca attatacatt taatacgcga tagaaaacaa aatatagcgc	4800
gcaaactagg ataaattatc gcgcgcgggt tcatctatgt tactagatcg ggaattcaag	4860
cttggcgtaa tcatggaccc agctttcttg tacaaagtgg ggtaccaatt cgaatccaaa	4920
aattacggat atgaatatag gcatatccgt atccgaatta tccgtttgac agctagcaac	4980
gattgtacaa ttgcttcttt aaaaaaggaa gaaagaaaga aagaaaagaa tcaacatcag	5040
cgttaacaaa cggccccgtt acggcccaaa cggtcataata gagtaacggc gttaagcggt	5100
gaaagactcc tatcgaaata cgtaaccgca aacgtgtcat agtcagatcc cctcttctct	5160
caccgcctca aacacaaaaa taatcttcta cagcctatat atacaacccc cccttctatc	5220
tctccttctt cacaattcat catcttctct tctctacccc caattttaag aaatcctctc	5280
ttctcctctt cattttcaag gtaaattctct ctctctctct ctctctctgt tattccttgt	5340
tttaattagg tatgtattat tgctagtttg ttaatctgct tatcttatgt atgccttatg	5400
tgaatatctt tatcttggtc atctcatccg tttagaagct ataaatttgt tgatttgact	5460
gtgtatctac acgtgggttat gtttatatct aatcagatat gaatttcttc atattgttgc	5520
gtttgtgtgt accaatccga aatcggtgat ttttttcatt 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 actttttgga	5820
gttgttgact tgattgtatt tgtgtgtgta tatgtgtgtt ctgatcttga tatgttatgt	5880

atgtgcagct gaaccatggc ggcggcaaca acaacaacaa caacatcttc ttcgatctcc	5940
ttctccacca aaccatctcc ttctctctcc aaatcaccat taccaatctc cagattctcc	6000
ctcccattct ccctaaaccc caacaaatca tcctctctct cccgccgccg cggtatcaaa	6060
tccagctctc cctcctccat ctccgccgtg ctcaacacaa ccaccaatgt cacaaccact	6120
ccctctccaa ccaaacctac caaacccgaa acattcatct cccgattcgc tccagatcaa	6180
ccccgcaaag gcgctgatat cctcgtcgaa gctttagaac gtcaaggcgt agaaaccgta	6240
ttcgcttacc ctggaggtag atcaatggag attcaccaag ccttaacccg ctcttcctca	6300
atccgtaacg tccttcctcg tcacgaacaa ggagggtgat tcgcagcaga aggatacgct	6360
cgatcctcag gtaaaccagg tatctgtata gccacttcag gtcccgagc taaaaatctc	6420
gttagcggat tagccgatgc gttgttagat agtgttcctc ttgtagcaat cacaggacaa	6480
gtccctcgtc gtatgattgg tacagatgcg tttcaagaga ctccgattgt tgaggtaacg	6540
cgttcgatta cgaagcataa ctatcttggt atggatgttg aagatatccc taggattatt	6600
gaggaagctt tctttttagc tacttctggt agacctggac ctgttttggt tgatgttcct	6660
aaagatatcc aacaacagct tgcgattcct aattgggaac aggctatgag attacctggt	6720
tatatgtcta ggatgcctaa acctccgga gattctcatt tggagcagat tgttaggttg	6780
atttctgagt ctaagaagcc tgtgtgtgat gttggtggtg gttgtttgaa ttctagcgat	6840
gaattgggta ggtttgttga gcttacgggg atccctgttg cgagtacgtt gatggggctg	6900
ggatcttata cttgtgatga tgagttgtcg ttacatatgc ttggaatgca tgggactgtg	6960
tatgcaaatt acgctgtgga gcatagtgat ttgttgttgg cgtttggggg aaggtttgat	7020
gatcgtgtca cgggtaagct tgaggctttt gctagtaggg ctaagattgt tcatattgat	7080
attgactcgg ctgagattgg gaagaataag actcctcatg tgtctgtgtg tggatgatgt	7140
aagctggctt tgcaagggat gaataagggt cttgagaacc gagcggagga gcttaagctt	7200
gattttggag tttggaggaa tgagtgaac gtacagaaac agaagtttcc gttgagcttt	7260
aagacgtttg gggaaagctat tcctccacag tatgcgatta aggtccttga tgagttgact	7320
gatggaaaag ccataataag tactggtgtc gggcaacatc aaatgtgggc ggcgcagttc	7380
tacaattaca agaaaccaag gcagtggcta tcatcaggag gccttgagac tatgggattt	7440
ggacttcctg ctgcgattgg agcgtctgtt gctaaccctg atgcgatagt tgtggatatt	7500
gacggagatg gaagctttat aatgaatgtg caagagctag ccactattcg tgtagagaat	7560
cttcagtgga aggtactttt attaaacaac cagcatcttg gcatgggttat gcaatgggaa	7620
gatcggttct acaaagctaa ccgagctcac acatttctcg gggatccggc tcaggaggac	7680
gagatattcc cgaacatgtt gctgtttgca gcagcttgcg ggattccagc ggcgagggtg	7740
acaaagaaag cagatctccg agaagctatt cagacaatgc tggatacacc aggaccttac	7800

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

gcacgcgtcg accctgtacc gcgcacttga gcgcagcgag gaagtgcgc ccaccgaggc	9780
caggcgggcg ggtgccttcc gtgaggacgc attgaccgag gccgacgccc tggcgggccg	9840
cgagaatgaa cgccaagagg aacaagcatg aaaccgcacc aggacggcca ggacgaaccg	9900
tttttcatta ccgaagagat cgaggcggag atgatcgcg cggggtacgt gttcgagccg	9960
cccgcgcacg tctcaaccgt gcggctgcat gaaatcctgg ccggtttgtc tgatgccaa	10020
ctggcgggcct ggccggccag cttggccgct gaagaaaccg agcgccgccc tctaaaaagg	10080
tgatgtgtat ttgagtaaaa cagcttgct catgcggtcg ctgcgtatat gatgcgatga	10140
gtaaataaac aaatacgcaa ggggaacgca tgaaggttat cgctgtactt aaccagaaag	10200
gcgggtcagg caagacgacc atcgcaaccc atctagcccc cgccctgcaa ctgcgcgggg	10260
ccgatgttct gttagtcgat tccgatcccc agggcagtg gcgcgattgg gcggccgtgc	10320
gggaagatca accgctaacc gttgtcggca tcgaccgcc gacgattgac cgcgacgtga	10380
aggccatcgg ccggcgcgac ttcgtagtga tcgacggagc gcccaggcg gcggacttgg	10440
ctgtgtccgc gatcaaggca gccgacttcg tgctgattcc ggtgcagcca agcccttacg	10500
acatatgggc caccgccgac ctggtggagc tggttaagca gcgcattgag gtcacggatg	10560
gaaggctaca agcggccttt gtcgtgtcgc gggcgatcaa aggcacgcgc atcggcgggtg	10620
aggttgccga ggcgctggcc gggtagcagc tgcccattct tgagtccgt atcacgcagc	10680
gcgtgagcta ccaggcact gccgccgccc gcacaaccgt tcttgaatca gaaccgagg	10740
gcgacgctgc ccgcgaggtc caggcgctgg ccgctgaaat taaatcaaaa ctcatattgag	10800
ttaatgaggt aaagagaaaa tgagcaaaa cacaacacg ctaagtgcc gccgtccgag	10860
cgcacgcagc agcaaggctg caacgttggc cagcctggca gacacgccag ccatgaagcg	10920
ggtcaacttt cagttgccg cgaggatca caccaagctg aagatgtacg cggtagcca	10980
aggcaagacc attaccgagc tgctatctga atacatcgcg cagctaccag agtaaataag	11040
caaatgaata aatgagtaga tgaattttag cggctaaagg aggcggcatg gaaaatcaag	11100
aacaaccagg caccgacgcc gtggaatgcc ccatgtgtgg aggaacgggc ggttggccag	11160
gcgtaagcgg ctgggttgtc tgccggccct gcaatggcac tggaaacccc aagcccagg	11220
aatcggcgtg agcggtcgca aaccatccgg cccggtacaa atcggcgcgg cgctgggtga	11280
tgacctgggt gagaagttga aggcgcgcga ggccgcccag cggcaacgca tcgaggcaga	11340
agcacgcccc ggtgaatcgt ggcaagcggc cgctgatcga atccgcaaag aatcccggca	11400
accgcgggca gccggtgcgc cgtcgattag gaagccgccc aagggcgacg agcaaccaga	11460
ttttttcggt ccgatgctct atgacgtggg caccgcgat agtcgcagca tcattggacgt	11520
ggccggtttt cgtctgtcga agcgtgaccg acgagctggc gaggtgatcc gctacgagct	11580
tccagacggg cacgtagagg tttccgcagg gccggccggc atggccagtg tgtgggatta	11640

cgacctggta ctgatggcgg tttcccatct aaccgaatcc atgaaccgat accgggaagg	11700
gaagggagac aagcccggcc gcgtgttccg tccacacgtt gcggacgtac tcaagttctg	11760
ccggcgagcc gatggcgga agcagaaaga cgacctggta gaaacctgca ttcggttaaa	11820
caccacgcac gttgccatgc agcgtacgaa gaaggccaag aacggccgcc tggtagcggg	11880
atccgagggg gaagccttga ttagccgcta caagatcgta aagagcgaaa ccgggcggcc	11940
ggagtacatc gagatcgagc tagctgattg gatgtaccgc gagatcacag aaggcaagaa	12000
cccggacgtg ctgacggttc accccgatta ctttttgatc gatcccggca tcggccgttt	12060
tctctaccgc ctggcacgcc gcgccgagg caaggcagaa gccagatggg tgttcaagac	12120
gatctacgaa cgcagtggca gcgccggaga gttcaagaag ttctgtttca ccgtgcgcaa	12180
gctgatcggg tcaaattgacc tgccggagta cgatttgaag gaggaggcgg ggcaggctgg	12240
cccgatccta gtcattgcgt accgcaacct gatcgagggc gaagcatccg ccggttccta	12300
atgtacggag cagatgctag ggcaaattgc cctagcaggg gaaaaaggtc gaaaaggctc	12360
ctttcctgtg gatagcacgt acattgggaa ccaaagccg tacattggga accggaaccc	12420
gtacattggg aacccaaagc cgtacattgg gaaccgggta cacatgtaag tgactgatat	12480
aaaagagaaa aaaggcgatt tttccgccta aaactcttta aaacttatta aaactcttaa	12540
aacccgcctg gcctgtgcat aactgtctgg ccagcgcaca gccgaagagc tgcaaaaagc	12600
gcctaccctt cggtcgctgc gtcacctacg ccccgccgct tcgcgctcggc ctatcgcggc	12660
cgctggccgc tcaaaaatgg ctggcctacg gccaggcaat ctaccagggc gcggacaagc	12720
cgcgccgtcg ccactcgacc gccggcgccc acatcaaggc accctgcctc gcgcgtttcg	12780
gtgatgacgg tgaaaacctc tgacacatgc agctcccga gacggtcaca gcttgtctgt	12840
aagcggatgc cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcggtgtc	12900
ggggcgagc catgaccag tcacgtagcg atagcggagt gtatactggc ttaactatgc	12960
ggcatcagag cagattgtac tgagagtgc ccatatgcgg tgtgaaatac cgcacagatg	13020
cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg	13080
ctcggtcgtt cggtgcggc gagcggatc agctcactca aaggcggtaa tacggttatc	13140
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag	13200
gaaccgtaaa aaggccgctg tgctggcggt tttccatagg ctccgcccc ctgacgagca	13260
tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat aaagatacca	13320
ggcgtttccc cctggaagct ccctcgtagc ctctcctgtt ccgaccctgc cgcttaccgg	13380
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag	13440
gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt	13500
tcagcccgac cgctgcgcct tatccggtaa ctatcgctctt gagtccaacc cggtaagaca	13560

cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 13620
 cgggtgctaca gagttcttga agtgggtggcc taactacggc tacactagaa ggacagtatt 13680
 tggatatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 13740
 cggcaaaaca accaccgctg gtagcgggtgg tttttttgtt tgcaagcagc agattacgcg 13800
 cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg acgctcagtg 13860
 gaacgaaaaac tcacgttaag ggatttttgt catgcatgat atatctccca atttgtgtag 13920
 ggcttattat gcacgcttaa aaataataaa agcagacttg acctgatagt ttggctgtga 13980
 gcaattatgt gcttagtgca tctaacgctt gagttaagcc gcgccgcgaa gcggcgctcg 14040
 cttgaacgaa tttctagcta gacattattt gccgactacc ttggtgatct cgcctttcac 14100
 gtagtggaca aattcttcca actgatctgc gcgcgaggcc aagcgatctt cttcttgctc 14160
 aagataagcc tgtctagctt caagtatgac gggctgatac tgggccggca ggcgctccat 14220
 tgcccagtcg gcagcgacat ccttcggcgc gatcttgccg gttactgccc tgtaccaaatt 14280
 gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg gcgagttcca 14340
 tagcgttaag gtttcattta gcgcctcaaa tagatcctgt tcaggaaccg gatcaaagag 14400
 ttctctcgcc gctggaccta ccaaggcaac gctatgttct cttgcttttg tcagcaagat 14460
 agccagatca atgtcgatcg tggctggctc gaagatacct gcaagaatgt cattgcgctg 14520
 ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga tgcgctcg 14580
 cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg aagccgaagt 14640
 ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta cggtcaccgt 14700
 aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgagg agccgtacaa 14760
 atgtacggcc agcaacgctg gttcgagatg gcgctcgatg acgccaacta cctctgatag 14820
 ttgagtcgat acttcggcga tcaccgcttc ccccatgatg tttaactttg ttttagggcg 14880
 actgccctgc tgcgtaacat cgttgctgct ccataacatc aaacatcgac ccacggcgta 14940
 acgcgcttgc tgcttggatg cccgagggat agactgtacc ccaaaaaaac agtcataaca 15000
 agccatgaaa accgccactg cgttccatg 15029

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

<220>
 <223> синтетична послідовність

<400> 68
 aataatggcg cgctgggtgc ttaaacactc tgggtgagt

38

<210> 69

<211> 40
 <212> ДНК
 <213> штучна

 <220>
 <223> синтетична послідовність

 <400> 69
 aataatggcg cgcctttgac ctacaaaatc aaagcagtca 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 acattaaaac 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
tatatataggcg cgccttatct cctgctcaaa gaaacca 37

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

<220>
<223> синтетична послідовність

<400> 80
tatatataggcg cgccactggt taagcttcac tgtct 35

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

<220>
<223> синтетична послідовність

<400> 81
tatatataggcg cgcctttctt ctaaagctga aagt 34

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

<220>
<223> векторна послідовність

<400> 82
gtgattttgt gccgagctgc cggtcgggga gctgttggtt ggctgggtggc aggatatatt 60
gtggtgtaaa caaattgacg cttagacaac ttaataacac attgcggacg tctttaatgt 120
actgaattta gttactgatc actgattaag tactgatatc ggtaccaagc ttccgcggct 180
gcagtgcagc gtgaccgggt cgtgccctc tctagagata atgagcattg catgtctaag 240
ttataaaaaa ttaccacata ttttttttgt cacacttggt tgaagtgcag tttatctatc 300
tttatacata tattttaaact ttactctacg aataatataa tctatagtag tacaataata 360
tcagtgtttt agagaatcat ataatgaac agttagacat ggtctaaagg acaattgagt 420
attttgacaa caggactcta cagttttatc ttttttagtgt gcatgtgttc tccttttttt 480
ttgcaaatag cttcacctat ataatacttc atccatttta ttagtacatc catttaggggt 540
ttaggggttaa tgggtttttat agactaattt ttttagtaca tctattttat tctatttttag 600
cctctaaatt aagaaaaacta aaactctatt ttagtttttt tatttaatag tttagatata 660
aaatagaata aaataaagtg actaaaaatt aaacaaatac cctttaagaa attaaaaaaa 720
ctaaggaaac atttttcttg tttcgagtag ataatgccag cctgttaaac gccgtcgacg 780
agtctaacgg acaccaacca gcgaaccagc agcgtcgcgt cgggccaagc gaagcagacg 840

gcacggcatc tctgtcgtg cctctggacc cctctcgaga gttccgctcc accgttggac	900
ttgtccgct gtcggcatcc agaaattgcg tggcggagcg gcagacgtga gccggcacgg	960
caggcggcct cctcctcctc tcacggcacc ggcagctacg ggggattcct tccccaccgc	1020
tccttcgctt tcccttcctc gcccgccgta ataaatagac acccctcca caccctcttt	1080
ccccaacctc gtgttggtcg gagcgcacac acacacaacc agatctcccc caaatccacc	1140
cgtcggcacc tccgcttcaa ggtacgccgc tegtctccc ccccccccc cctctctacc	1200
ttctctagat cggcggtccg gtccatggtt agggcccggt agttctactt ctgttcatgt	1260
ttgtgttaga tccgtgtttg tgtagatcc gtgctgctag cgttcgtaca cggatgcgac	1320
ctgtacgtca gacacgttct gattgctaac ttgccagtgt ttctctttgg ggaatcctgg	1380
gatggctcta gccgttccgc agacgggatc gatttcatga ttttttttgt ttcgttgcac	1440
agggtttggg ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc	1500
atcttttcat gctttttttt gtcttggttg tgatgatgtg gtctggttg gcggtcgttc	1560
tagatcggag tagaattctg tttcaaacta cctggtggat ttattaattt tggatctgta	1620
tgtgtgtgcc atacatattc atagttacga attgaagatg atggatggaa atatcgatct	1680
aggataggta tacatgttga tgcgggtttt actgatgcac atacagagat gctttttggt	1740
cgttggttg tgatgatgtg gtgtggttg gcggtcgttc attcgttcta gatcggagta	1800
gaatactgtt tcaaactacc tgggtgtattt attaatattt gaactgtatg tgtgtgtcat	1860
acatcttcat agttacgagt ttaagatgga tggaaatata gatctaggat aggtatacat	1920
gttgatgtgg gttttactga tgcataata tgatggcata tgcagcatct attcatatgc	1980
tctaaccttg 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 gccgcgccg	2220
cgtctaccgc gctcactggc gccactaccg ctgcgcccaa ggcgaggcgc cgggcgcacc	2280
tcctggccac ccgcgcgcc ctcgcgcgc ccatcaggtg ctcagcggcg tcaccgcga	2340
tgccgatggc tccccggcc acccgcctcc ggccgtgggg cccaccgat ccccgcaagg	2400
gcgcgcacat cctcgtcgag tccctcgagc gctgcggcgt ccgcgacgtc ttgcgcctacc	2460
ccggcggcac gtccatggag atccaccagg cactcaccgc cccccctc atcgccaacc	2520
acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgctcctcg	2580
gccgcgtcgg cgtctgcac gccacctccg gccccggcgc caccaacctt gtctccgcgc	2640
tcgccgacgc gctgctcgat tccgtcccca tggctgccat cacgggacag gtgccgcgac	2700
gcatgattgg caccgacgcc ttccaggaga cgccatcgt cgaggtcacc cgctccatca	2760

ccaagcacia ctacctggtc ctcgacgtcg acgacatccc ccgcgtcgtg caggaggctt	2820
tcttcctcgc ctcctctggt cgaccggggc cgggtcttgt cgacatcccc aaggacatcc	2880
agcagcagat ggcggtgcct gtctgggaca agcccatgag tctgcctggg tacattgcgc	2940
gccttcccaa gccccctgcg actgagttgc ttgagcaggt gctgcgtctt gttggtgaat	3000
cccggcgccc tgttctttat gttggcgggtg gctgcgcagc atctggtgag gagttgcgac	3060
gctttgtgga gctgactgga atccccgtca caactactct tatgggcctc ggcaacttcc	3120
ccagcgacga cccactgtct ctgcgcatgc taggtatgca tggcacgggtg tatgcaaatt	3180
atgcagtgga taaggccgat ctggtgcttg cacttggtgt gcggtttgat gatcgtgtga	3240
caggggaagat tgaggctttt gcaagcaggg ctaagattgt gcacgttgat attgatccgg	3300
ctgagattgg caagaacaag cagccacatg tgtccatctg tgcagatgtt aagcttgctt	3360
tgcagggcat gaatgctctt cttgaaggaa gcacatcaaa gaagagcttt gactttggct	3420
catggaacga tgagttggat cagcagaaga ggggaattccc ccttgggtat aaaacatcta	3480
atgaggagat ccagccacia tatgctattc aggttcttga tgagctgacg aaaggcgagg	3540
ccatcatcgg cacaggtgtt gggcagcacc agatgtgggc ggcacagtac tacacttaca	3600
agcggccaag gcagtgggtg tcttcagctg gtcttggggc tatgggattt ggtttgccgg	3660
ctgctgctgg tgcttctgtg gccaacccag gtgttactgt tgttgacatc gatggagatg	3720
gtagctttct catgaacgtt caggagctag ctatgatccg aattgagaac ctcccgtga	3780
aggtctttgt gctaaacaac cagcacctgg ggatgggtgt gcagtgggag gacaggttct	3840
ataaggccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatatc	3900
cagatttctg gacgatcgcc aaagggttca acattccagc ggtccgtgtg acaaagaaga	3960
acgaagtccg cgcagcgata aagaagatgc tcgagactcc agggccgtac ctcttgata	4020
taatcgcccc acaccaggag catgtgttgc ctatgatccc taatggtggg gctttcaagg	4080
atatgatcct ggatggtgat ggcaggactg tgtactgac taaaatccag caagcaactg	4140
atctaaaatc cagcaagcac cgctccctg ctagtacaag ggtgatatgt ttttatctgt	4200
gtgatgttct cctgtattct atcttttttt gtaggccgtc agctatctgt tatggtaatc	4260
ctatgtagct tccgaccttg taattgtgta gtctgttgtt ttcttcttgg catgtgtcat	4320
aagagatcat ttaagtgcct tttgtacat ataaataaga taataagcac tgctatgcag	4380
tggttctgaa ttggcttctg ttgccaaatt taagtgtcca actggtcctt gcttttgttt	4440
tcgctatttt tttctttttt tagttattat tatattggta atttcaactc aacatatgat	4500
gtatggaata atgctagggc tgcaatttca aactatttta caaaccagaa tggcattttc	4560
gtggtttgag gggagtgaat aaaaatgagg catttgactg aattagttac ctgatccatt	4620
ttcgtggttt ggatcattgg aattaaattc cattctaata atagtaattt tggcatatat	4680

caattaagtt aattcggttt tatgcaaaat atatttgtat actattatta tcaagatgtc	4740
ggagatattt atatgctaca tttttactat acaggagtga gatgaagagt gtcattgtaag	4800
ttacacagta gaaacaaatt ctattaatgc ataaaatcat ttccatcatc caccctatga	4860
atgtgagata gacctatata taaactttga aaagtgggtg aatatcaaatt tccaaattaa	4920
ataagttatt ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgccttc	4980
taaagctaatt ttggaaactc 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 caattgattc ccacttcatt tttctaaatg tagaaaacgg	5280
tgacgtataa gaaaaagaat gaattaggac ttttattccg tacactaatc tagagcggcc	5340
caagcttgta cactagtacg cgtcaattga tttaaattta attaatcccg tgtccgtcaa	5400
tgtgatacta ctagcatagt actagtacca tgcatacaca cagcaggctg gccgcctgga	5460
tggatogatg atgatactac atcatcctgt catccatcca ggcgatctag aaggggctg	5520
gctagctagc aaactgtgac cggtttttct acgccgataa taatactttg tcatggtaca	5580
gacgtacagt actgggtata tatatctgta gatttcaact gaaaagctag gatagctaga	5640
ttaatcctg agaaacacag ataaaattcg agcttggcta tagatgacaa aacggaagac	5700
gcatgcattg gacgacgtat gcaatgcgag cgcgtctcgt gtcgtcccgt ccaagtctgg	5760
cgatctcacg ccacgtgctc aacagctcaa ggactgttcg tcaccagcgt taaattcatt	5820
gaagggatga cgcatttcgg catttgtcat tgcttgtagc tatatatata tatccaacag	5880
atctctctca agcttttgta tgcgtgaatg taaagtctag cttatacgac agcacgtgca	5940
gatatattaa cgtcattatt aggtggagag caagatgcat gatctggtag aaattgtcga	6000
aaacacaaga gagagtgaag tgcacacttc tggatatagga gtgtatacgc cgctggttgg	6060
tgggcaatgc gcgccgcaat attggccaat gaaacctagc aacgcccact cgccacgccc	6120
catgaatggc ccccgacgag cagcgagcca gccagtgccg gcgcgaggcc cagccggagt	6180
cggcggaacg cgccacgggg gacgaggcgc ccgaggggcg aggcagcgcg gcatggcaag	6240
caagccgaag cgggcaagcg acctgcatgc agcccctgcc cctcgccctc gtcagtcgtc	6300
ccagcctccc actggaatcc acccaaccgc cccttctctc ccaaagcacg cgccccgga	6360
ctcgccctcg cctacgtgtc ggacgcgtcc ccgcgggtcg cccacgtacc ccgccccgtt	6420
ctcccacgtg cccctccctc tgcgcgcgtc cgattggctg acccgccctt cttaagccgc	6480
gccagcctcc tgtccggggc ccaacgcgct gctccgtcgt cgtctccgcc cccagagtga	6540
tcgagcccac tgacctggcc cccgagcctc agctcgtgag tccggcgcg cgtgtgctta	6600

aacactctgg	tgagttctag	tacttctgct	atgatcgatc	tcattaccat	ttcttaaatt	6660
tctctcccta	aatattccga	gttcttgatt	tttgataact	tcaggttttc	tctttttgat	6720
aaatctggtc	tttccatttt	tttttttttg	tggttaattt	agtttcctat	gttcttcgat	6780
tgtattatgc	atgatctgtg	tttggattct	gttagattat	gtattggtga	atatgtatgt	6840
gtttttgcat	gtctggtttt	ggtcttaaaa	atgttcaaat	ctgatgattt	gattgaagct	6900
tttttagtgt	tggtttgatt	cttctcaaaa	ctactgttaa	tttactatca	tgtttttcaa	6960
ctttgattca	tgatgacact	tttgttctgc	tttgttataa	aattttggtt	ggtttgattt	7020
tgtaattata	gtgtaatttt	gttaggaatg	aacatgtttt	aatactctgt	tttcgatttg	7080
tcacacattc	gaattattaa	tcgataattt	aactgaaaat	tcatggttct	agatcttggt	7140
gtcatcagat	tatttgtttc	gataattcat	caaatatgta	gtccttttgc	tgatttgcca	7200
ctgtttcatt	ttttctcaaa	attgtttttt	gttaagttaa	tctaacagtt	atcgttgtca	7260
aaagtctctt	tcattttgca	aaatcttctt	tttttttttg	tttgtaactt	tgttttttaa	7320
gctacacatt	tagtctgtaa	aatagcatcg	aggaacagtt	gtcttagtag	acttgcatgt	7380
tcttgtaact	tctatttggt	tcagtttggt	gatgactgct	ttgattttgt	aggtcaaagg	7440
cgcgcgggat	ccccgggtgg	tcagtccttt	atgttacgtc	ctgtagaaac	cccaaccggt	7500
gaaatcaaaa	aactcgacgg	cctgtgggca	ttcagtctgg	atcgcgaaaa	ctgtggaatt	7560
gatcagcggt	ggtgggaaag	cgcgttacaa	gaaagccggg	caattgctgt	gccaggcagt	7620
tttaacgatc	agttcgccga	tgcagatatt	cgtaattatg	cgggcaacgt	ctgggatcag	7680
cgcgaagtct	ttataccgaa	aggttgggca	ggccagcgta	tcgtgctgcg	tttcgatgcg	7740
gtcactcatt	acggcaaagt	gtgggtcaat	aatcaggaag	tgatggagca	tcagggcggc	7800
tatacgccat	ttgaagccga	tgtcacgccg	tatgttattg	ccgggaaaag	tgtacgtaag	7860
tttctgcttc	tacctttgat	atatatataa	taattatcat	taattagtag	taatataata	7920
tttcaaatat	ttttttcaaa	ataaaagaat	gtagtatata	gcaattgctt	ttctgtagtt	7980
tataagtgtg	tatattttaa	tttataactt	ttctaataa	tgacaaaaat	ttgttgatgt	8040
gcaggatatca	ccgtttgtgt	gaacaacgaa	ctgaactggc	agactatccc	gccgggaatg	8100
gtgattaccg	acgaaaacgg	caagaaaaag	cagtcttact	tccatgattt	ctttaactat	8160
gccggaatcc	atcgcagcgt	aatgctctac	accacgccga	acacctgggt	ggacgatatc	8220
accgtggtga	cgcattgtgc	gcaagactgt	aaccacgcgt	ctgttgactg	gcagggtgtg	8280
gccaatggtg	atgtcagcgt	tgaactgcgt	gatgcggatc	aacagggtgt	tgcaactgga	8340
caaggcacta	gcgggacttt	gcaagtgtgt	aatccgcacc	tctggcaacc	gggtgaaggt	8400
tatctctatg	aactgtgcgt	cacagccaaa	agccagacag	agtgtgatat	ctacccgctt	8460
cgcgtcggca	tccggtcagt	ggcagtgaag	ggcgaacagt	tcctgattaa	ccacaaaccg	8520

ttctacttta ctggcttttg tctcatgaa gatgaggact tgcgtggcaa aggattcgat	8580
aacgtgctga tgggtgcacga ccacgcatta atggactgga ttggggccaa ctctaccgt	8640
acctcgcat acccttacgc tgaagagatg ctcgactggg cagatgaaca tggcatcggtg	8700
gtgattgatg aaactgctgc tgtcggcttt aacctctctt taggcattgg tttcgaagcg	8760
ggcaacaagc cgaaagaact gtacagcgaa gaggcagtca acggggaaac tcagcaagcg	8820
cacttacagg cgattaaaga gctgatagcg cgtgacaaaa accaccaag cgtgggtgatg	8880
tggagtattg ccaacgaacc ggatacccg cgcgaagggt cacgggaata tttcgcgcca	8940
ctggcggaag caacgcgtaa actcgaccgg acgcgtccga tcacctgct caatgtaatg	9000
ttctgcgacg ctcacaccga taccatcagc gatctctttg atgtgctgtg cctgaaccgt	9060
tattacggat ggtatgtcca aagcggcgat ttggaaacgg cagagaagggt actggaaaaa	9120
gaacttcttg cctggcagga gaaactgcat cagccgatta tcatcaccga atacggcggtg	9180
gatacgttag ccgggctgca ctcaatgtac accgacatgt ggagtgaaga gtatcagtgt	9240
gcatggctgg atatgtatca ccgcgtcttt gatcgcgtca gcgcgctcgt cgggtgaacag	9300
gtatggaatt tcgcccattt tgcgacctcg caaggcatat tgcgcgttgg cggtacaag	9360
aaagggatct tcactcgca ccgcaaaccg aagtcggcgg cttttctgct gcaaaaacgc	9420
tggactggca tgaacttcgg tgaaaaaccg cagcaggag gcaaacaatg aatcaacaac	9480
tctcctggcg caccatcgtc ggctacagcc tcgggaattg ctaccgagct cctgcaggcc	9540
taggatcggt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgcgggtctt	9600
gcatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa	9660
tgcattgacgt tatttatgag atgggttttt atgattagag tcccgcgaatt atacatttaa	9720
tacgcgatag aaaaacaaat atagcgcgca aactaggata aattatcgcg cgcgggtgtca	9780
tctatgttac tagatcggcc ggccgtttta acttagttac taatcagtga tcagattgtc	9840
gttcccgcc ttcactttta actatcagtg tttgacagga tatattggcg ggtaaaccta	9900
agagaaaaga gcgtttatta gaataatcgg atatttaaaa gggcgtgaaa aggtttatcc	9960
gttcgtccat ttgtatgtca atattggggg ggggggaaag ccacgttgtg tctcaaaatc	10020
tctgatgtta cattgcacaa gataaaaata tatcatcatg aacaataaaa ctgtctgctt	10080
acataaacag taatacaagg ggtgttcgcc accatgagcc atatccagcg tgaaacctcg	10140
tgtcccgcc cgcgcctcaa ttccaatatg gatgccgacc tttatggcta caagtggcg	10200
cgcgacaacg tcggccagtc gggcgcgacc atttatcggc tttatggcaa acccgatgcc	10260
ccggaactgt tcctgaagca cggcaaaggc agcgtcgcaa acgatgtcac cgatgagatg	10320
gtccgcctga actggcttac cgagttcatg ccgtgccga cgattaagca tttcatccgt	10380
accccgagc atgcctggct cttgaccag gccattccgg gcaaaacggc ctttcaggtc	10440

cttgaagagt acccggactc cggtgagaat atcgtggacg ccctcgcggt cttcctccgc	10500
cgtttgcata gcatccccgt gtgcaactgc cccttcaact cggaccgggt tttccgcctg	10560
gcacaggccc agtcgcgcac gaataacggc ctcgttgacg cgagcgattt cgacgatgaa	10620
cggaaatggct ggccgggtgga acaggtttgg aaggaaatgc acaaactgct tccgttctcg	10680
ccggattcgg tggtcacgca tggtgatttt tccctggata atctgatctt tgacgagggc	10740
aagctgatcg gctgcatcga cgtgggtcgc gtcggtatcg ccgaccgcta tcaggacctg	10800
gcgatcttgt ggaattgcct cggcaggttc tcgccctcgc tccagaagcg cctgttccag	10860
aagtacggca tcgacaaccc ggatatgaac aagctccagt tccacctcat gctggacgaa	10920
tttttttgaa cagaattggt taattgggtg taacactggc agagcattac gctgacttga	10980
cgggacggcg gctttgttga ataaatcgaa cttttgctga gttgaaggat cgatgagttg	11040
aaggaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc	11100
tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag	11160
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgagataacc aaatactgtc	11220
cttctagtgt agccgtagt tggccaccac ttcaagaact ctgtagcacc gcctacatac	11280
ctcgtctcgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc	11340
gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt	11400
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt	11460
gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc	11520
ggcagggctg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt	11580
tatagtctcg tcgggtttcg ccacctctga cttgagcgtc gattttttgtg atgctcgtca	11640
ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacgggt cctggccttt	11700
tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt	11760
attaccgcct ttgagtgagc tgataccgct cggcgcagcc gaacgaccga gcgcagcgag	11820
tcagtgagcg aggaagcgga agagcgctg atgcggtatt ttctccttac gcatctgtgc	11880
ggtatttcac accgcatagg ccgcgatagg ccgacgcgaa gcggcggggc gtagggagcg	11940
cagcgaccga agggtaggag ctttttgacg ctcttcggct gtgcgctggc cagacagtta	12000
tgacacaggc aggcgggttt taagagtttt aataagtttt aaagagtttt aggcggaaaa	12060
atcgcctttt ttctctttta tatcagtcac ttacatgtgt gaccggttcc caatgtacgg	12120
ctttgggttc ccaatgtacg ggttcgggtt cccaatgtac ggctttgggt tcccaatgta	12180
cgtgctatcc acaggaaaga gaccttttcg acctttttcc cctgctaggg caatttgccc	12240
tagcatctgc tccgtacatt aggaaccggc ggatgcttcg ccctcgatca ggttgcggta	12300
gcgcatgact aggatcgggc cagcctgccc cgcctcctcc ttcaaactcg actccggcag	12360

gtcatttgac cccgatcagct tgcgcacggc gaaacagaac ttcttgaact ctccggcgct	12420
gccactgcgt tcgtagatcg tcttgaacaa ccatctggct tctgccttgc ctgcccgcgcg	12480
gcgtgccagg cggtagagaa aacggccgat gccggggctg atcaaaaagt aatcgggggtg	12540
aaccgtcagc acgtccgggt tcttgccttc tgtgatctcg cggtagatcc aatcagcaag	12600
ctcgatctcg atgtactccg gccgcccggc ttcgctcttt acgatcttgt agcggctaata	12660
caaggcttca ccctcgata ccgtcaccag gcggccgttc ttggccttct tggtagcgtg	12720
catggcaacg tgcgtgggtgt ttaaccgaat gcaggtttct accaggtcgt ctttctgctt	12780
tccgccatcg gctcgccggc agaacttgag tacgtccgca acgtgtggac ggaacacgcg	12840
gccgggcttg tctcccttcc ctcccggtta tcggttcatg gattcgggta gatgggaaac	12900
cgccatcagt accaggtcgt aatcccacac actggccatg ccggcggggc ctgcccgaac	12960
ctctacgtgc ccgtctggaa gctcgtagcg gatcacctcg ccagctcgtc ggtcacgctt	13020
cgacagacgg aaaacggcca cgtccatgat gctgcgacta tcgcccgtgc ccacgtcata	13080
gagcatcgga acgaaaaaat ctggttgctc gtcgcccttg ggccgcttcc taatcgacgg	13140
cgcaccggct gccggcggtt gccgggattc tttgcggatt cgatcagcgg ccccttgcca	13200
cgattcacgc gggcggtgct ctgcctcgat gcgttgccgc tgggcggcct gcgcggcctt	13260
caacttctcc accaggtcat caccacgcgc cgcgccgatt tgtaccgggc cggatgggtt	13320
gcgaccgctc acgccgattc ctccggcttg ggggttccag tgccattgca gggccggcag	13380
acaaccacgc cgcttacgcc tggccaaccg cccgttcctc cacacatggg gcattccacg	13440
gcgtcgggtg ctggttggtt ttgattttcc atgccgcctc ctttagccgc taaaattcat	13500
ctactcattt attcatttgc tcatttactc tggtagctgc gcgatgtatt cagatagcag	13560
ctcggtaatg gtcttgcctt ggcgtaccgc gtacatcttc agcttgggtg gatcctccgc	13620
cggcaactga aagttgacct gcttcatggc tggcggtgtc gccaggctgg ccaacgttgc	13680
agccttgctg ctgcgtgcgc tcggacggcc ggcacttagc gtgtttgtgc ttttgctcat	13740
tttctcttta cctcattaac tcaaagagt tttgatttaa tttcagcggc cagcgcctgg	13800
acctcgccgg cagcgtcgcc ctccgggttct gattcaagaa cggttgtgcc ggcggcggca	13860
gtgcctgggt agctcacgcg ctgcgtgata cgggactcaa gaatgggcag ctogtaccgc	13920
gccagcgcct cggcaacctc accgccgatg cgcgtgcctt tgatcgcccg cgacacgaca	13980
aaggccgctt gtagccttcc atccgtgacc tcaatgcgct gcttaaccag ctccaccagg	14040
tggcggtgg cccaaatgtc gtaagggtt ggctgcaccg gaatcagcac gaagtcggct	14100
gccttgatcg cggacacagc caagtccgcc gcctggggcg ctccgctgat cactacgaag	14160
tcgcgccggc cgatggcctt cagctcgcg tcaatcgctg ggccgctgat gccgacaacg	14220
gttagcgggt gatcttcccg caccgccgcc caatcgccgg cactgccttg gggatcgga	14280

tgcactaaca gaacatcggc cccggcgagt tgcagggcgc gggctagatg ggttgcgatg 14340
 gtcgtcttgc ctgaccgcgc tttctgggta agtacagcga taaccttcat gcgttcccct 14400
 tgcgtatttg tttatttact catcgcatca tatacgcagc gaccgcatga cgcaagctgt 14460
 ttactcaaaa tacacatcac ctttttagat gatca 14495

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

<220>
 <223> синтетична послідовність

<400> 83
 atatacgcggt ggtgcttaaa cactctggtg agt 33

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

<220>
 <223> синтетична послідовність

<400> 84
 atatggcgcg cctttgacct acaaaatcaa agcagtc 38

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

<220>
 <223> синтетична послідовність

<400> 85
 atatacgcggt 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 gctgttggtt ggctggtggc aggatatatt	60
gtggtgtaaa caaattgacg cttagacaac ttaataacac attgcggacg tctttaatgt	120
actgaattta gttactgac actgattaag tactgatatc ggtaccaagc ttccgcggt	180
gcagtgcagc gtgacccggt cgtgccctc tctagagata atgagcattg catgtctaag	240
ttataaaaaa ttaccacata ttttttttgt cacacttggt tgaagtgcag tttatctatc	300
tttatacata tattttaaact ttactctacg aataatataa tctatagtag tacaataata	360
tcagtgtttt agagaatcat ataaatgaac agttagacat ggtctaaagg acaattgagt	420
atthtgacaa caggactcta cagttttatc tttttagtgt gcatgtgttc tctttttttt	480
ttgcaaatac cttcacctat ataatacttc atccatttta ttagtacatc catttaggggt	540
ttaggggtta tggtttttat agactaattt ttttagtaca tctattttat tctatttttag	600
cctctaaatt aagaaaacta aaactctatt ttagtttttt tatttaatag tttagatata	660
aaatagaata aaataaagtg actaaaaatt aaacaaatac cctttaagaa attaaaaaaa	720
ctaaggaaac atthttcttg tttcgagtag ataatgccag cctgttaaac gccgtcgacg	780
agtctaacgg acaccaacca gcgaaccagc agcgtcgcgt cgggccaagc gaagcagacg	840
gcacggcatc tctgtcgtg cctctggacc cctctcgaga gttccgctcc accgttggac	900
ttgctccgct gtcggcatcc agaaattgct tggcggagcg gcagacgtga gccggcacgg	960
caggcggcct cctcctctc tcacggcacc ggcagctacg ggggattcct tcccaccgc	1020
tcttcgctt tcccttctc gcccgccgta ataaatagac acccctcca caccctcttt	1080
ccccaacctc gtgttggtcg gagcgcacac acacacaacc agatctcccc caaatccacc	1140
cgtcggcacc tccgcttcaa ggtacgccgc tcgtcctccc ccccccccc cctctctacc	1200
ttctctagat cggcgttccg gtccatgggt agggcccggt agttctactt ctgttcatgt	1260
ttgtgttaga tccgtgtttg tgtagatcc gtgtgctag cgttcgtaca cggatgcgac	1320
ctgtacgtca gacacgttct gattgctaac ttgccagtgt ttctctttgg ggaatcctgg	1380
gatggctcta gccgttcgc agacgggacg gatttcatga ttttttttgt ttggtgcat	1440
agggtttggt ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc	1500
atcttttcat gctttttttt gtcttggttg tgatgatgtg gtctggttgg gcggtcggtc	1560
tagatcggag tagaattctg tttcaacta cctggtggat ttattaattt tggatctgta	1620
tgtgtgtgcc atacatattc atagttacga attgaagatg atggatggaa atatcgatct	1680
aggataggta tacatgttga tgcgggtttt actgatgatc atacagagat gcttttttgt	1740
cgcttggttg tgatgatgtg gtgtggttgg gcggtcggtc attcgttcta gatcggagta	1800
gaatactggt tcaaaactacc tgggtgtattt attaatthtg gaactgtatg tgtgtgtcat	1860

acatcttcat agttacgagt ttaagatgga tggaaatatac gatctaggat aggtatacat	1920
gttgatgtgg gttttactga tgcataataca tgatggcata tgcagcatct attcatatgc	1980
tctaaccttg agtacctatc tattataata aacaagtatg ttttataatt atttcgatct	2040
tgatataactt ggatgatggc atatgcagca gctatatgtg gatTTTTTTta gccctgcctt	2100
catacgctat ttatttgctt ggtactgttt cttttgtcga tgetcaccct gttgtttggt	2160
gttacttctg cagccccggg gatccactag ttctagaaac catggccacc gccgcgcgcg	2220
cgtctaccgc gctcactggc gccactaccg ctgcgcccaa ggcgaggcgc cgggcgcacc	2280
tcctggccac ccgcgcgcgc ctgcgcgcgc ccatcaggtg ctcagcggcg tcaccgcga	2340
tgccgatggc tccccggcc accccgctcc ggccgtggg cccaccgat cccgcgaagg	2400
gcgcgcacat cctcgtcgag tccctcgagc gctgcggcgt ccgcgacgtc ttcgcctacc	2460
ccggcggcac gtccatggag atccaccagg cactcaccgc ctccccgtc atcgccaacc	2520
acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgctcctcgg	2580
gccgcgtcgg cgtctgcac gccacctccg gccccggcgc caccaacctt gtctccgcgc	2640
tcgccgacgc gctgctcgat tccgtcccca tggctgccat cacgggacag gtgccgcgac	2700
gcatgattgg caccgacgcc ttccaggaga cgcccatcgt cgaggtcacc cgctccatca	2760
ccaagcacia ctacctggtc ctgcacgtcg acgacatccc ccgcgtcgtg caggaggctt	2820
tcttctcgc ctctcttgtt cgaccggggc cgggtgctgt cgacatcccc aaggacatcc	2880
agcagcagat ggcgggtgcct gtctgggaca agcccatgag tctgcctggg tacattgcgc	2940
gccttcccaa gccccctgcg actgagttgc ttgagcaggt gctgcgtctt gttggtgaat	3000
cccggcgcgc tgttctttat gttggcggtg gctgcgcagc atctggtgag gagttgcgac	3060
gctttgtgga gctgactgga atcccgggtca caactactct tatgggcctc ggcaacttcc	3120
ccagcgacga cccactgtct ctgcgcacgc taggtatgca tggcacggtg tatgcaaatt	3180
atgcagtgga taaggccgat ctggtgcttg cacttggtgt gcggtttgat gatcgtgtga	3240
caggaagat tgaggctttt gcaagcaggg ctaagattgt gcacgttgat attgatccgg	3300
ctgagattgg caagaacaag cagccacatg tgtccatctg tgcagatgtt aagcttgctt	3360
tgcagggcat gaatgctctt cttgaaggaa gcacatcaaa gaagagcttt gactttggct	3420
catggaacga tgagttggat cagcagaaga gggaattccc ccttgggtat aaaacatcta	3480
atgaggagat ccagccacaa tatgctattc aggttcttga tgagctgacg aaaggcgagg	3540
ccatcatcgg cacaggtgtt gggcagcacc agatgtgggc ggcacagtac tacacttaca	3600
agcggccaag gcagtgggtg tcttcagctg gtcttggggc tatgggattt ggtttgccgg	3660
ctgctgctgg tgcttctgtg gccaacccag gtgttactgt tgttgacatc gatggagatg	3720
gtagctttct catgaacgtt caggagctag ctatgatccg aattgagaac ctcccggtga	3780

aggctcttgt gctaaacaac cagcacctgg ggatggtggt gcagtgggag gacaggttct	3840
ataaggccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatact	3900
cagatttcgt gacgacgcc aaagggttca acattccagc ggtccgtgtg acaaagaaga	3960
acgaagtccg cgcagcgata aagaagatgc tcgagactcc agggccgtac ctcttgata	4020
taatcgcccc acaccaggag catgtgttgc ctatgatccc taatggtggg gctttcaagg	4080
atatgatcct ggatggtgat ggcaggactg tgtactgatc taaaatccag caagcaactg	4140
atctaaaatc cagcaagcac cgctccctg ctagtacaag ggtgatatgt ttttatctgt	4200
gtgatgttct cctgtattct atcttttttt gtaggccgtc agctatctgt tatggtaatc	4260
ctatgtagct tccgacctg taattgtgta gtctgtgtt ttccttctgg catgtgtcat	4320
aagagatcat ttaagtgcct tttgctacat ataaataaga taataagcac tgctatgcag	4380
tggttctgaa ttggcttctg ttgccaaatt taagtgtcca actggtcctt gcttttgttt	4440
tcgctatttt tttccttttt tagttattat tatattggtg atttcaactc aacatatgat	4500
gtatggaata atgctagggc tgcaatttca aactatttta caaaccagaa tggcattttc	4560
gtggtttgag gggagtgaag aaaaatgagg catttgactg aattagttac ctgatccatt	4620
ttcgtggttt ggatcattgg aattaaattc cattctaata atagtaattt tggcatatat	4680
caattaagtt aattcggtt tatgcaaaat atatttgtat actattatta tcaagatgtc	4740
ggagatattt atatgctaca tttttactat acaggagtga gatgaagagt gtcatgtaag	4800
ttacacagta gaaacaaatt ctattaatgc ataaaatcat ttccatcatc caccctatga	4860
atttgagata gacctatata taaactttga aaagtggttg aatatcaa atccaaattaa	4920
ataagttatt ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgccctc	4980
taaagcta at ttggaaactc aaactttctt agcattggag gggattgaga aaaaatatta	5040
attcattttc atctcaatca ttcaatctcc aaagagattt gagttcctta ttagtctgtt	5100
ccatgcatca aatcggctca atgtgtcatt atttgccatg acgattgacg agttgttctg	5160
gggcctagcg ctttccacgc cgatgtgctg gggcctggtc ctggagaaga cagcttgata	5220
tttaaagcta tcaattgttt caattgattc ccacttcatt tttctaaatg tagaaaacgg	5280
tgacgtataa gaaaaagaat gaattaggac ttttattccg tacactaatc tagagcgcc	5340
caagcttgta cactagtacg cgtcaattga tttaaattta attaatcccg tgtccgtcaa	5400
tgtgatacta ctagcatagt actagtacca tgcatacaca cagcaggtcg gccgcctgga	5460
tggatcgatg atgatactac atcatcctgt catccatcca ggcgatctag aaggggcgtg	5520
gctagctagc aaactgtgac cggtttttct acgccgataa taatactttg tcatggtaca	5580
gacgtacagt actggttata tatatctgta gatttcaact gaaaagctag gatagctaga	5640
ttaattcctg agaaacacag ataaaattcg agcttggtc tagatgacaa aacggaagac	5700

gcatgcattg gacgacgtat gcaatgcgag cgcgtctcgt gtcgtcccgt ccaagtctgg	5760
cgatctcacg ccacgtgctc aacagctcaa ggactgttcg tcaccagcgt taaattcatt	5820
gaagggatga cgcatttcgg catttgtcat tgcttgtagc tatatatata tatccaacag	5880
atctctctca agcttttgta tgcgtgaatg taaagtctag cttatacgac agcacgtgca	5940
gatatattaa cgtcattatt aggtggagag caagatgcat gatctggtag aaattgtcga	6000
aaacacaaga gagagtgaag tgcacacttc tggatatagga gtgtatacgc cgctgggttg	6060
tgggcaatgc gcgcgcaat attggccaat gaaacctagc aacgcccact cgccacgccc	6120
catgaatggc ccccgacagg cagcgagcca gccagtgcc gcgcgcgggc cagccggagt	6180
cggcggaacg cgccacgggg gacgaggcgc ccgagggccg aggcagcgcg gcatggcaag	6240
caagccgaag cgggcaagcg acctgcatgc agcccctgcc cctcgccctc gtcagtcgtc	6300
ccagcctccc actggaatcc acccaaccgc cccttctctc ccaaagcacg cgccccgcga	6360
ctcgccctcg cctacgtgtc ggcagcgtcc ccgcgggtcg cccacgtacc ccgccccgtt	6420
ctcccacgtg cccctccctc tgcgcgcgtc cgattggctg acccgccctt cttaagccgc	6480
gccagcctcc tgtccggggc ccaacgcctg gtcctcgtcgt cgtctccgcc cccagagtga	6540
tcgagcccac tgacctggcc cccgagcctc agctcgtgag tccggcgctg ggtgcttaaa	6600
cactctggtg agttctagta cttctgctat gatcgatctc attaccattt cttaaatttc	6660
tctccctaaa tattccgagt tcttgatttt tgataacttc aggttttctc tttttgataa	6720
atctgggtctt tccatttttt tttttttgtg gttaatttag tttcctatgt tcttcgattg	6780
tattatgcat gatctgtgtt tggattctgt tagattatgt attggtgaat atgtatgtgt	6840
ttttgcatgt ctggttttgg tcttaaaaat gttcaaactc gatgatttga ttgaagcttt	6900
tttagtggtg gtttgattct tctcaaaact actgttaatt tactatcatg ttttccaact	6960
ttgattcatg atgacacttt tgttctgctt tgttataaaa ttttggttgg tttgattttg	7020
taattatagt gtaattttgt taggaatgaa catgttttaa tactctgttt tcgatttgtc	7080
acacattcga attattaatc gataatttaa ctgaaaattc atggttctag atcttggtgt	7140
catcagatta tttgtttcga taattcatca aatatgtagt ccttttgctg atttgcgact	7200
gtttcatttt ttctcaaaat tgttttttgt taagtttatc taacagttat cgttgtcaaa	7260
agtctctttc attttgcaaa atcttctttt tttttttgtt tgtaactttg ttttttaagc	7320
tacacattta gtctgtaaaa tagcatcgag gaacagttgt cttagtagac ttgcatgttc	7380
ttgtaacttc tatttgtttc agtttggtga tgactgcttt gatttttag gtcaaaggcg	7440
cgccaccatg gaagacgcca aaaacataaa gaaaggcccg gcgccattct atccgctgga	7500
agatggaacc gctggagagc aactgcataa ggctatgaag agatacgccc tggttcctgg	7560
aacaattgct tttacagatg cacatatcga ggtggacatc acttacgctg agtacttcga	7620

aatgtccggtt	cggttggcag	aagctatgaa	acgatatggg	ctgaatacaa	atcacagaat	7680
cgtcgtatgc	agtgaaaact	ctcttcaatt	ctttatgccg	gtgttgggcg	cgttatattat	7740
cggagttgca	gttgcgcccc	cgaacgacat	ttataatgaa	cgtgaattgc	tcaacagtat	7800
gggcatttcg	cagcctaccg	tgggtgttcgt	ttccaaaaag	gggttgcaaa	aaattttgaa	7860
cgtgcaaaaa	aagctcccaa	tcatccaaaa	aattattatc	atggattcta	aaacggatta	7920
ccagggattt	cagtcgatgt	acacgttcgt	cacatctcat	ctacctcccg	gttttaatga	7980
atacgatttt	gtgccagagt	ccttcgatag	ggacaagaca	attgcactga	tcatgaactc	8040
ctctggatct	actgggtctgc	ctaaagggtgt	cgctctgcct	catagaactg	cctgcgtgag	8100
attctcgcat	gccagagatc	ctatTTTTTg	caatcaaata	attccggata	ctgcgatttt	8160
aagtgttggt	ccattccatc	acggttttgg	aatgtttact	acactcggat	atttgatatg	8220
tggatttcga	gtcgtcttaa	tgtatagatt	tgaagaagag	ctgtttctga	ggagccttca	8280
ggattacaag	attcaaagtg	cgctgctggt	gccaaacctt	ttctccttct	tgcgcaaaag	8340
cactctgatt	gacaaatacg	atttatctaa	tttacacgaa	attgcttctg	gtggcgctcc	8400
cctctctaag	gaagtcgggg	aagcgggtgc	caagagggtc	catctgccag	gtatcaggca	8460
aggatatggg	ctcactgaga	ctacatcagc	tattctgatt	acacccgagg	gggatgataa	8520
accgggcgcg	gtcggtaaag	ttgttccatt	ttttgaagcg	aagggtgtgg	atctggatac	8580
cgggaaaacg	ctgggcgtta	atcaaagagg	cgaactgtgt	gtgagaggtc	ctatgattat	8640
gtccgggttat	gtaaacaata	cggaaagcgc	caacgccttg	attgacaagg	atggatggct	8700
acattctgga	gacatagctt	actgggacga	agacgaacac	ttcttcatcg	ttgaccgcct	8760
gaagtctctg	attaagtaca	aaggctatca	ggtggctccc	gctgaattgg	aatccatctt	8820
gtccaacac	ccaacatct	tcgacgcagg	tgtcgcagg	cttcccgcag	atgacgcg	8880
tgaacttccc	gccgcggttg	ttgttttgg	gcacggaaag	acgatgacgg	aaaaagagat	8940
cgtggattac	gtcgccagtc	aagtaacaac	cgcgaaaaag	ttgcgcggag	gagttgtgtt	9000
tgtggacgaa	gtaccgaaag	gtcttaccgg	aaaactcgac	gcaagaaaaa	tcagagagat	9060
cctcataaag	gccaagaagg	gcggaaagat	cgccgtgtaa	cctgcaggcc	taggatcggt	9120
caaacatttg	gcaataaagt	ttcttaagat	tgaatcctgt	tgccggtctt	gcgatgatta	9180
tcatataatt	tctgttgaat	tacgttaagc	atgtaataat	taacatgtaa	tgcatgacgt	9240
tatttatgag	atgggttttt	atgattagag	tcccgcatt	atacatttaa	tacgcgatag	9300
aaaacaaaat	atagcgcgca	aactaggata	aattatcgcg	cgcggtgtca	tctatgttac	9360
tagatcggcc	ggccgtttta	acttagttac	taatcagtga	tcagattgtc	gtttcccgcc	9420
ttcactttta	actatcagtg	tttgacagga	tatatggcg	ggtaaacctt	agagaaaaga	9480
gcgtttatta	gaataatcgg	atatttaaaa	gggcgtgaaa	aggtttatcc	gttcgtccat	9540

ttgtatgtca atattggggg ggggggaaag ccacgttggtg tctcaaaatc tctgatgtta	9600
cattgcacaa gataaaaata tatcatcatg aacaataaaa ctgtctgctt acataaacag	9660
taatacaagg ggtgttcgcc accatgagcc atatccagcg tgaaacctcg tgctcccgcc	9720
cgcgctcaa ttccaatatg gatgccgacc tttatggcta caagtgggcg cgcgacaacg	9780
tcggccagtc gggcgcgacc atttatcggc tttatggcaa acccgatgcc ccggaactgt	9840
tcctgaagca cggcaaaggc agcgtcgcaa acgatgtcac cgatgagatg gtccgcctga	9900
actggcttac cgagttcatg ccgctgccga cgattaagca tttcatccgt accccggacg	9960
atgcctggct cttgaccacg gccattccgg gcaaaacggc ctttcaggtc cttgaagagt	10020
acccggactc cggtgagaat atcgtggacg ccctcgcggt cttcctccgc cgtttgcata	10080
gcatccccgt gtgcaactgc cccttcaact cggaccgggt tttccgctg gcacaggccc	10140
agtcgcgcac gaataacggc ctggttgacg cgagcgattt cgacgatgaa cggaatggct	10200
ggccggtgga acaggtttgg aaggaaatgc acaaactgct tccgttctcg ccggattcgg	10260
tggtcacgca tggtgatttt tccctggata atctgatctt tgacgagggc aagctgatcg	10320
gctgcatcga cgtgggtcgc gtcggtatcg ccgaccgcta tcaggacctg gcgatcttgt	10380
ggaattgcct cggcgagttc tcgccctcgc tccagaagcg cctgttccag aagtacggca	10440
tcgacaaccc ggatatgaac aagctccagt tccacctcat gctggacgaa tttttttgaa	10500
cagaattggt taattggttg taacctggc agagcattac gctgacttga cgggacggcg	10560
gctttgttga ataaatcgaa cttttgctga gttgaaggat cgatgagttg aaggaccccg	10620
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc	10680
aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	10740
tttttccgaa ggtaactggc ttcagcagag cgagataacc aaatactgtc cttctagtgt	10800
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgcctctgc	10860
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	10920
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac	10980
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	11040
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg	11100
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtctcg	11160
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga	11220
gcctatggaa aaacgccagc aacgcggcct ttttacgggt cctggccttt tgctggcctt	11280
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attacgcct	11340
ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg	11400
aggaagcgga agagcgcctg atcggtattt ttctccttac gcatctgtgc ggtatttcac	11460

accgcatagg	cgcgatagg	cgcacgcgaa	gcggcggggc	gtagggagcg	cagcgaccga	11520
agggtagggc	ctttttgcag	ctcttcggct	gtgcgctggc	cagacagtta	tgcacaggcc	11580
aggcgggttt	taagagtttt	aataagtttt	aaagagtttt	aggcggaaaa	atcgcttttt	11640
ttctctttta	tatcagtcac	ttacatgtgt	gaccggttcc	caatgtacgg	ctttgggttc	11700
ccaatgtacg	ggttccgggt	cccaatgtac	ggctttgggt	tcccaatgta	cgtgctatcc	11760
acaggaaaga	gaccttttcg	acctttttcc	cctgctaggg	caatttgccc	tagcatctgc	11820
tccgtacatt	aggaaccggc	ggatgcttcg	ccctcgatca	ggttgcggtta	gcgcgatgact	11880
aggatcgggc	cagcctgccc	cgctcctccc	ttcaaactgt	actccggcag	gtcatttgac	11940
ccgatcagct	tgcgcacggg	gaaacagaac	ttcttgaact	ctccggcgct	gccactgcgt	12000
tcgtagatcg	tcttgaacaa	ccatctgggt	tctgccttgc	ctgcggcgcg	gcgtgccagg	12060
cggtagagaa	aacggccgat	gccggggctg	atcaaaaagt	aatcgggggtg	aaccgtcagc	12120
acgtccgggt	tcttgcttcc	tgtgatctcg	cggtagatcc	aatcagcaag	ctcgatctcg	12180
atgtactccg	gccgcccggg	ttcgctcttt	acgatcttgt	agcggctaata	caaggcttca	12240
ccctcgata	ccgtcaccag	gcggccggtc	ttggccttct	tggtacgctg	catggcaacg	12300
tgcgtgggtg	ttaaccgaat	gcaggtttct	accaggctcg	ctttctgctt	tccgccatcg	12360
gctcgccggc	agaacttgag	tacgtccgca	acgtgtggac	ggaacacgcg	gccgggcttg	12420
tctcccttcc	cttcccggta	tccgttcattg	gattcgggta	gatgggaaac	cgccatcagt	12480
accaggctcg	aatcccacac	actggccatg	ccggcggggc	ctgcggaaac	ctctacgtgc	12540
ccgtctggaa	gctcgtagcg	gatcacctcg	ccagctcgtc	ggtcacgctt	cgacagacgg	12600
aaaacggcca	cgtccatgat	gctgcgacta	tccgggggtg	ccacgtcata	gagcatcgga	12660
acgaaaaaat	ctgggttgctc	gtcgcccttg	ggcggttccc	taatcgacgg	cgcaccggct	12720
gccggcggtt	gccgggatcc	tttgcggtt	cgatcagcgg	ccccttgcca	cgattcaccg	12780
gggcgtgctt	ctgcctcgat	gcgttgccgc	tgggcggcct	gcgcggcctt	caacttctcc	12840
accaggctcat	caccacgcgc	cgcgccgatt	tgtaccgggc	cggatgggtt	gcgaccgctc	12900
acgccgatcc	ctcgggcttg	ggggttccag	tgccattgca	gggccggcag	acaaccacgc	12960
cgcttacgcc	tggccaaccg	cccgttcctc	cacacatggg	gcattccacg	gcgtcgggtgc	13020
ctgggttgctc	ttgattttcc	atgccgcctc	ctttagccgc	taaaattcat	ctactcattt	13080
attcattttgc	tcatttactc	tggtagctgc	gcgatgtatt	cagatagcag	ctcggtaatg	13140
gtcttgccct	ggcgtaccgc	gtacatcttc	agcttggtgt	gatcctccgc	cggcaactga	13200
aagttgaccc	gcttcattgg	tggcgtgtct	gccaggctgg	ccaacgttgc	agccttgctg	13260
ctgcgtgcgc	tccgacggcc	ggcacttagc	gtgtttgtgc	ttttgctcat	tttctcttta	13320
cctcattaac	tcaaagtgag	tttgatttaa	tttcagcggc	cagcgcttgg	acctcgcggg	13380


```

cagcgtcgcc ctcggttct gattcaagaa cgggttgcc ggcggcgga gtgcctgggt 13440
agctcacgcg ctgctgata cgggactcaa gaatgggcag ctctgacctg gccagcgctt 13500
cggcaacctc accgccgatg cgcgtgcctt tgatcgcccg cgacacgaca aaggccgctt 13560
gtagccttcc atccgtgacc tcaatgcgct gcttaaccag ctccaccagg tcggcggtgg 13620
cccaaagtgc gtaagggtt ggctgcaccg gaatcagcac gaagtcggct gccttgatcg 13680
cggacacagc caagtccgcc gcctggggcg ctccgtcgat cactacgaag tcgcgccggc 13740
cgatggcctt cacgtcgcg tcaatcgctg ggcggtcgat gccgacaacg gttagcggtt 13800
gatcttcccg cacggccgcc caatcgcggg cactgccttg gggatcgga tcgactaaca 13860
gaacatcggc ccggcgagt tgcaggcgcg gggctagatg ggttgcgatg gtcgtcttgc 13920
ctgaccgcgc tttctggtta agtacagcga taaccttcat gcgttccctt tgcgtatttg 13980
tttatttact catcgcatca tatacgcagc gaccgcatga cgcaagctgt tttactcaa 14040
tacacatcac ctttttagat gatca 14065

```

ФОРМУЛА ВИНАХОДУ

- 5 1. Спосіб одержання вискоекспресійного насінно-специфічного та/або насінно-селективного рослинного промотору, що включає:
 - a) одержання 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 промоторові,
- 10 причому NEENA має послідовність SEQ ID NO:1.
2. Спосіб за п. 1, який **відрізняється** тим, що додатково включає стадії:
 - a) включення NEENA, яка має послідовність SEQ ID NO:1, у рослину або її частину та
 - b) об'єднання зазначеної NEENA у геном зазначеної рослини або її частини, таким чином, щоб зазначена NEENA була функціонально зв'язаною з ендогенною насінно-специфічною та/або
 - 15 насінно-селективною експресованою нуклеїною кислотою, гетерологічною зазначеній NEENA, і, необов'язково,
 - c) регенерацію рослини або її частини, що включає зазначену NEENA з зазначеної трансформованої клітини.
- 20 3. Спосіб за п. 1 або 2, який **відрізняється** тим, що додатково включає етапи:
 - a) одержання експресійної конструкції, яка включає NEENA, яка має послідовність SEQ ID NO:1, 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 промотора, та
 - 25 b) об'єднання зазначеної експресійної конструкції, яка включає зазначену NEENA, у геном зазначеної рослини або її частини та, необов'язково,
 - c) регенерацію рослини або її частини, яка включає зазначені одну або кілька експресійних конструкцій, з зазначеної трансформованої рослини або її частини.
- 30 4. Спосіб за будь-яким з пп. 1-3, який **відрізняється** тим, що рослина є однодольною або дводольною рослиною.
5. Спосіб за п. 4, який **відрізняється** тим, що рослина є дводольною рослиною.
6. Спосіб за п. 4, який **відрізняється** тим, що рослина є однодольною рослиною.
7. Спосіб за будь-яким з пп. 1-6, який **відрізняється** тим, що зазначена NEENA є функціонально зв'язаною з насінно-специфічним та/або насінно-селективним промотором, наближеним до
- 35 сайту початку транскрипції зазначеної гетерологічної молекули нуклеїнової кислоти.
8. Спосіб за п. 7, 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.

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 перед сайтом початку трансляції молекули нуклеїнової кислоти, експресія якої перебуває під контролем зазначеного насінно-специфічного та/або насінно-селективного промотору.

10. 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 у межах 5'UTR молекули нуклеїнової кислоти, експресія якої перебуває під контролем зазначеного насінно-специфічного та/або насінно-селективного промотору.

11. Рекombінантна експресійна конструкція, яка включає NEENA, яка має послідовність SEQ ID NO:1.

12. Рекombінантна експресійна конструкція за п. 11, яка **відрізняється** тим, що включає NEENA, яка має послідовність SEQ ID NO:1, 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.

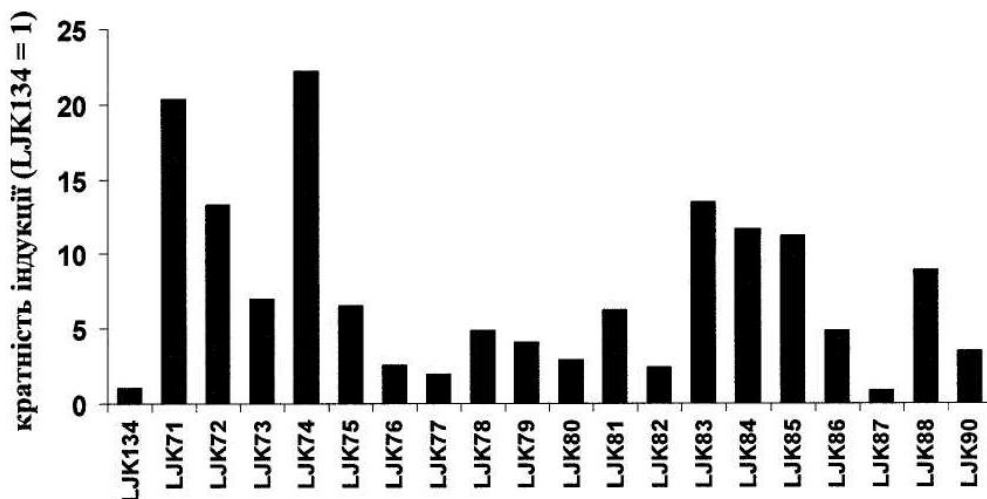
13. Трансгенна клітина або трансгенна рослина або її частина, яка включає рекombінантний вектор експресії, який включає одну або кілька рекombінантних експресійних конструкцій за п. 11 або 12 або рекombінантну експресійну конструкцію за п. 11 або 12.

14. Трансгенна клітина, яка включає рекombінантний вектор експресії, який включає одну або кілька рекombінантних експресійних конструкцій за п. 11 або 12 або рекombінантну експресійну конструкцію за п. 11 або 12, причому трансгенна клітина є вибраною або походить з групи, до якої належать бактерії, грибки, дріжджі або рослини.

15. Трансгенна рослина або її частина за п. 13, яка **відрізняється** тим, що зазначена рослина або її частина є дводольною рослиною.

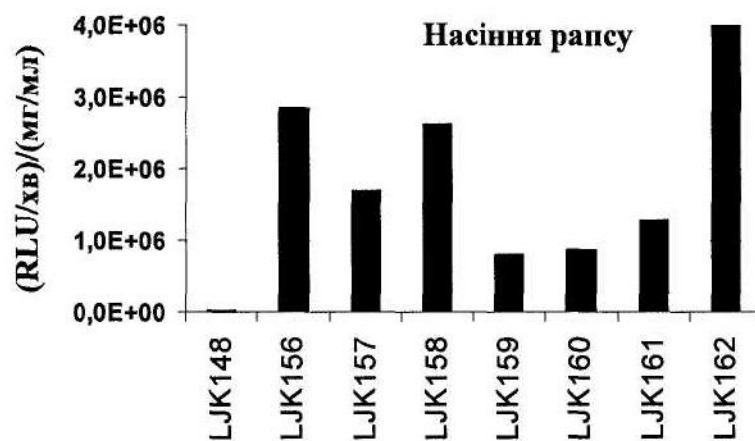
16. Трансгенна рослина або її частина за п. 13, яка **відрізняється** тим, що зазначена рослина або її частина є однодольною рослиною.

17. Застосування NEENA, яка має послідовність SEQ ID NO:1, або рекombінантної конструкції за п. 11 або 12 для посилення експресії у рослинах або їх частинах.

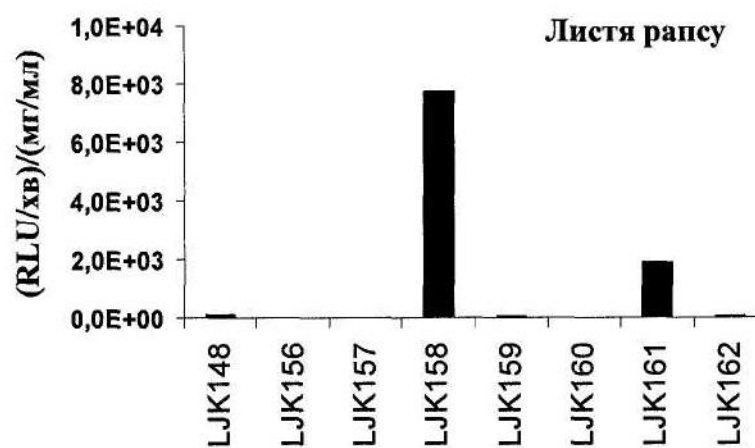


ФІГ. 1

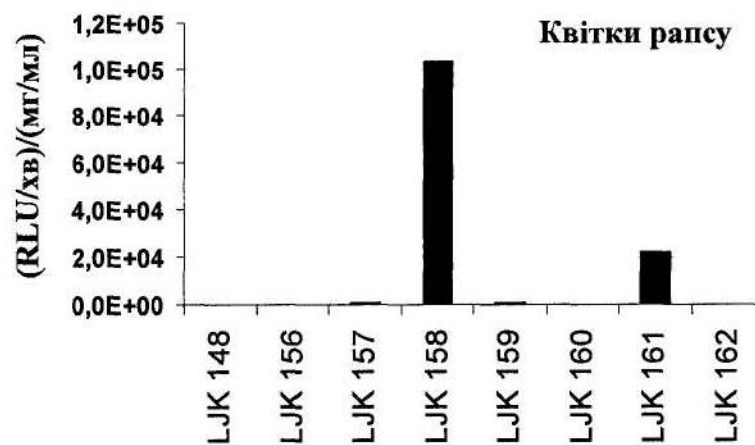
A)



B)

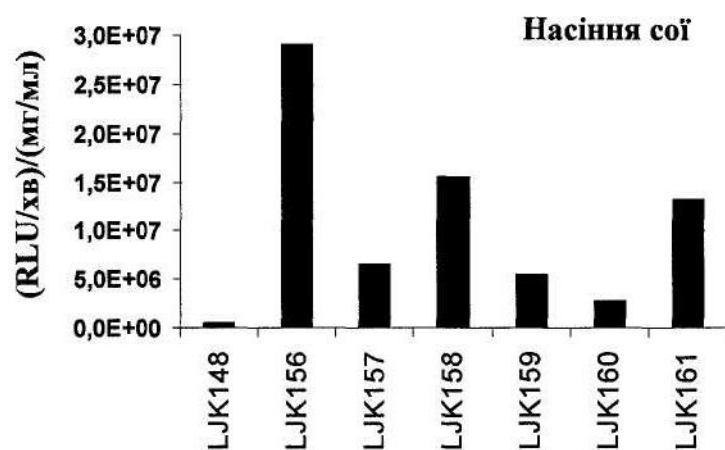


C)

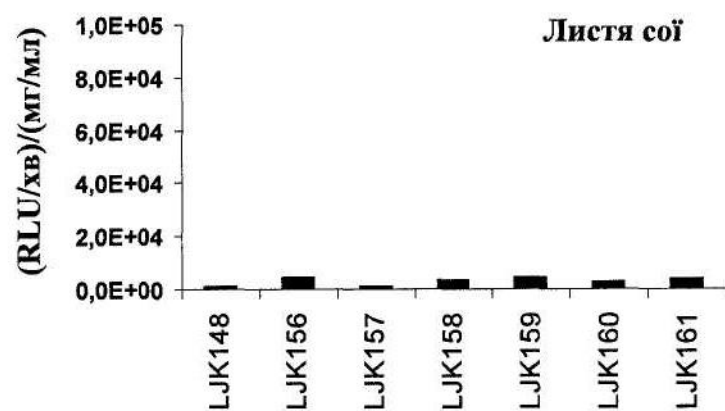


ФІГ. 2

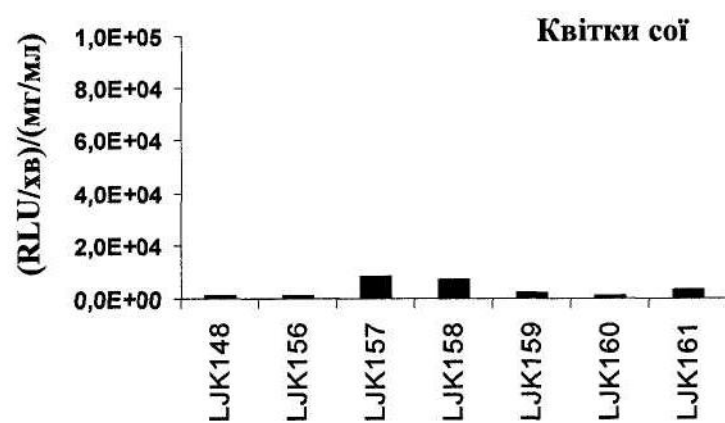
A)



B)

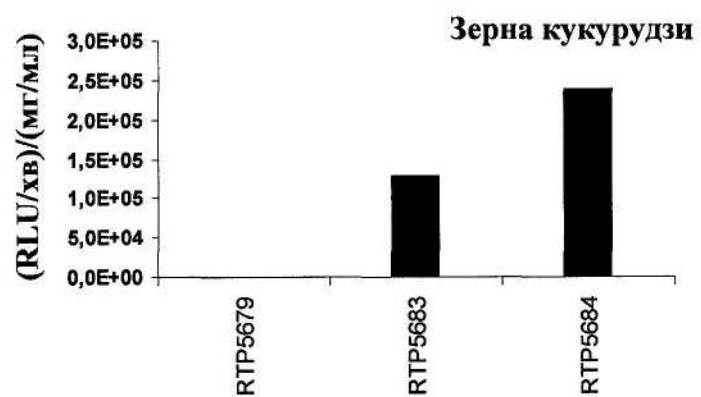


C)

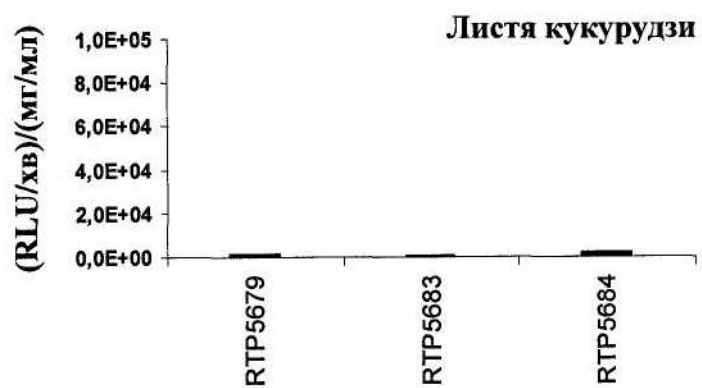


ФІГ. 3

A)

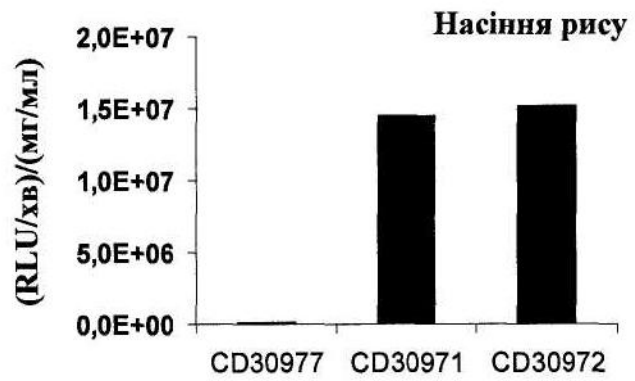


B)

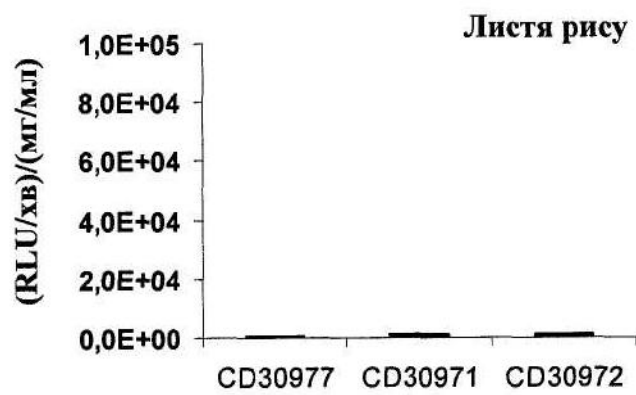


ФІГ. 4

A)



B)



ФІГ. 5

Комп'ютерна верстка О. Рябко

Державна служба інтелектуальної власності України, вул. Василя Липківського, 45, м. Київ, МСП, 03680, Україна

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