



УКРАЇНА

(19) **UA** (11) **122127** (13) **C2**
(51) МПК (2020.01)

C07K 14/47 (2006.01)

C12N 15/86 (2006.01)

A61K 48/00

C07K 14/755 (2006.01)

МІНІСТЕРСТВО РОЗВИТКУ
ЕКОНОМІКИ, ТОРГІВЛІ ТА
СІЛЬСЬКОГО ГОСПОДАРСТВА
УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

<p>(21) Номер заявки: а 2016 12986</p> <p>(22) Дата подання заявки: 23.04.2015</p> <p>(24) Дата, з якої є чинними права інтелектуальної власності: 26.09.2020</p> <p>(31) Номер попередньої заявки відповідно до Паризької конвенції: 1409089.8</p> <p>(32) Дата подання попередньої заявки відповідно до Паризької конвенції: 21.05.2014</p> <p>(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: GB</p> <p>(41) Публікація відомостей про заявку: 10.04.2017, Бюл.№ 7</p> <p>(46) Публікація відомостей про державну реєстрацію: 25.09.2020, Бюл.№ 18</p> <p>(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ: РСТ/GB2015/051201, 23.04.2015</p> <p>(72) Винахідник(и): Елтон Ерик Волтер Фредерик Вольфганг (GB), Гризенбах Юта (GB), Пітель Каміла Мальґожата (GB), Пол-Сміт Майкл Кристіан (GB), Принґл Іан Ендрю (GB), Гайд Стивен Чарльз (GB), Джил Дебора Ребека (GB), Девіс Лі Адріан (GB), Бойд Алан Кристофер (GB), МакЛеклен Джерард (GB), Іноуе Макото (JP)</p>	<p>(73) Володілець (володільці): ІМПЕРІАЛ ІННОВЕЙШНЗ ЛІМІТЕД, 52 Princes Gate, Exhibition Road, London SW7 2PG, United Kingdom (GB), АЙДІ ФАРМА КО., ЛТД., Techno Park Oho, 6 Ohkubo, Tsukuba, Ibaraki, 300-2611, Japan (JP)</p> <p>(74) Представник: Опанасенко Ольга Сергіївна, реєстр. №471</p> <p>(56) Перелік документів, взятих до уваги експертизою: EP 1950307 A1, 30.07.2008 WO 2007110628 A2, 04.10.2007 WO 03029274 A2, 10.04.2003 KATSUYUKI MITOMO et al. Toward gene therapy for cystic fibrosis using a lentivirus pseudotyped with sendai virus envelopes. Molecular therapy, 2010, Vol. 18, no. 6, P. 1173 – 1182 NAKAJIMA T et al. Development of novel simian immunodeficiency virus vectors carrying a dual gene expression system. Human gene therapy, Mary Ann Liebert, New York, NY, US, 2000, Vol. 11, no. 13, P. 1863 – 1874 DANIELA ZYCHLINSKI et al. Physiological promoters reduce the genotoxic risk of integrating gene vectors. Molecular therapy, 2008, Vol. 16, no. 4, P. 718 – 725 ENGELHARDT J.F. The lung as a metabolic factory for gene therapy. Journal of clinical investigation, 2002, Vol. 110, no. 4, P. 429 – 432 CHUAH M.K.L. et al. Gene therapy for hemophilia. Journal of gene medicine, John Wiley & sons, INC, US, 2001, Vol. 3, no. 1, P. 3 – 20</p>
---	---

(54) ЛЕНТИВІРУСНИЙ ВЕКТОР

(57) Реферат:

UA 122127 C2

Опис

Даний винахід стосується лентивірусних векторів для перенесення генів, псевдотипованих гемогліутинін-нейрамінідазою (HN) і гібридними (F) білками з респіраторного параміксовірусу, що містить промотор і трансген; і способів їхнього одержання. Даний винахід також стосується застосування вказаних векторів у генній терапії, зокрема, для лікування захворювань дихальних шляхів, таких як муковісцидоз (МВ).

ПОПЕРЕДНІЙ РІВЕНЬ ТЕХНІКИ

Лентивіруси належать до роду вірусів сімейства Retroviridae і характеризуються тривалим періодом інкубації. Лентивіруси можуть доставляти значну кількість вірусної РНК у ДНК клітину-хазяїна і мають унікальну здатність серед ретровірусів, маючи здатність інфікувати клітини, які не діляться, таким чином, вони є одним з найбільш ефективних методів створення вектора доставки гена.

Лентивірусні вектори, особливо, ті, які отримані з ВІЛ-1, є широко досліджуваними і часто застосовуваними векторами. Еволюція кістяка лентивірусних векторів і здатність вірусів доставляти молекули рекомбінантних ДНК (трансгени) у клітини-мішені привели до їх застосування в багатьох напрямках практичного використання. Два можливі напрямки використання вірусних векторів включають відновлення функціональних генів у генетичній терапії й одержання рекомбінантних білків *in vitro*.

Псевдотипування являє собою процес одержання вірусів або вірусних векторів у комбінації з чужорідними білками вірусної оболонки. Самі по собі, чужорідні білки вірусної оболонки можуть використовуватися для зміни тропізму хазяїна або збільшеної/зменшеної стабільності вірусних частинок. Наприклад, псевдотипування дозволяє встановлювати характер білків оболонки. Часто використовуваним білком є глікопротеїн G вірусу везикулярного стоматиту (VSV), скорочено VSV-G.

Зрозуміло, що для ефективною і контрольований ретровірусної експресії трансгена потрібна присутність інтронних послідовностей. Однак уведення таких інтронів у ретровірусні вектори включає ускладнені і потребуючі великих витрат часу способи, внаслідок використовуваних багатостадійних процесів.

Дотепер вірусні агенти перенесення генів були непридатними для лікування захворювань, без трансдукції популяцій стовбурових клітин, внаслідок адаптивної імунної відповіді хазяїна, що перешкоджає успішному повторному введенню.

Більше того, перенесення гена до епітелію дихальних шляхів виявився більш важким, ніж спочатку припускали. Наприклад, застосування лентивірусних псевдотипів, для якого потрібно руйнування епітеліальної цілісності, щоб здійснити трансдукцію в дихальні шляхи, наприклад, за допомогою використання детергентів, таких як лізофосфатидилхолін або етиленгліколь-біс(2-аміноетиловий ефір)-N,N,N'-тетраоцтова кислота, було пов'язане з підвищеним ризиком сепсису.

Одним прикладом клінічної ситуації, що могла б одержати сприятливий ефект від перенесення генів в епітелій дихальних шляхів є лікування муковісцидозу (МВ). МВ є летальним генетичним розладом, який викликається мутаціями в гені регулятора трансмембранної провідності при МВ (CFTR), що діє як хлоридний канал в епітеліальних клітинах дихальних шляхів. МВ характеризується рецидивуючими інфекціями дихальних шляхів, збільшеними секретами дихальних шляхів, і, в остаточному підсумку, дихальною недостатністю. У Великобританії в даний час середній вік померлих становить ~25 років. Для більшості генотипів, не існує терапій, націлених на базовий дефект; застосовувані в даний час терапії для полегшення симптомів вимагають годин самостійно ввідного лікування щодня. Генна терапія, на відміну від низькомолекулярних лікарських засобів, є незалежною від мутаційного класу CFTR і є, таким чином, застосовною до всіх уражених МВ індивідуумів. Однак, дотепер жоден вірусний вектор не відповідав вимогам для клінічного застосування, і те ж саме стосується інших захворювань, конкретно, багатьох інших захворювань дихальних шляхів.

У цьому зв'язку, зіштовхуються щонайменше з трьома основними проблемами. Ефективність перенесення генів звичайно є низькою, щонайменше почасти, оскільки відповідні рецептори для багатьох вірусних векторів, як видно, переважним чином локалізовані на базолатеральній поверхні епітелію дихальних шляхів. По-друге, проникність слизового шару дихальних шляхів звичайно є низькою. У кінцевому рахунку, здатність до повторного уведення вірусних векторів, обов'язкова для довгострокової обробки самооновлюваного епітелію є обмеженою.

Уведення векторів для клінічного застосування є ще одним актуальним фактором. Отже, для ефективності лікування повинна підтримуватися вірусна стабільність при використанні клінічно значимих пристроїв (наприклад, бронхоскопу і небулайзера).

Ще одним прикладом потенційної мішені для генної терапії є недостатність $\alpha 1$ -антитрипсину (A1AT). Недостатність A1AT являє собою спадкове порушення, яке може викликати захворювання легень і захворювання печінки. Симптоми включають задишку/свистяче дихання, знижену здатність до фізичного навантаження, втрату маси тіла, рецидивуючі респіраторні інфекції, стомлюваність і прискорене серцебиття при стоянні. У страждаючих індивідуумів часто розвивається емфізема. Приблизно в 10-15 % пацієнтів з недостатністю A1AT розвивається захворювання печінки. Індивідууми з недостатністю A1AT також мають ризик розвитку гепатоклітинної карциноми.

A1AT являє собою секретований білок, який виробляється, в основному, у печінці, і потім переноситься у легені, причому більш дрібні кількості також продукуються в самих легенях. Основною функцією A1AT є зв'язувати і нейтралізувати нейтрофіл-еластазу. Генна терапія A1AT, імовірно, буде мати терапевтичне значення в пацієнтів з недостатністю A1AT, MB і хронічним обструктивним захворюванням легень (ХОЗЛ), де підвищення або введення A1AT може поліпшити функцію легень.

Терапія A1AT також є потенційно цінною для лікування нереспіраторних/нелегенових захворювань, таких як діабет типу 1 і типу 2, гострий інфаркт міокарда, ревматоїдний артрит, запальне захворювання кишечника, відторгнення трансплантата, реакція трансплантат проти хазяїна (GvH), розсіяний склероз і інфекції, особливо, вірусні інфекції, унаслідок впливу недостатності A1AT на інші тканини/органи, такі як печінка і підшлункова залоза (див., наприклад, Lewis Mol. Med. 2012; 18:957-970, що включена в даний опис за допомогою посилання).

Недостатність A1AT є привабливим цільовим захворюванням для генної терапії, оскільки терапевтичні граничні рівні є добре визначеними. Порівняння рівнів A1AT у суб'єктів з ризиком розвитку емфіземи/ХОЗЛ дозволило визначити захисний граничний рівень, який дорівнює 11 мкМ у сироватці, причому рівні нижчі 11 мкМ використовують як граничні значення для ініціації стимулюючої білкової терапії, де це є доступним. Рівні A1AT в лаважній рідині дихальних шляхів становлять тільки ~10 % від сироваткового рівня, оскільки епітелій легень створює бар'єр, і терапевтичне граничне значення в лаважній рідині поверхні дихальних шляхів, отже, вважають таким, що дорівнює 1,1 мкМ (див. Ferraroti et al. Thorax. 2012 Aug; 67(8):669-74 і Abusriwil & Stockley 2006 Current Opinion in Pulmonary Medicine 12:125-131, кожна з яких включена в даний опис за допомогою посилання).

Шість дозволених FDA до застосування лікарських форм білка A1AT, що випускаються промислово, виділеного з об'єднаної крові людини, знаходять клінічне застосування в США для лікування пацієнтів з тяжкою недостатністю A1AT (за допомогою щотижневих внутрішньовенних ін'єкцій). Ферментозамісна терапія (ERT) є дорогою (~100000 доларів/рік), і, незважаючи на те, що біохімічна ефективність дії для стимулюючої білкової терапії ERT була доведена, доказ ефективності в клінічній практиці є значно більш важким. ERT A1AT у даний час недоступна у всіх країнах, і, у даний час, є недоступною у Великобританії. Крім того, використовуючи прийняті в даний час терапії, важко досягти достатніх перманентних тканинних рівнів, що може почасти бути причиною обмеженої ефективності в клінічній практиці, яка спостерігається дотепер.

Інші привабливі мішені для генної терапії включають серцево-судинні захворювання і захворювання крові, зокрема, порушення, пов'язані з недостатньою зсілістю крові, такі як гемофілія (A і B), хвороба фон Вілебранда і недостатність фактора VII.

Гемофілія, зокрема, гемофілія A, є привабливою мішенню для генної терапії. Гемофілія A являє собою спадкове порушення гемостазу, яке викликається недостатністю або мутацією Фактора VIII (FVIII). Його успадковування пов'язане зі статтю, причому майже всі пацієнти є чоловіками. Кровотеча звичайна відбувається в суглоби. Кровотеча в м'язи, слизову тканину і центральну нервову систему (ЦНС) є нетиповою, але може відбуватися. Тяжкість захворювання є зворотно пропорційною рівню FVIII: менше ніж 1 % (<0,01 МО/мл) приводить до тяжкого захворювання, із кровотечею після мінімального ушкодження; між 1-5 % (0,01 МО/мл-0,05 МО/мл) викликає помірне захворювання, із кровотечею після легкого ушкодження; і більше ніж 5 % (>0,05 МО/мл) викликає легке захворювання, із кровотечею тільки після значної травми або хірургічного втручання.

Відповідно, існує потреба у векторі для генної терапії, що уможливить уникнути однієї або декількох із проблем, описаних вище.

КОРОТКИЙ ОПИС СУТІ ВИНАХОДУ

Автори даного винаходу розробили лентивірусний вектор, який був псевдотипований гемаглютинин-нейрамінідазою (HN) і гібридними (F) білками з респіраторного параміксовірусу, що містить промотор і трансген. Звичайно кістяк вектора складається з вірусу імунодефіциту мавп (SIV), такого як SIV1 або вірусу SIV імунодефіциту африканської зеленої мавпи (SIV-AGM).

Переважно кістяк вірусного вектора за винаходом береться з SIV-AGM. HN і F-білки функціонують, відповідно, щоб приєднатися до сіалових кислот, і опосередковують злиття клітин для входження вектора в клітини-мішені. Автори даного винаходу виявили, що цей специфічно F/HN-псевдотипований лентівірусний вектор може ефективно трансдукуватися в епітелій дихальних шляхів, приводячи в результаті до експресії трансгена, перманентної протягом періодів, що знаходяться за межами передбачуваної тривалості життя епітеліальних клітин дихальних шляхів. Важливо відзначити, що, автори даного винаходу також виявили, що повторне введення не приводить до втрати ефективності дії. Ці ознаки роблять вектори згідно із даним винаходом привабливими кандидатами для лікування захворювань за допомогою їхнього застосування в експресії терапевтичних білків: (i) усередині клітин дихальних шляхів; (ii) секретованих у просвіт дихальних шляхів; і (iii) секретованих у систему кровообігу.

Даний винахід спрямований на задоволення однієї або декількох із вказаних вище потреб і стосується лентівірусних векторів, псевдотипованих з використанням гемаглютинин-нейрамінідази (HN) і гібридних (F) білків з респіраторного параміксовірусу, що містять промотор і трансген. В одному варіанті здійснення, промотор, переважно, являє собою гібридний енхансер людського CMV/промотор EF1a (hCEF). Даний винахід також стосується способів одержання вказаних векторів, композицій, що містять вказані вектори, і їхнього застосування в терапії.

Вектори згідно із даним винаходом забезпечують більш високу і перманентну експресію гена за допомогою ефективного перенесення генів. Ідентифіковані вище проблеми зважуються даним винаходом, що надає F/HN-псевдотиповані лентівірусні вектори, які мають здатність до: (i) трансдукції у повітряноносі шляхи без руйнування епітеліальної цілісності; (ii) персистувальної експресії гена; (iii) відсутності хронічної токсичності; і (iv) ефективного повторного введення. Довгострокова/персистувальна стабільна експресія гена, переважно на терапевтично ефективному рівні, може досягатися з використанням повторних доз вектора даного винаходу. Альтернативно, щоб досягти бажаної довгострокової експресії, може застосовуватися однократна доза.

На відміну від відомих лентівірусних векторів, лентівірусні вектори за винаходом виявляють ефективне включення в клітини дихальних шляхів, збільшену експресію трансгена, і не страждають втратою ефективності при повторному введенні.

Таким чином, переважно, лентівірусні вектори згідно із даним винаходом можуть використовуватися в генній терапії. Як приклад, властивості ефективного включення векторів за винаходом в клітини дихальних шляхів роблять їх у високому ступені придатними для лікування захворювань дихальних шляхів. Лентівірусні вектори за винаходом можуть також використовуватися в методах генної терапії, щоб стимулювати секрецію терапевтичних білків. Як додатковий приклад, винахід стосується секреції терапевтичних білків у просвіт дихальних шляхів або системи кровообігу. Таким чином, введення вектора за винаходом і його включення в клітини дихальних шляхів можуть забезпечити використання легень (або носа або дихальних шляхів) як "фабрики" для виробництва терапевтичного білка, який потім секретується і потрапляє в загальний кровообіг при терапевтичних рівнях, де він може переміщуватися до клітин/тканин, які представляють інтерес, щоб зробити терапевтичний вплив. На відміну від внутрішньоклітинних або мембранних білків, вироблення таких секретованих білків не ґрунтуються на специфічних для захворювання клітинах-мішенях, які є трансдукованими, що є значною перевагою, і досягає високих рівнів експресії білка. Таким чином, інші захворювання, які не є захворюваннями дихальних шляхів, такі як серцево-судинні захворювання і захворювання крові, особливо, порушення, пов'язані з недостатньою зсілістю крові, можуть також піддаватися лікуванню векторами даного винаходу.

Як приклад, альфа-1 антитрипсин (A1AT) являє собою секретовану анти-протеазу, яка продукується в основному в печінці і потім переміщується в легені, причому більш дрібні кількості також продукуються в самій легені. Основною функцією A1AT є зв'язувати і нейтралізувати/інгібувати нейтрофіл-еластазу. Генна терапія з використанням A1AT, згідно із даним винаходом, має відношення до пацієнта з недостатністю A1AT, а також пацієнта з іншими захворюваннями легень, такими як муковісцидоз або хронічне обструктивне захворювання легень (ХОЗЛ), і дає можливість перебороти деякі проблеми, з якими зіштовхується ферментозамісна терапія.

Автори даного винаходу раніше продемонстрували, що існує значна кореляція між нейтрофіл-еластазою (NE) і A1AT у зразках мокротиння від пацієнтів з муковісцидозом, що показує, що організм виробляє A1AT у відповідь на стимуляцію NE. Автори даного винаходу також показали, що існує статистично значима кореляція між NE і показником легеневого очищення, маркером захворювання дрібних дихальних шляхів, означаючи, що збільшена NE

має негативний вплив на функцію легень. Як показано в даному описі, автори винаходу в даний час несподівано продемонстрували, що лентивірусні вектори за винаходом можуть досягати високих концентрацій A1AT і довгострокової (щонайменше, 90 днів) експресії A1AT *in vivo*. Таким чином, генна терапія з використанням A1AT може нейтралізувати NE, поліпшуючи функцію легень пацієнтів з муковісцидозом і/або ХОЗЛ (і маючи терапевтичний ефект для інших показань, як описано в даному документі). Відповідно, даний винахід стосується застосування лентивірусного вектора, описаного в даному описі, для введення трансгена A1AT і генної терапії станів, що включають, але не обмежених перерахованими, недостатність A1AT, муковісцидоз і/або ХОЗЛ. Уведення лентивірусного A1AT безпосередньо в назальний епітелій і/або легеню може перебороти кілька обмежень, з якими в даний час зіштовхується ферментозамісна терапія (A1AT, який виділений з людської крові, і вводиться внутрішньовенно щотижня), забезпечуючи стабільну, тривалу експресію в тканині-мішені (легеня/назальний епітелій), простоту введення і необмежену доступність.

У декількох варіантах здійснення, трансдукція з використанням лентивірусного вектора за винаходом приводить до секреції рекомбінантного білка в просвіт легені, а також у кровообіг. Одним сприятливим ефектом цього є те, що терапевтичний білок досягає інтерстиціальної тканини. У випадку недостатності A1AT, цей ефект є переважним, оскільки інгібування NE також потрібне на даній ділянці. Генна терапія A1AT може, отже, також бути сприятливою при показаннях для інших захворювань, необмежувальні приклади яких включають діабет типу 1 і типу 2, гострий інфаркт міокарда, ішемічну хворобу серця, ревматоїдний артрит, запальне захворювання кишечника, відторгнення трансплантата, реакцію трансплантат проти хазяїна (GvH), розсіяний склероз, захворювання печінки, цироз, васкуліти й інфекції, такі як бактеріальні і/або вірусні інфекції.

A1AT має численні інші протизапальні і тканинозахисні ефекти, наприклад, на доклінічних моделях діабету, реакції трансплантат проти хазяїна і запального захворювання кишечника. Вироблення A1AT у легенях і/або носі після трансдукції, згідно із даним винаходом, може, отже, бути більш широко застосовною, із включенням цих показань.

Інші приклади захворювань, що можуть піддаватися лікуванню з використанням генної терапії секретованим білком, згідно із даним винаходом, включають серцево-судинні захворювання і захворювання крові, особливо, порушення, пов'язані з недостатньою зсілістю крові, такі як гемофілія (A і B), хвороба фон Вілебранда і недостатність Фактора VII.

У декількох варіантах здійснення, гемофілія A може піддаватися лікуванню згідно із даним винаходом. Тяжкість захворювання є зворотно пропорційною рівню FVIII, і збільшення FVIII, яке дорівнює 2-5 % (0,02-0,05 МО/мл) є достатнім, щоб бути терапевтично ефективним.

У декількох варіантах здійснення ніс є кращою ділянкою продукування для терапевтичного білка з використанням вектора за винаходом для генної терапії внаслідок щонайменше однієї з наступних причин: (i) позаклітинні бар'єри, такі як запальні клітини і мокротиння, є найменш вираженими в носі; (ii) простота введення вектора; (iii) більш дрібні кількості необхідного вектора; і (iv) етичні розуміння. Таким чином, трансдукція назальних епітеліальних клітин з використанням лентивірусного вектора за винаходом може приводити до ефективної (на високому рівні) і тривалої експресії терапевтичного трансгена, що представляє інтерес.

Вектори згідно із даним винаходом забезпечують довгострокову експресію гена, що приводить до довгострокової експресії терапевтичного білка. Як описано в даному описі, фрази "довгострокова експресія", "перманентна експресія" і "персистувальна експресія" використовують взаємозамінно. Довгострокова експресія, згідно із даним винаходом, означає експресію терапевтичного гена і/або білка, переважно, на терапевтичних рівнях, протягом щонайменше 45 днів, щонайменше 60 днів, щонайменше 90 днів, щонайменше 120 днів, щонайменше 180 днів, щонайменше 250 днів, щонайменше 360 днів, щонайменше 450 днів, щонайменше 730 днів або більше. Переважно, довгострокова експресія означає експресію протягом щонайменше 90 днів, щонайменше 120 днів, щонайменше 180 днів, щонайменше 250 днів, щонайменше 360 днів, щонайменше 450 днів, щонайменше 720 днів або більше, більш переважно, щонайменше 360 днів, щонайменше 450 днів, щонайменше 720 днів або більше. Ця довгострокова експресія може бути досягнута за допомогою повторних доз або однократної дози.

Повторювані дози можуть уводитися двічі на день, щодня, двічі на тиждень, щотижня, щомісяця, кожні два місяці, кожні три місяці, кожні чотири місяці, кожних шість місяців, щорічно, кожні два роки або більше. Дозування може продовжуватися настільки довго, наскільки потрібно, наприклад, протягом щонайменше шести місяців, щонайменше одного року, двох років, трьох років, чотирьох років, п'яти років, десяти років, п'ятнадцяти років, двадцяти років або більше, аж до всього часу життя пацієнта, що підлягає лікуванню.

Лентивірусні вектори, такі як вектори за винаходом, можуть інтегруватися в геном трансдукованих клітин і приводити до довгострокової експресії, що робить їхній придатними для трансдукції стовбурових клітин/клітин-попередників. У легенях, деякі типи клітин з регенеративною здатністю були ідентифіковані, як відповідальні за підтримку специфічних послідовностей клітинних поколінь у дихальних шляхах і альвеолах. Такі типи включають базальні клітини і підслизові клітини протоки залози у верхніх дихальних шляхах, клітини Клара і нейроендокринні клітини в бронхіальних повітроносних шляхах, бронхоальвеолярні стовбурові клітини в термінальних бронхіолах і пневмоцити типу II в альвеолах. Отже, і без обмеження теорією, вважають, що вектори згідно із даним винаходом викликають довгострокову генну експресію трансгена, що представляє інтерес, за допомогою введення трансгена в одну або декілька довгожиттєвих епітеліальних клітин або один або кілька типів клітин дихальних шляхів, таких як базальні клітини і підслизові клітини протоки залози верхніх дихальних шляхів, клітини Клара і нейроендокринні клітини в бронхіальних повітроносних шляхах, бронхоальвеолярні стовбурові клітини в термінальних бронхіолах і пневмоцити типу II в альвеолах.

Відповідно, лентивірусні вектори за винаходом можуть трансдукувати одну або кілька клітин або ліній клітин з регенеративним потенціалом усередині легені (включаючи дихальні шляхи) для досягнення довгострокової експресії гена. У переважному варіанті здійснення, лентивірусний вектор винаходу трансдукує базальні клітини, такі як клітини верхніх дихальних шляхів/дихальних шляхів. Базальні клітини відіграють головну роль у процесах епітеліальної підтримки і відновлення після ушкодження. Крім того, базальні клітини є широко розподіленими уздовж людського респіраторного епітелію, з відносним розподілом від 30 % (великі повітроносні шляхи) до 6 % (дрібні повітроносні шляхи).

Лентивірусні вектори за винаходом можуть використовуватися для трансдукції ізольованих і розмножених стовбурових клітин/клітин-попередників *ex vivo* до введення пацієнту. Переважно, лентивірусні вектори за винаходом використовують для трансдукції клітин у легенях (або дихальних шляхах) *in vivo*.

Вектори згідно із даним винаходом забезпечують високі рівні експресії гена, що приводять у результаті до високих рівнів (переважно, терапевтичних рівнів) експресії терапевтичного білка. Експресію можна виміряти будь-яким придатним методом (якісним або кількісним, переважно кількісним), і концентрації можуть бути приведені в будь-яких відповідних одиницях вимірювання, наприклад, нг/мл. Високий рівень експресії згідно із даним винаходом може означати експресію терапевтичного гена і/або білка при концентрації, яка дорівнює щонайменше 10 нг/мл, щонайменше 20 нг/мл, щонайменше 30 нг/мл, щонайменше 40 нг/мл, щонайменше 50 нг/мл, щонайменше 60 нг/мл, щонайменше 70 нг/мл, щонайменше 80 нг/мл, щонайменше 90 нг/мл, щонайменше 100 нг/мл, щонайменше 200 нг/мл, щонайменше 300 нг/мл, щонайменше 400 нг/мл, щонайменше 500 нг/мл, щонайменше 600 нг/мл, щонайменше 700 нг/мл, щонайменше 800 нг/мл, щонайменше 900 нг/мл, щонайменше 1000 нг/мл, щонайменше 2000 нг/мл, щонайменше 3000 нг/мл, щонайменше 4000 нг/мл, щонайменше 5000 нг/мл, щонайменше 10000, щонайменше 15000 нг/мл, щонайменше 20000 нг/мл або більше. Терапевтична експресія може бути задана з використанням таких же значень.

Лентивірусні вектори згідно із даним винаходом звичайно забезпечують високі рівні експресії трансгена при введенні пацієнту. Терміни висока експресія і терапевтична експресія в даному описі використовують взаємозамінно.

Високий рівень експресії згідно із даним винаходом може означати експресію терапевтичного гена і/або білка при концентрації, яка дорівнює щонайменше приблизно 100 нМ, щонайменше приблизно 200 нМ, щонайменше приблизно 300 нМ, щонайменше приблизно 400 нМ, щонайменше приблизно 500 нМ, щонайменше приблизно 600 нМ, щонайменше приблизно 700 нМ, щонайменше приблизно 800 нМ, щонайменше приблизно 900 нМ, щонайменше приблизно 1 мкМ, щонайменше приблизно 1,1 мкМ, щонайменше приблизно 1,2 мкМ, щонайменше приблизно 1,3 мкМ, щонайменше приблизно 1,4 мкМ, щонайменше приблизно 1,5 мкМ, щонайменше приблизно 2 мкМ, щонайменше приблизно 3 мкМ, щонайменше приблизно 4 мкМ, щонайменше приблизно 5 мкМ, щонайменше приблизно 6 мкМ, щонайменше приблизно 7 мкМ, щонайменше приблизно 8 мкМ, щонайменше приблизно 9 мкМ, щонайменше приблизно 10 мкМ, щонайменше приблизно 11 мкМ, щонайменше приблизно 12 мкМ, щонайменше приблизно 13 мкМ, щонайменше приблизно 14 мкМ, щонайменше приблизно 15 мкМ, щонайменше приблизно 20 мкМ, щонайменше приблизно 25 мкМ, щонайменше приблизно 30 мкМ, щонайменше приблизно 40 мкМ, щонайменше приблизно 50 мкМ, щонайменше приблизно 75 мкМ, або, щонайменше приблизно 100 мкМ або більше. Терапевтична експресія може бути задана з використанням цих же самих значень.

Високий рівень експресії, згідно із даним винаходом, може означати експресію терапевтичного гена (звичайно вимірювану по експресії мРНК), яка дорівнює щонайменше приблизно 1 %, щонайменше приблизно 2 %, щонайменше приблизно 3 %, щонайменше приблизно 4 %, щонайменше приблизно 5 %, щонайменше приблизно 6 %, щонайменше приблизно 7 %, щонайменше приблизно 8 %, по меншій мері приблизно 9 %, щонайменше приблизно 10 %, щонайменше приблизно 15 %, щонайменше приблизно 20 % або більше у порівнянні з рівнем експресії відповідної ендогенної (дефектної) мРНК. Терапевтична експресія може бути задана з використанням тих же самих значень. Наприклад, типовий рівень експресії мРНК ендогенного CFTR може бути кількісно оцінений з позицій числа копій мРНК на клітину легені, наприклад, одна копія мРНК ендогенного CFTR на клітину легені, дві копії мРНК ендогенного CFTR на клітину легені, три копії мРНК ендогенного CFTR на клітину легені, чотири копії мРНК ендогенного CFTR на клітину легені, п'ять копій мРНК ендогенного CFTR на клітину легені, або більш, переважно дві копії мРНК ендогенного CFTR на клітину легені. Експресія терапевтичного гена винаходу, такого як функціональний ген CFTR, може бути кількісно визначена відносно ендогенного гена, такого як ендогенні (дисфункціональні) гени CFTR з позицій числа копій мРНК на клітину або будь-якої іншої відповідної одиниці вимірювання.

Високий рівень експресії, згідно із даним винаходом, може означати експресію терапевтичного гена і/або білка при концентрації, яка дорівнює щонайменше приблизно 0,5 %, щонайменше приблизно 1 %, щонайменше приблизно 2 %, щонайменше приблизно 3 %, щонайменше приблизно 4 %, щонайменше приблизно 5 %, щонайменше приблизно 6 %, щонайменше приблизно 7 %, щонайменше приблизно 8 %, щонайменше приблизно 9 %, щонайменше приблизно 10 %, щонайменше приблизно 15 %, щонайменше приблизно 20 %, щонайменше приблизно 25 %, щонайменше приблизно 30 %, щонайменше приблизно 40 %, щонайменше приблизно 50 %, щонайменше приблизно 60 %, щонайменше приблизно 70 %, щонайменше приблизно 80 %, щонайменше приблизно 90 %, щонайменше 95 %, щонайменше приблизно 96 %, щонайменше приблизно 97 %, щонайменше приблизно 98 %, щонайменше приблизно 99 % або більше у порівнянні з рівнем терапевтичного гена і/або білка дикого типу, де рівень дикого типу являє собою рівень у нормального індивідуума без захворювання. У декількох варіантах здійснення, експресію дикого типу приймають за 100 %, причому будь-яке поліпшення експресії гена вимірюють відносно цього значення. Як необмежувальний приклад, якщо в нормального індивідуума без захворювання експресію функціонального гена приймають за 100 %, і в індивідуума з захворюванням, експресія функціонального гена дорівнює 0 %, терапевтичний рівень експресії гена або білка може становити щонайменше приблизно 0,5 %, щонайменше приблизно 1 %, щонайменше приблизно 2 %, щонайменше приблизно 3 %, щонайменше приблизно 4 %, щонайменше приблизно 5 %, щонайменше приблизно 6 %, щонайменше приблизно 7 %, щонайменше приблизно 8 %, щонайменше приблизно 9 %, щонайменше приблизно 10 %, щонайменше приблизно 15 %, щонайменше приблизно 20 %, щонайменше приблизно 25 %, щонайменше приблизно 30 %, щонайменше приблизно 40 %, щонайменше приблизно 50 % або більше у порівнянні з рівнем для терапевтичного гена і/або білка дикого типу. Як ще один необмежувальний приклад, якщо в нормального індивідуума без захворювання експресію функціонального гена приймають за 100 %, і в індивідуума з захворюванням, експресія функціонального гена становить 50 %, терапевтичний рівень експресії гена або білка може становити щонайменше приблизно 55 %, щонайменше приблизно 60 %, щонайменше приблизно 70 %, щонайменше приблизно 80 %, щонайменше приблизно 90 %, щонайменше 95 %, щонайменше приблизно 96 %, щонайменше приблизно 97 %, щонайменше приблизно 98 %, щонайменше приблизно 99 % або більше у порівнянні з рівнем для терапевтичного гена і/або білка дикого типу.

Для секретованих білків, таких як A1AT, звичайно концентрація в легені або бронхоальвеолярній лаважній рідині (виміряна з використанням BAL) є приблизно десятикратною відносно концентрації в сироватці. Як необмежувальний приклад, якщо концентрація секретованого білка в легені або бронхоальвеолярній лаважній рідині знаходиться в області 750 нг/мл, сироваткова концентрація білка знаходиться в області 75 нг/мл.

Рівні експресії терапевтичного гена і/або білка винаходу можуть бути виміряні в тканині легені, бронхоальвеолярній лаважній рідині і/або сироватці/плазмі відповідним чином. Високий і/або терапевтичний рівень експресії може, отже, стосуватися концентрації в легені, бронхоальвеолярній лаважній рідині і/або сироватці/плазмі.

Як необмежувальний приклад, терапевтичний рівень експресії CFTR звичайно становить 1-5 % від терапевтичного рівня мРНК CFTR у порівнянні з рівнем експресії ендогенної (дефектної) мРНК CFTR.

Як ще один необмежувальний приклад, терапевтичний рівень експресії A1AT звичайно дорівнює, щонайменше приблизно 1 мкМ у бронхоальвеолярній лаважній рідині, і/або, щонайменше приблизно 0,1 мкМ у сироватці. У переважному варіанті здійснення, терапевтичний рівень експресії A1AT у бронхоальвеолярній лаважній рідині дорівнює, щонайменше приблизно 1,1 мкМ, і/або терапевтичний сироватковий рівень експресії A1AT, згідно із даним винаходом, становить, щонайменше приблизно 11 мкМ. Як ще один необмежувальний приклад, терапевтичний рівень експресії A1AT у бронхоальвеолярній лаважній рідині (ELF, тобто промивній рідині повітроносних шляхів і повітряного простору в легенях) дорівнює 70 мкг/мл (у порівнянні з "нормальним" цільовим рівнем АТТ (A1AT) у ELF, який дорівнює 200 мкг/мл).

Як ще один необмежувальний приклад, терапевтичний рівень експресії білка FVIII звичайно становить, щонайменше приблизно 1-3 % або, щонайменше приблизно 1-6 % від рівня експресії в нормального індивідуума, що не страждає на гемофілію.

Терапевтичний ген, включений у вектор винаходу, може бути модифікований, щоб сприяти експресії. Наприклад, послідовність гена може знаходитися в CpG-збідненій і/або кодон-оптимізованій формі, щоб сприяти експресії гена. Стандартні методи для модифікації генної послідовності наступним чином є відомими в даній галузі.

Промотор, включений у вектор за винаходом, може бути специфічно вибраний і/або модифікований, щоб, надалі, оптимізувати регуляцію експресії терапевтичного гена. Також прийнятні промотори і стандартні методи для їхньої модифікації є відомими в даній галузі. Як необмежувальний приклад, ряд придатних (які не містять CpG) промоторів, придатних для застосування в даному винаході описані в Pringle et al. (J. Mol. Med. Berl. 2012, 90(12): 1487-96), яка повністю включена в даний опис за допомогою посилання.

Вектор за винаходом може бути модифікований для забезпечення припинення експресії гена. Стандартні методи модифікації вектора наступним чином є відомими в даній галузі. Як необмежувальний приклад, широко використовують Тет-чутливі промотори.

Вектори згідно із даним винаходом також демонструють чудову стійкість до зусиль зсуву лише з помірним зниженням здатності до трансдукції, при проходженні через клінічно значимі пристрої для доставки, такі як бронхоскопи, аерозольні балончики й інгалятори.

В одному варіанті здійснення, винахід стосується лентивірусних векторів F/ΔN, що містять промотор і трансген, який не має нітрону, розташованого між промотором і трансгеном. В одному варіанті здійснення, вектор згідно із даним винаходом доставляють у клітини дихальних шляхів. У варіанті здійснення, лентивірус являє собою SIV. В одному варіанті здійснення, промотор є гібридом енхансера людського CMV/промотору EF1α (hCEF). Звичайно у вказаному промоторі винаходу відсутній інтрон, який відповідає нуклеотидам 570-709, і екзон, що відповідає нуклеотидам 728-733 промотору hCEF. Переважний приклад послідовності промотору hCEF винаходу представлений SEQ ID NO: 6. Промотор може бути промотором CMV. Приклад послідовності промотору CMV представлений SEQ ID NO: 17. Інші промотори для експресії трансгена є відомими в даній галузі, і їхню придатність для лентивірусних векторів за винаходом визначають з використанням установлених методів, відомих у даній галузі. Необмежувальні приклади інших промоторів включають UbC і UCOE. Як описано в даному документі, промотор може бути модифікований, щоб надалі регулювати експресію трансгена винаходу.

В одному варіанті здійснення, трансген кодує CFTR. Приклад кДНК CFTR представлений SEQ ID NO: 7.

В одному варіанті здійснення, трансген кодує A1AT. Приклад трансгена A1AT представлений SEQ ID NO: 15 або послідовністю, комплементарною SEQ ID NO: 26. SEQ ID NO: 15 є кодон-оптимізованим збідненням по CpG трансгеном A1AT, сконструйованим авторами даного винаходу для посилення трансляції в людських клітинах. Було показано, що така оптимізація підсилює експресію гена аж до 15-кратного посилення. Таким чином, в одному варіанті здійснення, винахід стосується полінуклеотиду, що містить нуклеотидну послідовність або складається з нуклеотидної послідовності з SEQ ID NO: 15. Варіанти такої ж самої послідовності (визначеної в даному описі), що мають такий же технічний ефект посилення трансляції в порівнянні з немодифікованою (дикого типу) послідовністю гена A1AT, також є охопленими даним винаходом. Винахід додатково стосується поліпептиду, який кодується вказаним трансгеном A1AT, ілюстрованому поліпептидом з SEQ ID NO: 27, плазмід (зокрема, плазмід з геномом вектора, визначеного в даному описі) і лентивірусних векторів, що містять вказаний трансген A1AT. У переважному варіанті здійснення, в аспектах винаходу, які стосуються генної терапії A1AT, згідно із даним винаходом, використовують послідовність трансгена A1AT з SEQ ID NO: 15.

В одному варіанті здійснення, трансген кодує FVIII. Приклади трансгена FVIII представлені за допомогою SEQ ID NO: 16 і 30, або за допомогою відповідних послідовностей, комплементарних SEQ ID NO: 28 і 31.

Лентивірусні вектори, що підходять для використання в даному винаході, включають вірус імунodefіциту людини (BІЛ), вірус імунodefіциту мавп (SIV), вірус імунodefіциту котятих (FIV), вірус інфекційної анемії коней (EІAV), і вірус вісна/маеді. В одному варіанті здійснення винаходу, використовують вектор SIV, переважно SIV-AGM. У ще одному варіанті здійснення, використовують вектор BІЛ.

Лентивірусні вектори згідно із даним винаходом є псевдотипованими гемаглютинін-нейрамінідазою (HN) і гібридними (F) білками з респіраторного параміксовірусу. В одному варіанті здійснення, респіраторний параміксовірус являє собою вірус Сендай (вірус мишачого парогрипу типу 1).

В одному варіанті здійснення винаходу, лентивірусний вектор є інтеграза-компетентним (IC). В альтернативному варіанті здійснення, лентивірусний вектор є інтеграза-дефіцитним (ID).

У ще одному іншому варіанті здійснення винаходу, трансген за винаходом являє собою будь-який один або декілька з DNAH5, DNAH11, DNAI1 і DNAI2 або інший відомий споріднений ген.

В одному варіанті здійснення винаходу, мішенню для доставки вектора є епітелій дихальних шляхів. У цьому варіанті здійснення, трансген кодує альфа-1 антитрипсин (A1AT), поверхнево-активний білок В (SFTPВ) або гранулоцитарно-макрофагальний колонієстимулюючий фактор (GM-CSF). У ще одному іншому варіанті здійснення, трансген кодує моноклональне антитіло (mAb) до збудника інфекції. В одному варіанті здійснення, трансген кодує антитіло до ФНО-альфа. У додатковому варіанті здійснення, трансген кодує терапевтичний білок, що бере участь у запальному, імунному або метаболічному стані.

В одному варіанті здійснення винаходу, вектор доставляють у клітини дихальних шляхів, щоб забезпечити вироблення білків, що підлягають секреції в систему кровообігу. У цьому варіанті здійснення, трансген кодує Фактор VII, Фактор VIII, Фактор IX, Фактор X, Фактор XI і/або фактор фон Вілебранда. Такий вектор може використовуватися в лікуванні захворювань, особливо, серцево-судинних захворювань і захворювань крові, переважно, розладів, пов'язаних з недостатньою зсілістю крові, таких як гемофілія. У ще одному варіанті здійснення, трансген кодує моноклональне антитіло (mAb) проти збудника інфекції. В одному варіанті здійснення, трансген кодує білок, що бере участь у запальному, імунному або метаболічному стані, такому як хвороба лізосомального нагромадження.

Відповідно до винаходу, він стосується лентивірусного вектора F/HN-SIV, що містить промотор hCEF і трансген CFTR, що не має інтрона, розташованого між промотором і трансгеном. Аналогічно, відсутній інтрон між промотором і трансгеном у геномі вектора (пДНК1) плазміді (наприклад, pGM326, описаної в даному документі, яка ілюстрована на Фіг. 1A і з послідовністю SEQ ID NO: 1).

Винахід також стосується лентивірусного вектора F/HN-SIV, що містить промотор hCEF і трансген A1AT, що не має інтрона, розташованого між промотором і трансгеном. Такий лентивірусний вектор може бути отриманий за допомогою способу, описаного в даному описі, з використанням плазміді, що несе трансген A1AT і промотор. Аналогічно, відсутній інтрон між промотором і трансгеном A1AT у геномі вектора (пДНК1) плазміді. Ілюстративна послідовність такий плазміді приведена в SEQ ID NO: 9(F/HN-SIV-hCEF-so1AT, яка ілюстрована на Фіг. 15A).

Винахід також стосується лентивірусного вектора F/HN-SIV, що містить (i) промотор hCEF або промотор CMV; і (ii) трансген FVIII; де відсутній інтрон між промотором і трансгеном. Такий лентивірусний вектор може бути отриманий за допомогою способу, описаного в даному описі, з використанням плазміді, що несе трансген FVIII і промотор. Аналогічно, відсутній інтрон між промотором і трансгеном FVIII у геномі вектора (пДНК1) плазміді. Ілюстративні послідовності таких плазмід приведені в SEQ ID NO: 11-14 (ілюстрованих на Фіг. 22A-E).

Лентивірусний вектор, описаний вище, містить трансген. Трансген містить послідовність нуклеїнової кислоти, що кодує генний продукт, наприклад, білок.

Наприклад, в одному варіанті здійснення, послідовність нуклеїнової кислоти, що кодує CFTR, A1AT або FVIII, містить послідовність нуклеїнової кислоти (або складається з послідовності нуклеїнової кислоти), що має щонайменше 90 % (таким чином, як, щонайменше 90, 92, 94, 95, 96, 97, 98, 99 або 100 %) ідентичності послідовності з послідовністю нуклеїнової кислоти CFTR, A1AT або FVIII, відповідно. У додатковому варіанті здійснення, послідовність нуклеїнової кислоти, що кодує CFTR, A1AT або FVIII, містить послідовність нуклеїнової кислоти (або складається з послідовності нуклеїнової кислоти), що має щонайменше 95 % (таким чином, як, щонайменше 95, 96, 97, 98, 99 або 100 %) ідентичності послідовності з послідовністю

нуклеїнової кислоти CFTR, A1AT або FVIII, відповідно. В одному варіанті здійснення, послідовність нуклеїнової кислоти, що кодує CFTR, представлена SEQ ID NO: 7, послідовність нуклеїнової кислоти, що кодує A1AT, представлена SEQ ID NO: 15, або послідовністю, комплементарною SEQ ID NO: 26 і/або послідовність нуклеїнової кислоти, що кодує FVIII представлена SEQ ID NO: 16 і 30, або за допомогою відповідних послідовностей, комплементарних SEQ ID NO: 28 і 31, або їхніх варіантів.

Термін "поліпептид", використовуваний у даному описі, також охоплює варіантні послідовності. Таким чином, поліпептид, який кодується трансеном винаходу, може мати щонайменше 90 % (наприклад, щонайменше 90, 92, 94, 95, 96, 97, 98, 99 або 100 %) ідентичності послідовності з функціональною поліпептидною послідовністю CFTR, A1AT або FVIII, відповідно. У додатковому варіанті здійснення, амінокислотна послідовність CFTR, A1AT або трансгена FVIII містить амінокислотну послідовність (або складається з амінокислотної послідовності), що має щонайменше 95 % (наприклад, щонайменше 95, 96, 97, 98, 99 або 100 %) ідентичності послідовності з функціональною поліпептидною послідовністю CFTR, A1AT або FVIII, відповідно. В одному варіанті здійснення, амінокислотна послідовність білка A1AT, який кодується трансеном винаходу, містить амінокислотну послідовність (або складається з амінокислотної послідовності) з SEQ ID NO: 27, або її варіанти. Переважно, вказаний варіант білків A1AT за винаходом має щонайменше 90 % (наприклад, щонайменше 90, 92, 94, 95, 96, 97, 98, 99 або 100 %), більш переважно, щонайменше 95 % або більш, ідентичності послідовності з SEQ ID NO: 27.

В одному варіанті здійснення, послідовність нуклеїнової кислоти, що кодує CFTR, A1AT або FVIII містить послідовність ДНК (або складається з послідовності ДНК), що є комплементарною послідовності ДНК CFTR, A1AT або FVII, відповідно. В одному варіанті здійснення, CFTR, A1AT або трансген FVIII представляє CFTR, A1AT або FVIII з оптимізованою послідовністю (soCFTR2, so1AT або FVIII). Приклад представлений SEQ ID NO: 7, 15 і 16 відповідно. Ілюстративна послідовність, комплементарна послідовності A1AT з оптимізованою послідовністю, приведена в SEQ ID NO: 26. Ілюстративні послідовності, комплементарні оптимізованим послідовностям FVIII, приведені в SEQ ID NO: 28 і 31.

В одному варіанті здійснення винаходу, вектор F/HN експресії трансгена спрямовується промотором цитомегаловірусу (CMV). У ще одному варіанті здійснення, вектор експресії трансгена спрямовується промотором фактора елонгації 1a (EF1a). У переважному варіанті здійснення, вектор експресії трансгена спрямовується гібридом енхансера людського CMV/промотору EF1a (hCEF). В одному варіанті здійснення, промотор hCEF має заміну всіх динуклеотидів CG, на будь-який один з AG, TG або GT. Таким чином, в одному варіанті здійснення, промотор hCEF є таким, що не містить CpG.

В одному варіанті здійснення, лентивірусний вектор може бути отриманий з використанням плазміди F/HN-SIV-hCEF-soCFTR2-IC. У цьому варіанті здійснення, CFTR експресується під керуванням промотору hCEF. Цей лентивірусний вектор може бути описаний, як такий, що містить F/HN-SIV-hCEF-soCFTR2-IC, тому що він містить F/HN елементи SIV, а також компетентну по інтеграції касету експресії, що містить CFTR під керуванням промотору hCEF. Цей лентивірусний вектор за винаходом здатний робити довгостроковий, повторюваний, високий рівень експресії в клітинах дихальних шляхів, без індукування надмірної імунної відповіді. Отже, винахід стосується ефективного засобу для генної терапії *in vivo*, наприклад, для перенесення гена CFTR у легеню, уражену муковісцидозом, для лікування муковісцидозу легень.

У переважному варіанті здійснення, лентивірусний вектор може бути отриманий з використанням плазміди F/HN-SIV-hCEF-so1AT. У цьому варіанті здійснення, A1AT експресується під керуванням промотору hCEF. Даний лентивірусний вектор може бути описаний, як такий, що містить F/HN-SIV-hCEF-so1AT, тому що він містить F/HN елементи SIV, а також касету експресії, що містить A1AT під керуванням промотору hCEF. Цей лентивірусний вектор за винаходом здатний робити довгостроковий, повторюваний, високий рівень експресії в клітинах дихальних шляхів, без індукування надмірної імунної відповіді. Отже, винахід стосується ефективного засобу для генної терапії *in vivo*, наприклад, для перенесення гена A1AT у легеню або ніс пацієнта для вироблення A1AT, що потім секретується в систему кровообігу (як описаний у даному документі). Таким чином, цей вектор і інші вектори за винаходом, що містять трансген A1AT, можуть використовуватися в лікуванні недостатності A1AT, або при інших показаннях, як описано в даному описі.

У ще одному переважному варіанті здійснення, лентивірусні вектори можуть бути отримані з використанням F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-ко і/або F/HN-SIV-hCEF-HFVIII-N6-коплазмід. HFVIII стосується людського FVIII. У даному варіанті

здійснення, FVIII експресується під керуванням промотору hCEF або промотору CMV. Ці лентивірусні вектори можуть бути описані, як такі, що містять F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-ко і F/HN-SIV-hCEF-HFVIII-N6-ко, відповідно, тому що вони включають F/HN елементи SIV, а також касету експресії, що містить FVIII під керуванням промотору hCEF/промотору CMV. Вірусні векторні продукти, одержувані з використанням F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-3 i/або F/HN-SIV-hCEF-HFVIII-N6-коплазміди є також відомими також, як vGM126, vGM127, vGM142 і vGM129 (див. Фігуру 22). Ці лентивірусні вектори за винаходом здатні виробляти довгостроковий, повторюваний, високий рівень експресії в клітинах дихальних шляхів, без індукування надмірної імунної відповіді. Отже, винахід надає ефективний засіб генної терапії in vivo, наприклад, для перенесення гена FVIII у легень або ніс пацієнта для вироблення FVIII, що потім секритується в систему кровообігу (як описано в даному описі). Таким чином, ці вектори й інші вектори за винаходом, що містять трансген FVIII, можуть використовуватися в лікуванні гемофілії або при інших показаннях, як описано в даному описі.

Лентивірусні вектори за винаходом не містять інтрон між промотором і трансгеном. Аналогічно, геном вектора плазмід за винаходом (використовуваний для генерації, вказаних лентивірусних векторів, описаних у даному описі) також не містять інтрон між промотором і трансгеном. Винахід, отже, стосується, в одному варіанті здійснення, відсутності інтрона між промотором hCEF і кодувальними послідовностями, які підлягають експресії. В одному переважному варіанті здійснення, кодувальна послідовність, яка підлягає експресії, являє собою послідовність нуклеїнової кислоти CFTR, A1AT i/або FVIII.

В одному варіанті здійснення, вектори за винаходом містять центральний поліпуриновий шлях (сРРТ) і посттранскрипційні регуляторні елементи вірусу гепатиту бабаків (WPRE). В одному варіанті здійснення, послідовність WPRE представлена SEQ ID NO: 8.

В одному варіанті здійснення вектор за винаходом використовують для генної терапії. В одному варіанті здійснення захворювання, яке підлягає лікуванню, являє собою МБ. У ще одному іншому варіанті здійснення винаходу, захворювання, яке підлягає лікуванню, являє собою первинну циліарну дискінезію (PCD). В одному варіанті здійснення, вектор застосовують для лікування гострого ушкодження легень. В одному варіанті здійснення винаходу, захворювання, яке підлягає лікуванню, являє собою недостатність поверхнево-активного білка В (ПАБ-В), недостатність альфа-1-антитрипсину (A1AD), легеневий альвеолярний протеїноз (РАР), хронічне обструктивне захворювання легень (ХОЗЛ). У ще одному іншому варіанті здійснення, захворювання являє собою запальний, імунний або метаболічний стан.

Захворювання, яке підлягає лікуванню, може бути серцево-судинним захворюванням або хворобою крові, особливо, недостатньою зсілістю крові. Таким чином, у декількох варіантах здійснення, захворювання, яке підлягає лікуванню, являє собою гемофілію А, гемофілію В або гемофілію С, недостатність Фактора VII i/або хворобу фон Вілебранда. У ще одному іншому варіанті здійснення, захворювання, яке підлягає лікуванню, являє собою запальне захворювання, інфекційне захворювання або метаболічний стан, такий як хвороба лізосомального нагромадження.

Необмежувальні приклади захворювань, що можуть піддаватися лікуванню з використанням генної терапії A1AT, згідно із даним винаходом, включають діабет типу 1 і типу 2, гострий інфаркт міокарда, ішемічну хворобу серця, ревматоїдний артрит, запальне захворювання кишечника, відторгнення трансплантата, реакцію трансплантат проти хазяїна (GvH), розсіяний склероз, захворювання печінки, цироз, васкуліти й інфекції, такі як бактеріальні i/або вірусні інфекції.

В одному аспекті винаходу, за допомогою вектора можна ефективно лікувати захворювання, за допомогою одержання трансгена для корекції захворювання. Наприклад, за допомогою уведення функціональної копії гена CFTR для полегшення протікання або запобігання захворюванню легень у пацієнтів з муковісцидозом, незалежно від мутації, яка лежить в основі.

У ще одному іншому варіанті здійснення винахід стосується способу одержання лентивірусу. У цьому варіанті здійснення, спосіб за винаходом являє собою масштабований, що відповідає вимогам GMP спосіб. Таким чином, спосіб за винаходом забезпечує генерацію очищених F/HN векторів з високим титром.

Спосіб за винаходом включає наступні стадії:

- (a) вирощування клітин у суспензії;
- (b) трансфекція клітин з використанням однієї або декількох плазмід;
- (c) додавання нуклеази;
- (d) збирання лентивірусу;
- (e) додавання трипсину; і

(f) очищення.

В одному варіанті здійснення способу за винаходом, одна або декілька плазмід надають геном вектора, Gag-Pol, Rev, F і HN. Таким чином, можуть використовуватися п'ять плазмід для кожного з генома векторів, Gag-Pol, Rev, F і HN, відповідно. У переважному способі за винаходом з використанням 5 плазмід, плазміда з геномом вектора кодує весь генетичний матеріал, що упакований у кінцевий лентивірусний вектор, що включає трансген. Звичайно тільки частина генетичного матеріалу, виявленого в плазміді з геномом вектора, закінчується у вірусі. Плазміда з геномом вектора може позначатися в даному описі як "пДНК1". Інші чотири плазміди є виробляючими плазмідами, що кодують білки Gag-Pol, Rev, F і HN. Ці плазміди можуть позначатися як "пДНК2a", "пДНК2b", "пДНК3a" і "пДНК3b" відповідно.

В одному варіанті здійснення винаходу, лентивірус являють собою SIV, такий як SIV1, переважно SIV-AGM. В одному варіанті здійснення, білки F і HN є похідними параміксовірусу, такого як вірус Сендай. В одному варіанті здійснення, плазміда з геномом вектора (пДНК1) містить трансген і промотор трансгена.

У конкретному варіанті здійснення, що стосується CFTR, п'ять плазмід характеризуються фігурами 1A-1E, таким чином, пДНК1 є плазмідною рGM326 фігури 1A, пДНК2a є плазмідною рGM299 фігури 1B, пДНК2b є плазмідною рGM299 фігури 1C, пДНК3a є плазмідною рGM301 фігури 1D і пДНК3b є плазмідною рGM303 фігури 1E. У даному варіанті здійснення, остаточний CFTR-вмісний лентивірусний вектор може позначатися як vGM058 (див. Приклади). Вектор vGM058 являє собою переважний варіант здійснення винаходу.

У варіанті здійснення, що стосується A1AT, п'ять плазмід можуть характеризуватися фігурами 15A (таким чином, плазміда пДНК1 може бути рGM407) і 1B-E (як вказано вище для конкретного варіанта здійснення з CFTR).

У варіанті здійснення, що стосується FVIII, п'ять плазмід можуть характеризуватися фігурами 22C-F (у такий спосіб плазміда пДНК1 може являти собою рGM411, рGM412, рGM413 або рGM414) і 1B-E.

У цих варіантах здійснення винаходу, плазміда, визначена на Фіг. 1A, представлена за допомогою SEQ ID NO: 1; плазміда, визначена на Фіг. 1B, представлена за допомогою SEQ ID NO: 2; плазміда, визначена на Фіг. 1C, представлена за допомогою SEQ ID NO: 3; плазміда, визначена на Фіг. 1D, представлена за допомогою SEQ ID NO: 4; плазміда, визначена на Фіг. 1E, представлена за допомогою SEQ ID NO: 5; плазміда, визначена на Фіг. 15A, представлена за допомогою SEQ ID NO: 9 і F/HN-SIV-CMV-HFVIII-V3, F/HN-SIV-hCEF-HFVIII-V3, F/HN-SIV-CMV-HFVIII-N6-ко i/або F/HN-SIV-hCEF-HFVIII-N6-коплазміди, визначені на Фіг. 22B представлені за допомогою SEQ ID NO: 11-14, відповідно.

У 5-плазмідному способі за винаходом всі п'ять плазмід сприяють утворенню остаточного лентивірусного вектора. Під час одержання лентивірусного вектора, плазміда з геномом вектора (пДНК1) забезпечує енхансер/промотор, Psi, RRE, cPPT, mWPPE, SIN LTR, SV40 поліA (див. Фігуру 1A), що є важливими для виробництва вірусу. При використанні рGM326 як необмежувальний приклад пДНК1, CMV енхансер/промотор, SV40 poly, colE1 Ori і KanR беруть участь у продукуванні лентивірусного вектора за винаходом (наприклад, vGM058), але не виявляються в остаточному вірусному векторі. RRE, cPPT (центральний поліпуриновий шлях), hCEF, soCFTR2 (трансген) і mWPPE з рGM326 виявляються в остаточному вірусному векторі. SIN LTR (довгі кінцеві повтори, самоінактивованій SIN/IN) і Psi (сигнал упаковки) можуть виявлятися в остаточному вірусному векторі.

Для інших лентивірусних векторів за винаходом, відповідні елементи з інших плазмід з геномом вектора (пДНК1) вимагаються для одержання (але не виявляються в остаточному векторі), або присутні в остаточному вірусному векторі.

Білки F і HN із пДНК3a і пДНК3b (переважно, білки Sendai F і HN) є важливими для інфікування цільових клітин кінцевим лентивірусним вектором, тобто для входу в епітеліальні клітини пацієнтів (звичайно клітини легень або назальні клітини, описані в даному описі). Продукти плазмід пДНК2a і пДНК2b є важливими для трансдукції вірусу, тобто для введення ленти вірусної ДНК у геном хазяїна. Промотор, регуляторні елементи (такі як WPPE) і трансген є важливими для експресії транс гена в клітині (клітинах) - мішенях.

В одному варіанті здійснення, стадії (a)-(f) виконують послідовно. В одному варіанті здійснення, клітини являють собою клітини HEK293 або клітини 293T/17. В одному варіанті здійснення, клітини вирощують у середовищах, які не містять компонентів із тварин (безсироваткових). В одному варіанті здійснення, трансфекцію виконують за допомогою застосування PEIPro™. В одному варіанті здійснення, нуклеаза являє собою ендонуклеазу, наприклад, Бензоназу®. В одному варіанті здійснення, активність трипсину представлена рекомбінантним ферментом, який не має тваринного походження, таким як TrypLE Select™.

В одному варіанті здійснення винаходу, додавання нуклеази відбувається на стадії до збирання. В альтернативному варіанті здійснення, додавання нуклеази відбувається на стадії після збирання. У ще одному іншому варіанті здійснення, додавання трипсину відбувається на стадії до збирання. У ще одному іншому варіанті здійснення, додавання трипсину відбувається на стадії після збирання.

В одному варіанті здійснення, стадія очищення включає стадію хроматографії. У цьому варіанті здійснення використовують гель-проникну хроматографію (SEC) з комбінованим режимом. В одному варіанті здійснення використовують аніонообмінну хроматографію. У цьому варіанті здійснення, для стадії елювання не використовують ніякого градієнта солі.

В одному варіанті здійснення, цей метод використовують для одержання лентивірусних векторів за винаходом. У цьому варіанті здійснення, вектор за винаходом містить ген CFTR, A1AT і/або FVIII. В альтернативному варіанті здійснення, вектор за винаходом містить будь-який з вказаних вище генів, або генів, що кодують вказані вище білки.

В одному варіанті здійснення способу за винаходом, будь-яку комбінацію з однієї або декількох специфічних конструкцій плазмід, представлених на Фіг. 1A-1E, Фіг. 15A і/або Фіг. 22C-22F, використовують для одержання вектора за винаходом.

Винахід додатково стосується способу лікування захворювання, що включає введення лентивірусного вектора за винаходом пацієнту. У цьому варіанті здійснення, винахід стосується лентивірусного вектора за винаходом для застосування в лікуванні захворювання легень. В одному варіанті здійснення, захворювання є хронічним захворюванням. У конкретному варіанті здійснення, винахід стосується способу лікування МВ. В інших варіантах здійснення, винахід стосується способу лікування первинної циліарної дискінезії (PCD), недостатності поверхнево-активного білка В (ПАБ-В), недостатності альфа-1-антитрипсину (A1AD), легеневого альвеолярного протеїнозу (РАР), хронічного обструктивного захворювання легень (ХОЗЛ). У ще одному іншому варіанті здійснення, захворювання являє собою запальний, імунний або метаболічний стан.

У ще одному іншому варіанті здійснення, захворювання може бути серцево-судинним захворюванням або хворобою крові, особливо, недостатньою зсілістю крові, такою як гемофілія А, гемофілія В, гемофілія С, недостатність Фактора VII і/або хвороба фон Вілебранда, запальне захворювання, інфекційне захворювання або метаболічний стан, такий як хвороба лізосомального нагромадження.

Захворювання може являти собою діабет типу 1 і типу 2, гострий інфаркт міокарда, ішемічну хворобу серця, ревматоїдний артрит, запальне захворювання кишечника, відторгнення трансплантата, реакцію трансплантат проти хазяїна (GvH), розсіяний склероз, захворювання печінки, цироз, васкуліти й інфекції, такі як бактеріальні і/або вірусні інфекції.

Лентивірусні вектори за винаходом можуть вводитися в будь-якому дозуванні, що підходить для досягнення бажаного терапевтичного ефекту. Відповідні дозування можуть визначатися клініцистом або іншим практикуючим лікарем з використанням стандартних методів і в ході їхньої звичайної роботи. Необмежувальні приклади придатних дозувань включають 1×10^8 одиниць трансдукції (ОТ), 1×10^9 ОТ, 1×10^{10} ОТ, 1×10^{11} ОТ або більше.

Винахід також стосується композицій, що містять лентивірусні вектори, описані вище, і фармацевтично прийнятний носій. Необмежувальні приклади фармацевтично прийнятних носіїв включають воду, фізрозчин, і забуферений фосфатом фізрозчин. У декількох варіантах здійснення, однак композиція знаходиться в ліофілізованій формі, і в цьому випадку, вона може включати стабілізатор, такий як альбумін бичачої сироватки (BSA). У декількох варіантах здійснення, може бути бажаним складати композицію разом з консервантом, таким як тимеросал або азид натрію, щоб сприяти довгостроковому збереженню.

Вектори за винаходом можуть уводитися за допомогою будь-якого відповідного шляху. Може бути бажаним спрямовувати композиції згідно із даним винаходом (описані вище) у дихальну систему пацієнта. Ефективна передача терапевтичної/профілактичної композиції або лікарського засобу до ділянки інфекції в дихальному шляху може досягатися за допомогою перорального або інтраназального введення, наприклад, у вигляді аерозолів (наприклад, назальних спреїв), або за допомогою катетерів. Звичайно лентивірусні вектори за винаходом є стабільними в клінічно значимих інгаляторах, катетерах і аерозолях і т. д.

Лікарські форми для інтраназального введення можуть знаходитися у формі назальних крапель або назального спрею. Інтраназальна лікарська форма може містити краплі, що мають приблизні діаметри в інтервалі 100-5000 мкм, такі як 500-4000 мкм, 1000-3000 мкм або 100-1000 мкм. Альтернативно, з урахуванням об'єму, краплі можуть знаходитися в інтервалі приблизно 0,001-100 мкл, такому як 0,1-50 мкл або 1,0-25 мкл або такому як 0,001-1 мкл.

Аерозольна лікарська форма може приймати форму порошку, суспензії або розчину. Розмір частинок аерозолю має відношення до здатності аерозолю до доставки. Більш дрібні частинки можуть переміщуватися далі вниз по дихальних шляхах у напрямку альвеол, ніж могли б більш великі частинки. В одному варіанті здійснення, аерозольні частинки мають такий розподіл по

діаметру, щоб сприяти доставці протягом повної довжини бронхів, бронхіол і альвеол. Альтернативно, розподіл частинок по розмірах може бути вибраний, щоб цілеспрямовано діяти на конкретну частину дихального повітроносного шляху, наприклад, альвеоли. У випадку аерозольної доставки лікарського засобу, частинки можуть мати діаметри в приблизному інтервалі 0,1-50 мкм, переважно 1-25 мкм, більш переважно 1-5 мкм.

Аерозольні частинки можуть використовуватися для доставки з використанням інгалятора (наприклад, через рот) або назального спрею. Аерозольна лікарська форма може необов'язково містити пропелент і/або поверхнево-активну речовину.

Використовувані в даному описі терміни "послідовність нуклеїнової кислоти" і "полінуклеотид" використовуються взаємозамінно і не мають на увазі якого-небудь обмеження довжини. Використовувані в даному описі терміни "нуклеїнова кислота" і "нуклеотид" використовують взаємозамінно. Терміни "послідовність нуклеїнової кислоти" і "полінуклеотид" охоплюють послідовності ДНК (включаючи κДНК) і РНК. Терміни "транс ген" і "ген" також використовують взаємозамінно, і обидва терміни охоплюють їхні фрагменти або варіанти, що кодують цільовий білок.

Трансгени згідно із даним винаходом включають послідовності нуклеїнової кислоти, які були видалені з їхнього природного оточення, ізоляти рекомбінантної або клонованої ДНК, і хімічно синтезовані аналоги або аналоги, біологічно синтезовані гетерологічними системами.

Полінуклеотиди згідно із даним винаходом можуть бути отримані за допомогою будь-яких засобів, відомих у даній галузі. Наприклад, великі кількості полінуклеотидів можуть бути отримані за допомогою реплікації в придатній клітині-хазяїні. Фрагменти природних або синтетичних ДНК, що кодують бажаний фрагмент, будуть включатися в конструкції рекомбінантних нуклеїнових кислот, звичайно конструкції ДНК, здатні до введення в прокаріотичну або еукаріотичну клітину і реплікації в ній. Звичайно конструкції ДНК будуть придатними для автономної реплікації в одноклітинному хазяїні, такому як дріжджі або бактерії, але також можуть бути призначені для введення в геном культивованих клітинних ліній комах, ссавців, рослин або інших еукаріотичних клітинних ліній і інтеграції в їхній геном.

Полінуклеотиди згідно із даним винаходом також можуть бути отримані хімічним синтезом, наприклад, за допомогою фосфорамідитного методу або триєфірного методу, що можуть виконуватися на автоматизованих олігонуклеотидних синтезаторах, які промислово випускаються. Дволанцюжковий фрагмент може бути отриманий з одноланцюжкового продукту хімічного синтезу або за допомогою синтезування комплементарного ланцюга і спільного відпалювання ланцюга при відповідних умовах або за допомогою додавання комплементарного ланцюга з використанням ДНК-полімерази з послідовністю відповідного праймера.

Відносно послідовності нуклеїнової кислоти, термін "ізольована" у контексті даного винаходу означає, що полінуклеотидна послідовність була видалена з її природного генетичного оточення, і таким чином, не містить інших чужорідних або небажаних кодувальних послідовностей (але може включати природні 5'- і 3'-нетрансльовані області, такі як промотори і термінатори), і знаходиться у формі придатних для застосування в генетично сконструйованих системах продукування білка. Такі ізольовані молекули є молекулами, які відділяють від їхнього природного оточення.

З урахуванням виродженості генетичного коду, серед полінуклеотидів згідно із даним винаходом можлива значна варіабельність послідовностей. Вироджені кодони, що охоплюють усі можливі кодони для даної амінокислоти, приведені нижче:

Амінокислота	Кодони	Вироджений кодон
Cys	TGC TGT	TGY
Ser	AGC AGT TCA TCC TCG TCT	WSN
Thr	ACA ACC ACG ACT	ACN
Pro	CCA CCC CCG CCT	CCN
Ala	GCA GCC GCG GCT	GCN
Gly	GGA GGC GGG GGT	GGN
Asn	AAC AAT	AAV
Asp	GAC GAT	GAY
Glu	GAA GAG	GAR
Gln	CAA CAG	CAR

Амінокислота	Кодони	Вироджений кодон
His	CAC CAT	CAY
Arg	AGA AGG CGA CGC CGG CGT	MGN
Lys	AAA AAG	AAR
Met	ATG	ATG
Ile	ATA ATC ATT	ATH
Leu	CTA CTC CTG CTT TTA TTG	YTN
Val	GTA GTC GTG GTT	GTN
Phe	TTC TTT	TTY
Tyr	TAC TAT	TAY
Trp	TGG	TGG
Ter	TAA TAG TGA	TRR
Asn/Asp		RAY
Glu/Gln		SAR
Будь-яка		NNN

Звичайний фахівець в даній галузі зможе зрозуміти, що при визначенні виродженого кодону, представницького для всіх можливих кодонів, що кодують кожну амінокислоту, існує пластичність. Наприклад, декілька полінуклеотидів, які охоплюються виродженою послідовністю, можуть кодувати варіантні амінокислотні послідовності, але звичайний фахівець у даній галузі зможе легко ідентифікувати такі варіантні послідовності при порівнянні з амінокислотними послідовностями даного винаходу.

"Варіантна" послідовність нуклеїнової кислоти має суттєву гомологію або суттєву подібність з еталонною послідовністю нуклеїнової кислоти (або її фрагментом). Послідовність нуклеїнової кислоти або її фрагмент є "гомологічними по суті" (або "істотно ідентичними") з еталонною послідовністю, якщо, при оптимальному вирівнюванні (з відповідними нуклеотидними вставками або делеціями) з іншою нуклеїновою кислотою (або її комплементарним ланцюгом), ідентичність нуклеотидної послідовності становить щонайменше приблизно 70 %, 75 %, 80 %, 82, 84, 86, 88, 90, 92, 94, 96, 98 або 99 % нуклеотидних основ. Методи визначення гомології послідовностей нуклеїнових кислот є відомими в даній галузі.

Альтернативно, "варіантна" послідовність нуклеїнової кислоти є по суті гомологічною (або по суті ідентичною) еталонній послідовності (або її фрагменту), якщо "варіантна" і еталонна послідовність здатні до гібридизації при твердих (наприклад, дуже твердих) умовах гібридизації. На гібридизацію послідовності нуклеїнової кислоти будуть впливати такі умови, як концентрація солі (наприклад, NaCl), температура, або органічні розчинники, на додаток до складу основ, довжини комплементарних ланцюгів, і кількості розбіжностей нуклеотидних основ між гібридизованими нуклеїновими кислотами, що фахівці в даній галузі зможуть легко зрозуміти. Переважно використовуються тверді температурні умови, і вони включають температури, як правило такі, що перевищують 30 °C, що звичайно перевищують 37 °C і переважно перевищують 45 °C. Тверді сольові умови будуть, як правило, становити менше ніж 1000 mM, звичайно менше, ніж 500 mM, і переважно менше ніж 200 mM. pH звичайно знаходиться між 7,0 і 8,3. Комбінація параметрів є набагато більш важливою, ніж будь-який одиничний параметр.

Методи визначення відсотка ідентичності послідовності нуклеїнової кислоти є відомими в даній галузі. За допомогою приклада, при оцінці ідентичності послідовності нуклеїнової кислоти, послідовність, що має задане число суміжних нуклеотидів може бути вирівняна з послідовністю нуклеїнової кислоти (яка має таке ж саме число суміжних нуклеотидів) з відповідної частини послідовності нуклеїнової кислоти згідно із даним винаходом. Програмні засоби, відомі в даній галузі для визначення відсотка ідентичності послідовності нуклеїнової кислоти, включають Nucleotide BLAST.

Звичайний фахівець у даній галузі зрозуміє, що різні види проявляють "переважну частоту використання кодону". Використовуваний у даному описі, термін "переважна частота використання кодону" стосується кодонів, що найбільш часто використовуються в клітинах деяких видів, таким чином, сприяючи одному або декільком представникам можливих кодонів, що кодують кожну амінокислоту. Наприклад, амінокислота треонін (Thr) може кодуватися ACA, ACC, ACG, або ACT, але в клітинах-хазяях ссавців ACC є найбільш часто використовуваним кодоном; в інших видів, переважними можуть бути інші кодони. Переважні кодони для конкретного виду клітин-хазяїн можуть вводитися в полінуклеотиди згідно із даним винаходом за допомогою різних методів, відомих у даній галузі. Уведення послідовностей переважного кодону в рекомбінантну ДНК може, наприклад, збільшити вироблення білка, роблячи трансляцію білка більш ефективною в конкретному типі або виді клітин.

Таким чином, в одному варіанті здійснення винаходу, послідовність нуклеїнової кислоти є кодон-оптимізованою для експресії в клітині-хазяїні.

"Фрагмент" полінуклеотиду, що представляє інтерес, містить ряд послідовних нуклеотидів з послідовності вказаного непроцесованого полінуклеотиду. За допомогою приклада, "фрагмент" полінуклеотиду, що представляє інтерес, може мати (або складатися з) щонайменше 30 послідовних нуклеотидів з послідовності вказаного полінуклеотиду (наприклад, щонайменше 35, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 або 1000 послідовних залишків нуклеїнової кислоти вказаного полінуклеотиду). Фрагмент може включати, щонайменше одному антигенну детермінанту і/або може кодувати щонайменше один антигенний епітоп, що відповідає поліпептиду, який представляє інтерес. Звичайно, фрагмент, визначений у даному описі, зберігає таку ж функцію, як і непроцесовані полінуклеотид або поліпептид.

ОПИС КРЕСЛЕНЬ

Даний винахід тепер буде описаний за допомогою приклада тільки з посиланням на супровідні креслення, на яких:

Фіг. 1 показує ілюстративні плазмід, використовуваних у винаході. Фіг. 1A-1E показують схематичні креслення плазмід, використовуваних для одержання векторів за винаходом. В одному варіанті здійснення винаходу, Фіг. 1A надає інструмент винаходу.

Фіг. 2 показує тривалість експресії трансгена F/HN-SIV у мишачій назальній тканині, перфузію якої виконували 4×10^8 OT F/HN-SIV-CMV-EGFP, і експресію EGFP, визначену через вказану кількість днів після лікування. Показаний негативний контроль, де перфузію назальної тканини виконували розріджувачем вектора (PBS).

Фіг. 3 демонструє тривалість експресії трансгена F/HN-SIV, показуючи її через 1 рік після обробки. Експресію EGFP визначали у 10 незалежних мишей через 360 днів після обробки.

Фіг. 4 показує клітинний розподіл експресії трансгена F/HN-SIV. Експресію EGFP визначали в гістологічних зрізах носової порожнини мишей (2 мм від кінчика носа) через 30 днів після обробки. EGFP-позитивні клітини виробляли білий точковий сигнал.

Фіг. 5 показує типи клітин, трансдуковані з використанням обробки мишачого носа F/HN-SIV. 69 % клітин, трансдукованих у носову порожнину мишей являли собою війчасті респіраторні епітеліальні клітини. Інші трансдуковані клітинні типи включали нейрональні клітини в нюховому епітелії (21 %) і сквамозні клітини (7 %).

Фіг. 6 показує повторне введення F/HN-SIV у ніс мишей. Мишача назальна тканина була трансдукована (як на Фіг. 1) з використанням однієї дози F/HN-SIV-CMV-Lux або двох доз F/HN-SIV-CMV-EGFP, з наступним введенням однієї дози F/HN-SIV-CMV-Lux при 28-денних інтервалах. Таким чином, повторне введення F/HN-SIV у ніс миші не змінює рівнів експресії гена. Експресію трансгена порівнюють з ведучою невірусною лікарською формою для перенесення генів (плазмід CMV-Lux у комплексі з GL67A).

Фіг. 7 відображає трансдукцію людських респіраторних клітинних культур поверхні розділення повітря-рідини (ALI). Людські культури ALI трансдукували з використанням F/HN-SIV-Lux при вказаній множинності зараження (MOI) і візуалізували по експресії Lux через 5 днів після обробки.

Фіг. 8 демонструє, що F/HN-SIV може спрямовувати функціональну експресію CFTR. Клітини HEK293T трансдукували з використанням F/HN-SIV-CMV-EGFP-CFTR при 500 MOI і функціональну активність CFTR визначали по ефлюксу йодиду. F/HN-SIV-CMV-EGFP служила як негативний контроль.

Фіг. 9 показує, що F/HN-SIV ефективно трансдукує овечі і людські первинні клітини легені і легеню миші. Фіг. 9A показує, що трансдукція людських клітин зіскрібка слизової носа (MOI 250) і зрізів людської й овечої легені, культивованих ex vivo (1×10^5 OT/зріз) разом з F/HN-SIV-CMV-Lux приводить до значної експресії трансгена люциферази через 24-48 годин після трансдукції.

Фіг. 9B і 9C відбивають трансдукцію ($\sim 1 \times 10^5$) первинних людських клітин легені, ураженої MB, культивованих на поверхні розділення повітря-рідини (CF hALIs) з використанням векторів (3E7 OT) F/HN-SIV-soCFTR2, що містять CMV- і hCEF промотори трансгена. (B) Число копій вектора (копій провірусної ДНК на копію ендогенної ДНК CFTR) через 6-8 днів після трансдукції. (C) Рівень експресії мРНК CFTR (%VE: копії мРНК CFTR на копію ендогенної мРНК CFTR $\times 100$)

через 6-8 днів після трансдукції. Горизонтальна лінія на (C) представляє цільовий рівень експресії, який дорівнює 5 % VE - як вважають, що представляє терапевтичний поріг. Після in vivo доставки векторів F/HN-SIV-EGFPLux, що містять CMV-, EF1aS і промотор hCEFs у дефектній по інтеграції (ID) або компетентній по інтеграції формі (IC або без мітки) експресію трансгена в клітинах дихальних шляхів визначали в назальному епітелії (D) і епітелії легень (E) мишей (n=6-10/групу). Часову динаміку експресії трансгена люциферази реєстрували за

допомогою візуалізації біоломінесценції *in vivo* і нормалізували по дозі, що доставляється. Фіг. 9F показує представницькі зображення біоломінесценції після трансдукції мишам *in vivo* на день 14 після трансдукції. Фіг. 9G відбиває представницькі зображення біоломінесценції після трансдукції *in vitro* не-Мв hALI на день 5-6 після трансдукції. Фіг. 9H представляє експресію EGFP через 14 днів після трансдукції в мишачий назальний епітелій після доставки 1,6E8 ОТ F/HN-SIV-hCEF-EGFP_{Lux} (vGM020). EGFP візуалізували за допомогою імуногістохімії, ядра візуалізували за допомогою DAPI. Фіг. 9I показує часову динаміку експресії трансгена люциферази в не-Мв ALI, яку реєстрували за допомогою повторної візуалізації біоломінесценції і нормалізували по дозі, що доставляється.

Фіг. 10 показує, що F/HN-SIV ефективно трансдукує легеню вівці *in vivo*. Фіг. A показує, що для моделювання вірусної доставки в легеню вівці, автори винаходу інстальювали (аліквоти 3×100 мкл протягом (5 хвилин) акрифлавін у проксимальні дихальні шляхи при прямій бронхоскопічній візуалізації. Розподіл акрифлавіну можна оцінити по жовтогарячому забарвленню вирізаних дихальних шляхів після смерті. Слід зазначити, що акрифлавін у значній мірі обмежений провідними повітряними шляхами й відсутній в альвеолярних областях. Стрілка вказує приблизну ділянку інстиляції. Числа на лінійці являють собою див. Фіг. 10B є представленням у вигляді діаграми легені вівці (центр/верхня частина трахеї). Зелене коло представляє область на (A). На (B), лінія вказує проходження бронхоскопу для доставки аліквот 3×100 мкл 2,2E9 ОТ/мл (всього 6,6E8 ОТ) F/HN SIV CMV-EGFP_{Lux} n=3 індивідуальним вівцям (коди тварин T121, T156 і T251). Через сім днів після доставки, 5-6 блоків зразків тканини були узяті при розкритті з інтервалами ≈1 см від ділянки інстиляції. Блоки були розділені на 2-3 приблизно еквівалентних зразка і проаналізовані на предмет експресії трансгена за допомогою аналізу люциферази (C), нормалізованого по вмісту білка; і (D) кількісна РВ-ПЛР, нормалізована по рівнях ендогенної мРНК CFTR. Горизонтальна лінія на (C) представляє найвищу активність люциферази, відзначену в якому-небудь зразку, обробленому з використанням невірусного вектора перенесення генів, і (D) цільовий рівень експресії, який дорівнює 5 % VE - як вважають, що представляє терапевтичний поріг.

Фіг. 11 зображує одержання й очищення F/HN Векторів SIV. F/HN Вектори SIV одержували за допомогою 5-плазмідної (пДНК) PEI-опосередкованої часової трансфекції клітин 293T, вирощених у суспензії в масштабі 1 л у біореакторах WAVE з контрольованим рН (GE), із застосуванням масштабованих способів винаходу. Вектори прояснювали за допомогою глибинної/кінцевої фільтрації (GE/Pall), забруднювальні нуклеїнові кислоти видаляли з використанням Бензонази® (Merck), вектори активували з використанням TrypLE Select™ (Life Technology), очищали і концентрували за допомогою аніонообмінної мембранної хроматографії (Pall) і тангенціальної потокової фільтрації (Spectrum). Усі ємності, контейнери і колонки для процесу являли собою одноразові вироби відповідно до вимог cGMP. Усі реагенти, за винятком плазмідної ДНК, не мали тваринного походження відповідно до вимог cGMP. Показано дані для різних конфігурацій вектора (промотор трансгена, трансген, статус інтеграції). Фізичні і функціональні титри визначали, використовуючи К-ПЛР. (A) Фізичний титр від первинного проясненого збору і кінцевий очищений продукт. Середній вихід процесу становить ~44 %. (B) Функціональний титр кінцевого продукту. Середній функціональний титр становить ~2×10⁹ ОТ/мл продукту і ~3×10⁹ ОТ/л об'єму біореактора. Цільова продуктивність і виходи були перевищені. У нижніх вихідних векторах CFTR використовують промотор CMV трансгена. У верхніх вихідних векторах CFTR використовують промотори EF1aS і hCEF трансгена. (C) Частинки кінцевого продукту: співвідношення інфекційності є щільно кластеризованим і аналогічним значенням від інших виробників векторів з високою якістю (Oxford BioMedica, BlueBirdBio). Медіанна частка: співвідношення інфекційності становить ~300. Консистенції продукту досягали з використанням методів "Дизайну експериментів", що підтверджує остаточний перехід до процесу виробництва на підставі дозволу контролюючого відомства на основі QbD. (D) Функціональний титр остаточного продукту суворо корелює з первинним фізичним титром, що вказує на ненасичувальні умови з використанням концентрації вектора, необмежувальної стадії процесу - припускають, що масштабування очищення буде ефективним.

Фіг. 12 відображає профілювання ділянки введення (IS) і виживаність трансдукованих мишей. Фіг. 12A-F представляє порівняння профілів IS від F/HN-SIV-трансдукованої мишачої легені і трансдукованої вектором VSV-G-HIV мишачої сітківки (Bartholomae et al, Mol Ther (2011) 19:703-710). Профілі IS для (A, C,E) були отримані, виходячи з результатів глибокого секвенування ДНК легені від двох мишей, трансдукованих F/HN-SIV, і для (B, D,F) за результатами для послідовностей IS сітківки з VSV-G-HIV (Schmidt laboratory). (A, B) Агреговані ділянки IS (легеня, 2862; сітківка, 262) нанесені на каріограми, які генеруються за допомогою геномного браузера Ensembl. (C, D) Відстані від IS до сайтів початку транскрипції (TSS). Число

IS у кожному біні відстаней показані смужками вище. Сума перевищує загальне число проаналізованих IS, оскільки типовий IS розташований поруч з декількома TSS. Графіки генерували за допомогою застосування геномного браузеру UCSC і геномного аналізатора GREAT (great.stanford.edu). (E, F) QuickMap (www.gts.org) і порівняння між випадковими (ромб) і спостережуваними (квадрат) частотами введення на хромосому. (G) Вживаність мишей, оброблених попередником F/HN-вектор SIV (вірусний вектор, виготовлений відповідно до відомого способу з використанням клітин, що прилипають: чорна лінія, 24 місяці від даних) і вектор поточної генерації (GTC: темно-сіра лінія, дані для 8 місяців) у порівнянні з мишами, обробленими буфером (світло-сіра лінія, дані для 24 місяців). Були об'єднані дані від різних експериментів, що включають мишей, оброблених буфером або $\sim 10^7$ ОТ вірусу за допомогою назального нюхання.

Фіг. 13 показує hCEF-опосередковану респіраторну експресію трансгена - з використанням лентивірусного перенесення гена з використанням збагаченого CpG трансгена - експресію трансгена (середнє світіння п/с/см²/sr) залежно від днів після дозування. Високі рівні експресії репортерного гена люциферази Gaussia у порівнянні з контролем (інтактна легеня) спостерігали як у легені (квадрат), так і носі (коло) протягом щонайменше 168 днів після дозування.

Фіг. 14 показує ефект трансдукції людських кишкових органодів з використанням лентивірусного вектора CFTR (vGM058) винаходу. Лівостороння панель показує, що індуковане форсколіном набрякання значно ($p < 0,001$) знижувалося в трансдукованих vGM058 органодах. На правобічній панелі, клітини A549 трансдукували vGM058 або контрольним вірусом, і функцію CFTR кількісно оцінювали з використанням аналізу ефлюксу радіоактивного йодиду. Значні ($p < 0,05$) рівні CFTR-опосередкованого ефлюксу йодиду виявляли в трансдукованих vGM058 клітинах.

Фіг. 15A показує схематичні зображення плазмід, застосовуваної для одержання векторів A1AT винаходу. В одному варіанті здійснення винаходу, Фіг. 15A представляє інструмент винаходу. Фіг. 15B показує контрольну плазмиду, яка кодує ген-репортер люциферази Gaussia.

Фіг. 16 показує, що F/HN-SIV ефективно трансдукує первинні клітини легені людини в культурі ALI. Зокрема, трансдукція людських культур ALI приводить до значної експресії трансгена люциферази протягом щонайменше 80 днів після трансдукції. Кожна точка представляє середнє значення BCO/мкл у середовищі для $n=6$ ALI у показаній часовій оцінці. Вертикальні риси представляють стандартну помилку середнього.

Фіг. 17 показує експресію гена після трансдукції зрізів людської легені ($n=6$ на групу) з використанням SIV1 hCEF-sogLux (Фіг. 17A) і SIV1 hCEF-sohAAT (Фіг. 17B). Спостерігали високі рівні експресії. Кожна точка представляє середнє значення BCO/мкл у середовищі для $n=6$ зрізів легені в показаній часовій оцінці. Вертикальні риси представляють стандартну помилку середнього.

Фіг. 18 показує довгострокову експресію (>12 місяців) люциферази Gaussia після опосередкованого лентивірусом перенесення генів (SIV1 hCEF-soGLux) in vivo. A: гомогенат тканини легені; B: рідина бронхоальвеолярного лаважу (BAL); C: сироватка. Кожна точка представляє середнє значення BCO/мкл в одній групі тварин ($n=5$ або 6 на групу) при зборі у вказаній часовій оцінці. Вертикальні риси представляють стандартну помилку середнього.

Фіг. 19 показує високі рівні експресії A1AT після опосередкованого лентивірусом перенесення гена (SIV1 hCEF-sohAAT) in vivo. Кожна точка представляє одну тварину. Горизонтальні риси представляють медіану для кожної групи.

Фіг. 20 показує довгострокову експресію A1AT після опосередкованого лентивірусом перенесення гена AAT (A1AT) in vivo. A: гомогенат легені; B: BAL; і C: сироватка.

Фіг. 21 показує рівень A1A1 у бронхоальвеолярній лаважній рідині після опосередкованого лентивірусом перенесення гена (SIV1 hCEF-sohAAT) in vivo. Опосередковане лентивірусом перенесення гена (SIV1 hCEF-sohAAT) in vivo.

Фіг. 22A показує схематичні зображення конструкцій кДНК FVIII, використовуваних для одержання векторів FVIII винаходу. Фіг. 22B показує вірусні вектори за винаходом. В одному варіанті здійснення винаходу, Фіг. 22B представляє інструмент винаходу. Фіг. 22C-F показує схематичні зображення плазмід пДНК1, використовуваних для отримання векторів FVIII винаходу. В одному варіанті здійснення винаходу, Фіг. 22C-F представляє інструменти за винаходом.

Фіг. 23 показує ефективність трансдукції HEK293T з використанням vGM142 (SIV-F/HN-FVIII-N6-ко) для завантаження 1 (A) і завантаження 2 (B). Графіки показують активність FVIII для кожного MOI через 48 і 72 годин після трансдукції. Кожен символ представляє незалежний експеримент ($n=5-6$ експериментів). Горизонтальна риса вказує на середнє для групи \pm SD.

Аналіз виконували з використанням однофакторного дисперсійного аналізу Anova (GraphPad Prism) з множиною порівнянь з необробленим контролем ** $p < 0,01$; *** $p < 0,001$; **** $p < 0,0001$.

Фіг. 24 показує оцінку vGM142 у системі in vivo (мишача модель). А: легеня; В: Рідина BAL; С: плазма. Графіки показують рівень hFVIII (у вигляді процентної частинки від нормального рівня) у порівнянні з різними обробками для Груп 1-3 (Групи 1 і 2-10 днів обробки, 3 дози/на тиждень, по 100 мкл вектора на мишу (загальна кількість доз бути дорівнює 3 на тварину); Група 3-28 днів обробки, 3 дози/на тиждень, по 100 мкл вектора на мишу (загальна кількість доз бути дорівнює 12 на тварину)). Кожен символ представляє індивідуальну мишу (Група 1 ($n=4$), оброблена загальною дозою вектора, яка дорівнює $1,4 \times 10^6$ ЕТТ/миша; Група 2 ($n=3$), оброблена загальною дозою вектора, яка дорівнює $1,57 \times 10^8$ ЕТТ/миша і Група 3 ($n=4$), оброблена загальною дозою вектора, яка дорівнює $3,36 \times 10^8$ ЕТТ/миша). Горизонтальна риска показує середнє значення FVIII: рівні Ag \pm SD. Аналіз виконували з використанням однофакторного дисперсійного аналізу Anova (GraphPad Prism) із множинними порівняннями між групами обробки **** $p < 0,0001$.

ПРИКЛАДИ

Винахід тепер буде описаний з посиланням на Приклади, приведені нижче. Ці приклади не є обмежувачами обсягу винаходу, і фахівець у даній галузі зміг би зрозуміти, що могли б використовуватися прийнятні еквіваленти в межах обсягу даного винаходу. Таким чином, Приклади можуть розглядатися як компонент частин винаходу, і індивідуальні аспекти, описані в них, можуть розглядатися, як розкриті незалежно або в будь-якій комбінації.

Приклад 1: Культура клітин

Клітини HEK293T, Freestyle 293F (Life Technologies, Paisley, UK) і 293T/17 (CRL-11268; ATCC, Manassas, VA) підтримували в мінімальному поживному середовищі Ігла в модифікації Дульбекко (Invitrogen, Carlsbad, CA), що містить 10 % сироватки крові ембріонів корів і доповненому пеніциліном (100 Од/мл) і стрептоміцином (100 мкг/мл) або середовищі для експресії Freestyle™ 293 (Life Technologies).

Приклад 2: Побудова плазмід

pCAGGS-Fct4 і pCAGGS-SIVct+HN були побудовані наступним чином:

(i) Плазміда SIVct/HN містить ген, який кодує цитоплазматичний кінцевий сегмент SIVagm TMP (зворотний), гібридизований з позаклітинною частиною вірусу і трансмембранних областей білка SeV HN. Були синтезовані три олігонуклеотидні пари: пара 1, 5'-TCGAGATGTGGTCTGAGTTAAAAATCAGGAGCAACGACGGAGGTGAAGGACCAGACGCCAACGA CCC-3' (SEQ ID NO: 18) і пара 2, 5'-CCGGGGGTCGTTGGCGTCTGGTCCTTCACCTCCGTCGTTGCTCCTGATTTTAACTCAGACCACA TC-3' (SEQ ID NO: 19); пара 3, 5'-CCGGGGAAAGGGGGTGCAACACATCCATATCCAGCCATCTCTACCTGTTTATGGACAGA-3' (SEQ ID NO: 20) і пара 4, 5'-ACCCTCTGTCCATAAACAGGTAGAGATGGCTGGATATGGATGTGTTGCACCCCTTTCC-3' (SEQ ID NO: 21); і пара 5, 5'-GGGTTAGGTGGTTGCTGATTCTCTCATTCACCCAGTGGG-3' (SEQ ID NO: 22) і 5'-GATCCCCCACTGGGTGAATGAGAGAATCAGCAACCACCTA-3' (SEQ ID NO: 23).

Ці олігонуклеотидні пари піддавали відпалюванню і клонували в сайти XhoI і BamHI pBluescript KS+ (Stratagene) з одержанням pKS+SIVct. pCAGGS-SIVct/HN була побудована за допомогою клонування фрагмента XhoI-DraIII з 160 по з pKS+SIVct і фрагмента DraIII-Bsu36I з 1500 по з pCAGGS-HN, що несе ген HN дикого типу (HNwt) у сайті XhoI вектора pCAGGS, у сайті XhoI і Bsu36I pCAGGS. Ця плазміда була побудована таким чином, що цитоплазматичний кінцевий сегмент білка HN був замінений на цитоплазматичний кінцевий сегмент SIVagm TMP.

Для побудови pCAGGS-SIVct+HN, гени, що кодують цитоплазматичний кінцевий сегмент SIVagm TMP і N-кінець білка HN, спочатку ампліфікували за допомогою ПЛР із pCAGGS-SIVct/HN з використанням пари праймерів 5'-GAGACTCGAGATGTGGTCTGAGTTAAAAATCAGG-3' (SEQ ID NO: 24) і 5'-AGAGGTAGACCAGTACGAGTCACGTTTGGCCCTATCACCATCCCTAACCCTCTGTCATAAAC-3' (SEQ ID NO: 25). Отриманий у результаті ПЛР фрагмент клонували в сайти XhoI і Acc pKS+SIVct для генерації pKS+SIVct-H. Потім фрагмент XhoI-DraIII з pKS+SIVct-H і фрагмент DraIII-Bsu36I з pCAGGS-HN клонували в сайти XhoI і Bsu36I pCAGGS з одержанням pCAGGS-SIVct+HN.

Послідовності cPPT і WPRE вводили в плазмиду, яка походить від SIV, для перенесення генів. Приклад використаної послідовності WPRE представлений у SEQ ID NO: 8.

Плазміда pGM101 містить точку початку реплікації colE1, канаміцин-стійкий ген і промотор і була створена за допомогою синтезу синтетичного гена (GeneArt, Regensburg, Germany; тепер Life Technologies Ltd).

Гібрид CMV/SIV R U5 LTR, часткові послідовності Gag, RRE, cPPT, SIN U3 і R з pBS/CG2-Rc/s-CMV-D U (Nakajima et al. 2000 Human Gene Therapy 11:1863) ампліфікували за допомогою ПЛР і вводили поряд з послідовністю енхансер hCEF/Промотор, ампліфікованою за допомогою ПЛР із pGM169 (Hyde et al. Nature Biotechnology 26:549), і кДНК soCFTR2, виділеною з pGM169 на фрагмент ферменту рестрикції NheI-ApaI у pGM101 для створення pGM326.

Енхансер CMV/промотор курячого бета-актину поряд з асоційованими послідовностями екзон/інтрон, послідовностями SIV GagPol і RRE і полі A SV40/точка початку реплікації ампліфікували за допомогою ПЛР із pCAGGS/Sagm-gtr (Nakajima et al. 2000 Human Gene Therapy 11:1863) для створення pGM297. Енхансер CMV/промотор поряд з асоційованими послідовностями екзон/інтрон і послідовністю поліA SV40 з pCI (Promega, Madison, WI, USA) ізолювали на фрагмент ферменту рестрикції BglII-BamHI і послідовність SIV Rev, отриману з pCAGGS/Sagm-gtr, ампліфіковану за допомогою ПЛР, вводили в pGM101 для створення pGM299.

Енхансер CMV/промотор курячого бета-актину поряд з асоційованими послідовностями екзон/інтрон, кДНК Fct4 і поліA SV40/точка початку реплікації з pCAGGS-Fct4 ізолювали на фрагмент ферменту рестрикції SalI-HindIII за допомогою комбінації генного синтезу, ПЛР і рекомбінації фрагмента ферменту рестрикції і вводили в pGM101 для створення pGM301.

Енхансер CMV/промотор курячого бета-актину поряд з асоційованими послідовностями екзон/інтрон, кДНК SIVct+HN і поліA SV40/точка початку реплікації з pCAGGS-SIVct+HN ізолювали на фрагмент ферменту рестрикції SalI-HindIII за допомогою комбінації і вводили в pGM101 для створення pGM303.

Інші pGM-плазмиди винаходу одержували, використовуючи стандартні методи, і відповідно до приведеного вище розкриття.

На всьому протязі цих варіантів виконання зборки плазмідних ДНК, ферменти рестрикції і ПЛР-полімерази поставлялися New England Biolabs (Ipswich, MA, USA), а реагенти для очищення ДНК поставлялися Qiagen (Limburg, Netherlands).

Приклад 3: Одержання вектора SIV

Система з чотирьох плазмід:

Дефектний по реплікації самоінактивований вектор SIV був побудований з мінорними модифікаціями. Коротко, SeV-F/HN-псевдотипований вектор SIV одержували за допомогою трансфекції клітин 293T/17 (чашки для культивування з діаметром 15 см) з використанням чотирьох плазмід, у комплексі з реагентами Ліпофектамін/Плюс (Invitrogen) відповідно до рекомендацій виготовлювача [Плазмід-1:10 мкг отриманої з SIV плазмиди перенесення, що несе GFP, ген-репортер люциферази (lux), або гібридна конструкція GFP-CFTR, Плазмід-2:3 мкг пакувальної плазмиди, Плазмід-3:2 мкг pCAGGS-Fct4, Плазмід-4:2 мкг pCAGGS-SIVct+HN; Фіг. 1A-E показують схематичні креслення плазмід, використовуваних для одержання векторів за винаходом]. Псевдотипований VSV-G вектор SIV одержували, використовуючи аналогічний протокол, але плазмиду pVSV-G (2 мкг; Clontech, Mountain View, CA) використовували замість pCAGGS-Fct4 і pCAGGS-SIVct+HN. Через 12 годин після трансфекції культуральне середовище замінювали на 30 мл безсироваткового модифікованого Дульбекко середовища Ігла, що містить 5 ммоль/л бутирату натрію. Бутират натрію стимулює одержання вектора, щоб інгібувати деацетилазу гістону. Супернатант культури, що містить вектор SIV збирали через 48 годин після трансфекції, фільтрували через фільтрувальну мембрану 0,45 мкм, і додатково концентрували за допомогою високошвидкісного центрифугування (20000 g протягом 4 годин при 4 °C, ротор Avanti JA18; Beckman Coulter, Brea, CA). Згустки вектора суспендували в PBS (Invitrogen) до 100-200-кратної концентрації і зберігали при -80 °C.

Система з п'яти плазмід (переважна):

SeV-F/HN-псевдотипований вектор SIV одержували за допомогою трансфекції клітин HEK293T або 293T/17, культивованих у середовищі для експресії FreeStyle™ 293 з використанням суміші з п'яти плазмід, що мають наступні характеристики: пДНК1 (наприклад, pGM326; Фіг. 1A) кодує мРНК лентивірусного вектора; пДНК2a (наприклад, pGM297; Фіг. 1B) кодує білки SIV Gag і Pol; пДНК2b (наприклад, pGM299; Фіг. 1C) кодує білки SIV Rev; пДНК3a (наприклад, pGM301; Фіг. 1D) кодує білок Fct4 з Вірусу Сендай [Kobayashi et al., 2003 J. Virol. 77:2607]; і пДНК3b (наприклад, pGM303; Фіг. 1E) кодує SIVct+HN з вірусу Сендай [Kobayashi et al., 2003 J. Virol. 77:2607] у комплексі з PEIpro (Polyplus, Illkirch, France).

Середовище для культури клітин доповнювали через 12-24 годин після трансфекції бутиратом натрію. Бутират натрію стимулює одержання вектора за допомогою інгібування деацетилази гістону, що приводить до збільшення експресії компонентів гібридних білків SIV і вірусу Сендай, які кодуються п'ятьма плазмідами. Середовище для культури клітин доповнювали через 44-52 годин і/або 68-76 годин після трансфекції з використанням 5

одиниць/мл Нуклеази Бензоази (Merck Millipore, Nottingham, UK). Супернатант культури, що містить вектор SIV збирали через 68-76,5 годин після трансфекції, і просвітлювали за допомогою фільтрації через мембрану 0,45 мкм. Вектор SIV обробляли за допомогою обробки протеазою, що містить трипсинову активність - наприклад, що не має тваринного походження, рекомбінантним ферментом, таким як TrypLE Select™. Згодом, вектор SIV звичайно додатково очищують і концентрують за допомогою аніонообмінної хроматографії і/або ультрафільтрації/діафільтрації з фільтрацією тангенціальним потоком.

Такий же метод застосовували для генерації лентивірусних векторів, що містять трансгени A1AT і FVIII, з використанням плазмід Фіг. 15А і Фіг. 22В, при заміні плазмід Фіг. 1А для забезпечення відповідного трансгена (див. Приклади 15 і 20).

Приклад 4: Титрування вектора

Метод 1:

Титр частинок визначали, використовуючи ПЛР зі зворотною транскриптазою у режимі реального часу. Вірусну РНК очищали, застосовуючи міні-набір для вірусних РНК, QIAamp (QIAGEN, Strasse, Germany), і зворотно транскрибували, використовуючи Superscript II (Invitrogen). Систему ПЛР із зондом QuantiTest (QIAGEN) і праймери для ампліфікації 131 нуклеотидів (по) перекривних послідовностей WPRE (прямий праймер: 5'-ggatcacgtgcttaatgcc-3', зворотний праймер: 5'-acgccacgtgctgcagacaac-3') застосовували відповідно до протоколу виготовлювача на детекторній системі для послідовностей ABI PRISM 7700 (PE Applied Biosystems, Foster City, CA). ДНК плазмиди для перенесення генів SIV ($3 \times 10^4 \times 10^6$ молекул) застосовували як стандарт.

Одиниці трансдукції (ОТ/мл) визначали за допомогою трансдукування клітин 293Т серійними розведеннями вихідного розчину вектора і кількісної оцінки трансдукованих клітин за допомогою флуоресценції GFP (для F/HN-SIV-GFP і VSV-G-SIV-GFP) або фарбування з використанням антитілу проти люциферази (для F/HN-SIV-lux).

Метод 2 (переважний):

Титр частинок (ВЧ/мл) звичайно визначали, використовуючи ПЛР зі зворотною транскриптазою у режимі реального часу. Вірусну РНК очищали, використовуючи міні-набір для вірусних РНК, QIAamp (QIAGEN, Strasse, Germany), і зворотно транскрибували, використовуючи зворотну транскриптазу (Life Technologies). Систему для кількісної ПЛР, TaqMan (Life Technologies), з використанням праймерів, ампліфікуючи частину послідовності WPRE, застосовували на детекторній системі для послідовностей ABI PRISM 7700 (Life Technologies). Транскрибовані *in vitro* молекули РНК WPRE застосовували як кількісні стандарти.

Одиниці трансдукції (ОТ/мл) визначали за допомогою трансдукування клітин 293Т/17 або Freestyle 293F серійними розведеннями вектора SIV і кількісної оцінки WPRE, що містить провірусну ДНК, з використанням системи для кількісної ПЛР, TaqMan (Life Technologies), використовуючи праймери, ампліфікуючи частину послідовності WPRE на детекторній системі для послідовностей ABI PRISM 7700 (Life Technologies). Молекули ДНК плазмід, що містять послідовності WPRE, застосовували як кількісні стандарти.

Приклад 5: Генерація культур tEC, збагачених базальними клітинами

Мишачі трахеальні епітеліальні клітини (tEC) виділяли наступним чином. Мишей C57BL/6N умертвляли і трахеї вирізували від гортані до основних бронхіальних гілок, використовуючи стерильні хірургічні інструменти. Тканини вміщували в пробірку, що містить холодне середовище Хема F-12 з 100 Од/мл пеніциліну (P), 100 мкг/мл стрептоміцину (S) і 2,5 мг/мл амфотерицину В (A) (середовище Хема F12/PSA) і витримували на льоді. У стерильній витяжній шафі з ламінарним потоком повітря, трахеї очищали від прилягаючих м'язів і сполучної тканини, розрізали в подовжньому напрямку, щоб експонувати внутрішній респіраторний епітелій, і вміщували в 0,15 % розчин пронази в середовищі F-12 (~5 мл у пробірці 15 мл). Розщеплення тканини виконували протягом ночі (15-18 година) при 4 °С. Щоб блокувати ферментативну реакцію, 10 % сироватку крові ембріонів корів (FBS) додавали до зразка обробленої тканини. Після плавного перекидання пробірки для більшого від'єднання клітин, трахеї вміщували в нову пробірку, що містить розчин 10 % FBS/Хема F-12/PS, і перевертали, як раніше. Цю стадію повторювали ще два рази. Вміст чотирьох пробірок об'єднували разом і центрифугували при 500 g протягом 10 хв при 4 °С. Згусток ресуспендували в розчині ДНКази (0,5 мг/мл неочищеної панкреатичної ДНКази плюс 10 мг/мл BSA у розчині FBS/Хема F-12/PS, приблизно 200 мкл/трахею), інкубували на льоді протягом 5 хвилин і центрифугували як раніше.

Після видалення супернатанта, tEC ресуспендували в цільовому середовищі для клітин-попередників (PCT), (CnT-17, CELLnTEC, Bern, Switzerland), які не містять антибіотиків і протигрибкових засобів в складі, конкретно розроблених для виділення і проліферації клітин-попередників людських і мишачих повітряноносних шляхів. tEC потім переносили в чашки для

культур тканин Primaria (Becton Dickinson Labware, Franklin Lakes, NJ, USA) і інкубували протягом 3-4 година в 5 % CO₂ при 37°C. Неприлиплі клітини збирали і центрифугували при 500 g протягом 5 хв при 4°C і підраховували на гемоцитометрі. Для генерації культур, збагачених базальними клітинами, tEC суспендували в середовищі PCT і висівали на чашку Nunclon™Δ (Nunc A/S, Roskilde, Denmark), покриту 50 мкг/мл колагену типу 1 із сухожилля щурячого хвоста при рекомендованій щільності висівання, яка дорівнює 4×10³ клітин/см². tEC також культивували в контрольному основному середовищі, що містить DMEM/Хема F12, доповненому L-глутаміном (4 мМ), HEPES (15 мМ) і NaHCO₃ (3,4 мМ). Чашки інкубували при 37°C з 5 % CO₂. Щоб визначити частку базальних клітин у популяції tEC до і після розмноження в середовищі PCT, цитоспінові препарати забарвлювали антитілом проти KRT5 і відповідним вторинним антитілом.

Пул продуктованих tEC містив як повністю диференційовані клітини (клітини Клара і війчасті клітини), так і базальні клітини. Як тільки їх висівали на проникні мембрани, tEC були здатні генерувати систему культивування на межі розділення повітря/рідини (ALI) у результаті проліферації базальних клітин і редиференціації в секреторні і війчасті клітини. Для встановлення альтернативного і швидкого протоколу для розмноження базальних клітин, двовимірні (2D) культури tEC з використанням доступного для придбання патентованого цільового середовища для клітин-попередників (PCT), специфічно складені для підтримки проліферації клітин попередників дихальних шляхів при підтримці їх у недиференційованому статусі, піддавали оцінці. Як негативний контроль, tEC піддавали впливу складу основного середовища без додавання специфічних факторів росту. tEC, висіяні на покриті колагеном пластикові поверхні (4×10³ клітин/см²) і піддані впливу середовища PCT були здатні швидко рости і ставали конфлюентними в межах 5-8 днів, у той час, як tEC, піддані впливу контрольного середовища без факторів росту були нездатні проростати і розмножуватися. Щоб установити, чи приводить застосування середовища PCT до значного збагачення популяції базальних клітин, tEC збирали при приблизно 80 % злитті (n=6 ямок), фіксували й обробляли антитілами проти кератину 5 (Krt5), специфічним маркером базальних клітин. Свіжовиділені tEC застосовували як контролі (n=3 унікальних препаратів). Частка Krt5-позитивних базальних клітин після розмноження в середовищі PCT була вища (78±1,4 %), ніж у свіжовиділених пулах tEC (33±0,6 %), демонструючи, що базальні клітини мишачих дихальних шляхів можуть селективно і швидко розмножуватися зі змішаного пулу tEC з використанням промислового середовища.

Приклад 6: Трансдукція культур tEC, збагачених базальними клітинами ex vivo з використанням F/HN-SIV-GFP

Щоб визначити чи може F/HN-вектор SIV ефективно трансдукувати базальні клітини ex vivo, tEC, отримані в Прикладі 5, вирощували до приблизно 70 % злиття протягом 7 днів у середовищі PCT і трансдукували з використанням F/HN-SIV, що несе ген-репортер зеленого флуоресцентного білка (F/HN-SIV-GFP) при MOI 100 і інкубували при 37°C з 5 % CO₂ протягом 3 днів. tEC, отримані з тварин дикого типу, і GFP-трансгенних тварин культивували в тих же самих умовах і використовували як негативну (вірусна трансдукція відсутня) і позитивну контрольні групи, відповідно (n=3-6 ямок/групу).

Щоб кількісно оцінити частку GFP-позитивних клітин, культури tEC, збагачені базальними клітинами, відділяли за допомогою ферменту акутази (CELLnTEC), ресуспендували в PBS/1 % BSA і піддавали аналізу методом FACS, з підрахунком у середньому 20,277±2,478 клітин/групу. F/HN-вектор SIV трансдукував 26 %±0,9 % tEC, збагачених базальними клітинами (p<0,0001 при порівнянні з нетрансдукованими контролями).

Для оцінки того, або є трансдуковані GFP-експресувальні клітини базальними клітинами, через три дні після інфікування клітини двічі забарвлювали антитілами проти Krt5 і GFP. Імунофлуоресцентне фарбування культивованих клітин показало, що приблизно 40 % Krt5-експресувальних клітин також експресували GFP ген-репортер, показуючи, що F/HN-вектор SIV може трансдукувати попередники базальних клітин ex vivo.

Приклад 7: Введення в ніс миші in vivo

Використовували мишей C57BL/6N (самки, 6-8 тижнів). Мишей анестезували, вміщували горизонтально на спину на нагрітій дошці, і тонкий катетер (зовнішній діаметр <0,5 мм) вводили на глибину ~2,5 мм від кінчика носа в ліву ніздрю. Використовуючи шприцевий насос (Cole-Parmer, Vernon Hills, IL), вектор (100 мкл) потім повільно перфузували на поверхню назального епітелію (1,3 мкл/хв) протягом 75 хвилин. Незважаючи на перфузію вірусу в ліву ніздрю, автори винаходу звичайно спостерігали трансфекцію як у лівій, так і правій ніздрях, що було зумовлено дисперсією розчинів через цілу носову порожнину. Мишей, трансдукованих PBS і VSV-G-SIV, прекодиційованих з використанням 1 % лізофосфатидилхоліну, як описано Limberis et al., 2002, застосовували як контролі. У вказані відмітки часу (3-360 днів після трансдукції), мишей

умертвляли для візуалізації експресії GFP. Як показано на Фіг. 2, експресію GFP спостерігали протягом щонайменше 449 днів після трансдукції, у той час як негативний контроль не показував експресії GFP. Як показано на Фіг. 3, експресія трансгена з F/HN-вектором SIV була стійкою, зі спостережуваною експресією GFP, протягом щонайменше 360 днів після трансдукції в 10 незалежно тестованих мишей.

Аналогічно, як показано на Фіг. 13, трансдукція F/HN-вектором SIV винаходу, що містить промотор hCEF, приводила до довгострокової експресії CpG-збагаченого гена-репортера (люциферази). Високі рівні експресії відносно контролю спостерігали як у легенях, так і носі, протягом щонайменше 169 днів після трансдукції.

В експериментах з повторним введенням групи мишей трансдукували з використанням або однієї дози F/HN-SIV-lux (група з однією дозою), або двох доз F/HN-SIV-GFP (день 0, день 28), з наступною обробкою F/HN-SIV-lux у день 56 (група з повторною дозою). Важливо відзначити, що миші, які одержували F/HN-SIV-lux (група з однією дозою), і F/HN-SIV-lux у день 56 (група з повторною дозою), мали подібний вік і були трансдуковані в той самий час. Експресію гена аналізували через 30 днів після введення F/HN-SIV-lux. Для порівняння, мишей трансфікували катіонним ліпідом GL67A у комплексі з геном-репортером люциферази, як описано раніше (Griesenbach, U. et al., *Methods Mol Biol.* 2008; 433:229-42) і експресію люциферази вимірювали через 2 дні після трансфекції.

Як показано на Фіг. 6, повторне введення F/HN-SIV у ніс миші не змінювало рівнів експресії гена. Експресію трансгена порівнюють з лідируючою невірусною лікарською формою для перенесення генів (плазмід CMV-Lux у комплексі з GL67A).

Профілювання ділянки введення виконували на трансдукованих мишах, і досліджували час виживаності, як викладено при описі Фіг. 12 вище. Спостерігали, що трансдукція з використанням F/HN-вектора SIV винаходу не чинить яких-небудь несприятливих впливів на виживаність мишей у порівнянні з існуючим F/HN-вектором SIV або негативним контролем (тільки буфер) (див. Фіг. 12G).

Приклад 8: Індукована регенерація назальних епітеліальних клітин при обробці полідоканолом

Назальні епітеліальні клітини очищали за допомогою обробки полідоканолом відповідно до описаного методу (Borthwick et al., *Am J Respir Cell Mol Biol.* 2001 Jun; 24(6):662-70), з деякою модифікацією. Коротко, мишей анестезували і 10 мкл полідоканолу (2 %) вводили в ніс у вигляді болюса за допомогою "назального нюхання". Для підтвердження очищення і регенерації назальних епітеліальних клітин, назальну тканину перфузували 10 мкл 2 % (об./об. у PBS) полідоканолу (монододецилового ефіру нонаетиленгліколю; SIGMA, St Louis, MO), і виконували гістологічний аналіз через 24 години і 7 днів після обробки (n=3/група). Для аналізу трансдукції можливих клітин-попередників або стовбурових клітин, автори винаходу спочатку вводили вектор F/HN-SIV-GFP (4×10^8 ОТ/миша) у назальний епітелій миші. Через сім днів після трансдукції, назальну тканину перфузували 10 мкл 2 % (об./об. у PBS) полідоканолу, і дану обробку повторювали знову через 3 тижні. Гістологічні зрізи аналізували через 58 днів після введення вектора (30 днів після останньої обробки полідоканолом).

Приклад 9: Біолоюмінесцентна візуалізація

Мишам за допомогою внутрішньоочеревинної ін'єкції вводили 150 мг/кг D-люциферину (Xenogen, Alameda, CA) за 10 хвилин до візуалізації й анестезували ізофлураном. Біолоюмінесценцію (фотони/с/см²/sr) від живих мишей вимірювали, використовуючи систему IVIS50 (Xenogen) при групуванні по 4 протягом 10 хвилин, застосовуючи програму з програмного забезпечення Living Image (Xenogen). Для анатомічної локалізації генерували псевдоколірне зображення, що представляє інтенсивність світла (синій: найменш інтенсивний, червоний: найбільш інтенсивний), використовуючи програмне забезпечення Living Image, і накладене на напівтонове стандартне зображення. Для кількісного вимірювання біолоюмінесценції в носі, емісію фотонів у заданій області (червона коробочка) вимірювали за допомогою маркування стандартизованої області для кількісної оцінки. Розмір червоної коробочки зберігали постійним, і її вміщували над головами тварин, як вказано на Фіг. Важливо відзначити, що області маркували, використовуючи напівтонове стандартне зображення, щоб уникнути зміщення.

Приклад 10: Препарат тканини для гістологічної оцінки експресії GFP і/або виявлення базальних клітин

Мишей умертвляли і шкіру видаляли. Голови відрізали на рівні очей і шкіру, щелепи, язик, і м'який кінчик носа обережно видаляли. Для візуалізації *in situ* експресії GFP у носовій порожнині, флуоресценцію GFP виявляли, використовуючи флуоресцентну стереоскопічну мікроскопію (Leica, Ernst Leitz Optische Werke, Germany). Згодом, тканину фіксували в 4 % параформальдегіді (pH 7,4) протягом ночі при кімнатній температурі і потім занурювали в 20 %

ЕДТА (рН 7,5 протягом 5 днів) для декальцифікації. Розчин ЕДТА замінювали, щонайменше через два дні. Після декальцифікації, тканину інкубували в 15 % сахарозі протягом ночі при кімнатній температурі, і потім занурювали в Tissue Mount (Chiba Medical, Soka, Japan). Зрізи товщиною десять мікрметрів були нарізані в шести різних положеннях у кожній мишачій голові (~0-6 мм від кінця носової кістки). Експресію GFP спостерігали, застосовуючи флуоресцентний мікроскоп (Leica). Кількісну оцінку й ідентифікацію клітинних типів виконували на шести рівнях на мишу, використовуючи об'єктив $\times 40$ або $\times 63$. Пролонгована експозиція зображення була необхідною, щоб зафіксувати структуру назального епітелію з використанням флуоресцентної мікроскопії. Це привело до піксельного насичення GFP-позитивних клітин і викликало те, що GFP-позитивні клітини здаються майже білими, у більшому ступені, ніж мають звичайний зелений зовнішній вигляд, що автори винаходу й інші, спостерігають при більш високому збільшенні.

Клітинний розподіл експресії трансгена F/NH-SIV досліджували в гістологічних зрізах. Конкретно, експресію EGFP визначали в гістологічних зрізах порожнини носа миші (2 мм від кінчика носа) через 30 днів після обробки. Фіг. 4 показує розташування експресії EGFP (Фіг. 4, білий точковий сигнал).

Фіг. 5 показує клітинні типи, трансдуковані за допомогою обробки трансгеном /NH-SIV носа миші. 69 % клітин, трансдукованих у порожнині носа миші, являли собою війчасті респіраторні епітеліальні клітини. Інші трансдуковані клітинні типи включали нейрональні клітини в нюховому епітелії (21 %) і сквамозні клітини (7 %).

Щоб знайти базальні клітини після обробки полідоканолом, виконували імунозabarвлення на основі пероксидази хрому (HRP), використовуючи набір Envision (Dako, Glostrup, Denmark). Коротко, предметні стекла обробляли 0,6 % пероксидом водню в метанолі протягом 15 хв, промивали у водопровідній воді і інкубували з 1,5 % нормальною козячою сироваткою (Abcam) протягом 30 хв. Предметні стекла потім інкубували з кролячими поліклональними антитілами проти Цитокератину 5 (1:500) (Abcam) протягом 1 години з наступною інкубацією з козячими антитілами проти кролячих IgG, кон'югованими з HRP (наданими разом з набором) протягом 30 хв. Предметні стекла потім промивали в PBS і інкубували із субстратом пероксидази, 3-аміно-9-етилкарбазолом (AEC) (наданим разом з набором) протягом 5 хв. Остаточо, предметні стекла промивали в дистильованій H_2O , дозabarвлювали водним гематоксилином Харріса (BDH) протягом 15 секунд, промивали у водопровідній воді і потім у дистильованій H_2O .

Імунофлуоресцентне виявлення GFP-позитивних трансдукованих назальних епітеліальних клітин і Krt5 позитивних базальних клітин виконували, використовуючи наступні первинні і вторинні антитіла: козячі поліклональні антитіла проти GFP (1:250) (Abcam), кролячі моноклональні антитіла проти KRT5 (1:500) (Abcam), ослині антитіла Alexa Fluor 488 проти козячих IgG (1:200) (Invitrogen, Paisley, UK) і козячі антитіла Alexa Fluor 594 проти кролячих IgG (1:200) (Invitrogen). Щоб поліпшити характеристики антитіл, зрізи піддавали опосередкованому нагріванню демаскуванню антигену в цитратному буфері (10 mM лимонної кислоти, 0,05 % Твін 20, рН 6,0) протягом 20 хв на водяній бані протягом 100 °C. Забарвлені зрізи встановлювали в реагенті, що перешкоджає вигорянню флуоресценції ProLong® Gold Antifade, разом з DAPI (Invitrogen) і аналізували з використанням конфокального мікроскопа, як раніше (усі на Zeiss). GFP-позитивні базальні клітини (ідентифікація основана на морфології і розташуванні в епітеліальному шарі) кількісно оцінювали усього на 13 зрізах/миша. Зрізи, які проявляли передбачувані GFP-позитивні базальні клітини, були відібрані для подвійного забарвлювання з використанням антитіл проти KRT5 і проти GFP, щоб підтвердити фенотип базальних клітин.

Приклад 11: Трансдукція культур ALI

Повністю диференційовані епітеліальні клітини дихальних шляхів, вирощені як культури ALI, купували у Epithelix (Geneva, Switzerland). ALI трансфікували з використанням F/HN-SIV-lux при множинності зараження в інтервалі від ~25 до ~300 ОТ/клітину. Через 6 годин, вірус видаляли і ALI інкубували протягом 10-26 днів. Базолатеральне середовище замінювали кожні 48 годин під час цього періоду інкубації. У встановлені відмітки часу, ALI лізували в 100 мкл буфера для лізису репортера й експресію люциферази кількісно оцінювали, використовуючи систему для аналізу люциферази (Promega, Southampton, UK) відповідно до інструкцій виробника. Вміст загального білка в культурах кількісно оцінювали, використовуючи набір для аналізу білка BioRad (BioRad, Hemel Hempstead, UK). Кожен зразок аналізували в двох повторях. Експресію люциферази потім представляли у вигляді відносних світлових одиниць/мг загального білка. Для біоломінесцентної візуалізації 100 мкг люциферину в PBS додавали до апікальної мембрани.

Як показано на Фіг. 7, клітини в культурах ALI, були успішно трансдуковані з використанням F/HN-SIV-Lux, про що свідчить експресія люциферази. Експресія люциферази була вища при MOI 250 у порівнянні з MOI 25.

Приклад 12: Аналіз ефлюксу йодиду

Клітини HEK293T трансфікували з використанням F/HN-SIV-GFP-CFTR або контрольного вірусу F/HN-SIV-GFP при множинності зараження, яка дорівнює 500 ОТ/клітину і культивували протягом 2 днів. Активність хлоридного каналу CFTR аналізували за допомогою вимірювання ступеня ефлюксу ^{125}I йодиду, як описано раніше (Derand, R., et al., 2003). Швидкості ефлюксу ^{125}I йодиду нормалізували за часом додавання суміші форсколін/IBMX (час 0). Криві були побудовані за допомогою нанесення на графік залежності швидкості ефлюксу ^{125}I йодиду від часу. Щоб відбити кумулятивні рівні ефлюксу ^{125}I йодиду після стимуляції агоністом, усі порівняння основані на площях під кривою залежності час-ефлюкс ^{125}I йодиду. Площу під кривою розраховували за правилом трапеції. Експерименти виконували в двох повторях (n=6 ямок/групу/експеримент).

Як показано на Фіг. 8, F/HN-SIV може спрямовувати функціональну експресію CFTR, з відносним ефлюксом ^{125}I йодиду, який дорівнює 0,65. Навпаки, плазмідний вектор GL67A дозволяв досягти більш низького значення ефлюксу ^{125}I йодиду, який дорівнює 0,33.

Приклад 13: Трансдукція овечих і людських первинних клітин легені і легені миші

F/HN-SIV ефективно трансдукує овечі і людські первинні клітини легень і легеню миші.

F/HN-SIV-CMV-Lux застосовували, щоб трансдукувати людські клітини назальних зіскрібків (MOI 250) і клітини зрізів людської й овечої легені, культивовані *ex vivo* (1×10^7 ОТ/slice). Як показано на Фіг. 9A, трансдукція як людських клітин назальних зіскрібків, так і клітин зрізів людської й овечої легені приводила до істотної експресії трансгена люциферази (середні значення в області 2×10^2 BCO/мг білка для людських назальних зіскрібків, 1×10^7 BCO/мг білка для зрізів людської легені і 2×10^7 BCO/мг білка для зрізів овечої легені) 24-48 годин після трансдукції.

Первинні клітини легені людини з MB, культивовані на межі розділення фаз повітря-рідина (CF hALI, $\sim 1 \times 10^5$) трансдукували з використанням (3×10^7 ОТ) векторів F/HN-SIV-soCFTR2, що містять промотори CMV- і hCEF-трансгена. Число копій вектора (копій провірусної ДНК на копію ендогенної ДНК CFTR) вимірювали через 6-8 днів після трансдукції. Промотори, як CMV, так і hCEF були здатні досягати числа копій вектора, що дорівнює щонайменше 1×10^1 (Фіг. 9B).

Також вимірювали рівень експресії мРНК CFTR (%VE: копій мРНК CFTR на копію ендогенної мРНК CFTR $\times 100$) через 6-8 днів після трансдукції. Горизонтальна пунктирна лінія на Фіг. 9C представляє цільовий рівень експресії, який дорівнює 5 % VE, що, як вважають, представляє терапевтичний поріг. Як F/HN-SIV-soCFTR-CMV, так і F/HN-SIV-soCFTR2-hCEF індукували експресію, яка була істотно вищою цієї цільової оцінки (в області 40583 ± 10687 і 18509 ± 13588 відповідно, середнє \pm SD, n=4).

Після доставки *in vivo* векторів F/HN-SIV-EGFPLux, що містять CMV, EF1a і промотори hCEF у дефіцитній по інтегразі (ID) або компетентній по інтегразі формі (IC або без мітки), у клітини дихальних шляхів, експресію трансгена визначали в назальному епітелії (Фіг. 9D) і епітелії легень (Фіг. 9E) мишей (n=6-10/група). Часову динаміку експресії трансгена люциферази реєстрували за допомогою повторної біоломінесцентної візуалізації *in vivo* і нормалізували по дозі, що доставляється. Чотири з п'яти тестованих векторів (vGM012 CMV, vGM014 EF1a, vGM020 hCEF і vGM076 hCEF ID) дозволяли досягти експресії в носі вище цільового рівня протягом усього часу протікання експерименту. П'ятий вектор, vGM074 CMV ID, дозволяв досягти експресії в носі вище прийнятого рівня експресії протягом усього часу протікання експерименту.

Два з п'яти тестованих векторів (vGM014 EF1a і vGM020 hCEF) дозволяли досягти експресії в легенях вище цільового рівня протягом усього часу протікання експерименту. Один вектор, vGM012 CMV, дозволяв досягнути експресії в легенях вище прийнятого рівня експресії протягом усього часу протікання експерименту.

Біоломінесценцію виявляли після трансдукції мишам *in vivo* на день 14 після трансдукції. Представницькі зображення показані на Фіг. 9F. З п'яти тестованих векторів вектор vGM020 hCEF дозволяв досягти найвищого рівня експресії *in vivo*.

Біоломінесценцію також виявляли після трансдукції *in vitro* не-MB hALI на день 5-6 після трансдукції. Представницькі зображення показані на Фіг. 9G. З п'яти тестованих векторів, знову vGM020 hCEF дозволяв досягти найвищого рівня експресії.

Фіг. 9H показує експресію EGFP через 14 днів після трансдукції в мишачому назальному епітелії після доставки $1,6 \times 10^8$ ОТ F/HN-SIV-hCEF-EGFPLux (vGM020), як візуалізовано за допомогою імуногістохімії (ядра, забарвлені DAPI).

Часову динаміку експресії трансгена люциферази в не-Мв ALI реєстрували за допомогою повторної біолюмінесцентної візуалізації і нормалізували по дозі, що доставляється. Як показано на Фіг. 9I, вектори vGM014 EF1a і vGM020 hCEF дозволяли досягти найвищого рівня експресії.

F/HN-SIV також ефективно трансдукує овечу легеню *in vivo*. Акрифлавін інстилювали (аліквоти 3×100 мкл протягом (5 хвилин) у проксимальні дихальні шляхи при прямій бронхоскопічній візуалізації. Розподіл акрифлавіну можна було оцінити по жовтогарячому забарвленню вирізаних дихальних шляхів після розкриття (Фіг.10A).

Акрифлавін був у значній мірі обмежений провідними повітроносними шляхами і був відсутній в альвеолярних областях. Стрілка на Фіг.10A указує приблизну ділянку інстиляції. Числа на лінійці приведені в см.

Фіг. 10B є представленням у вигляді діаграми овечої легені (центр/верхівка трахеї). Коло представляє область на (Фіг. 10A). На Фіг. 10B стрілка вказує проходження бронхоскопу для доставки аліквот 3×100 мкл 2,2E9 ОТ/мл (6,6E8 ОТ усього) F/HN SIV CMV-EGFPLux n=3 окремим вівцям (коди тварин T121, T156 і T251). Через сім днів після доставки, 5-6 блоків зразків тканин були узяті після розкриття при ~1 см інтервалах від ділянки інстиляції.

Блоки зразків були розділені на 2-3 приблизно еквівалентних зразка і піддані аналізам на предмет експресії трансгена, результати яких показані на Фіг. 1C у вигляді аналізів люциферази, нормалізованих по вмісту білка; і на Фіг. 10D у вигляді кількісної РВ-ПЛР, нормалізованої по рівнях ендогенної мРНК CFTR. Горизонтальна лінія на Фіг. 1C представляє найвищу активність люциферази, відзначену в якого-небудь зразка, обробленого невірусним вектором перенесення генів, і на Фіг. 10D цільовий рівень експресії, який дорівнює 5 % VE (як вважають, представляє терапевтичний поріг, див. вище). Для кожної групи обробки середнє значення було вище, ніж порівняльне значення для невірусного вектора (Фіг. 10C), і також досягало експресії вище цільового 5 % порога VE.

Приклад 14: Експресія і функція CFTR, виміряні з використанням людських кишкових органоїдів при CF

Кишкові органоїди людини з МВ генерували, як описано Dekkers JF et al, (Nature Medicine 2013, 19(7): 939-945). Коротко, кишкові біопсії промивали в розчинах, які містять ЕДТА, для дисоціації клітин крипти. Клітини крипти потім трансдукували з використанням vGM058 (приблизно 1×10^7 одиниць трансдукції) або контрольного вірусу (n=3 ямок/стан), заливали в Matrigel і давали можливість утворити органоїди протягом 3-4 днів. Функцію CFTR оцінювали за допомогою піддавання органоїдів впливу форсколіну (приблизно 5 мкМ форсколіну), що збільшує внутрішньоклітинні рівні цАМФ і наступним чином активує CFTR. У відповідь на активацію CFTR органоїди збільшують транспорт хлориду, що приводить до поглинання води і набрякання. Органоїди (мінімально 10/ямку) безпосередньо аналізували за допомогою конфокальної мікроскопії живих клітин (LSM710, Zeiss, ×5 об'єктив). Стимульоване форсколіном набрякання органоїдів автоматично кількісно визначали з використанням програмного забезпечення Volocity для візуалізації (Improvision). Збільшення загальної площі органоїдів (площина ху) відносно площі при t=0 обробки форсколіном розраховували й усереднювали. Індуковане форсколіном набрякання значно ($p < 0,001$) збільшувалося в трансдукованих vGM058 органоїдах у порівнянні з контролями (див. Фігуру 14A). Значні рівні CFTR-опосередкованого ефлюксу йодиду ($p < 0,05$) були також виявлені в трансдукованих vGM058 клітинах (Фіг. 14B) з використанням аналізу ефлюксу йодиду, розкритого в даному описі (див. Приклад 12).

Приклад 15: Генерація лентивірусних векторів для A1AT

Лентивірусні вектори одержували, використовуючи кістак SIV і 5-плазмідний метод, описаний вище в Прикладах 2 і 3 (для лентивірусу CFTR) і з використанням промотору hCEF, як описано в даному описі. Генерували дві окремі лентивірусні конструкції: одну з трансгеном людського альфа-1-антитрипсину (hAAT); одну з трансгеном *Gaussia luciferase* (Glux) (див. Фігуру 15A і B, SEQ ID NO: 9 і 10 відповідно). кДНК, що містяться в цих векторах, були кодон-оптимізованими і CpG-збідненими.

Приклад 16: Культивування на межі розділення фаз (ALI) з використанням лентивірусного вектора гена-репортера lux

Повністю диференційовані дикого типу людські ALI культури (MucilAir) купували у Epithelix SARL (Geneva, CH). ALI культивували при 37°C і 5 % CO₂, і базолатеральне культуральне середовище замінювали кожні 2-3 дні. Культуральне середовище зберігали при -20°C до додаткового аналізу.

ALI трансдукували (у день 0) за допомогою піпетування 100 мкл вірусу (1×10^7 одиниць трансфекції Tagman (EET) на апікальну поверхню. Вірус видаляли через 4 години інкубації при 37 °C, і базолатеральне середовище замінювали.

У вказані відмітки часу після трансдукції, апікальну поверхню ALI промивали за допомогою інкубації зі стерильним PBS протягом однієї години. Промивання видаляли і зберігали при -20°C до подальшого аналізу.

Аналіз люциферази Gaussia (New England Biolabs, Ipswich, USA) виконували відповідно до рекомендацій виготовлювача. 15 мкл зразка аналізували в двох повторях, і люмінесценцію визначали на зчитувальному пристрої для планшетів Appliskan. Експресію Glux виражали у вигляді BCO/мкл рідини (BCO=відносні світлові одиниці).

Фіг. 16 представляє результати, причому кожна точка представляє середнє значення BCO/мкл у середовищі для $n=6$ ALI при показаній часовій оцінці (стандартна помилка вказана за допомогою планок погрішностей). Опосередковане лентивірусом перенесення генів на поверхнях розділення фаз повітря-рідини в людини приводив до тривалої експресії секретованого репортерного білка люциферази Gaussia.

Приклад 17: Експресія трансгена в культурах зрізів легені з використанням A1AT і лентивірусних векторів гена-репортера і lux

Нарізані з високою точністю зрізи легені людини одержували, як описано в Moreno L et al, Respir Res 2006 Aug 21; 7:111. Зрізи легені вміщували в 12-ямкові планшети для культур тканин (1 зріз на ямку) у 1 мл середовища і інкубували при 37°C і 5 % CO₂. Середовище змінювали щодня і зберігали при -20°C до подальшого аналізу.

У день 0 зрізи легені ($n=6$ на групу) трансдукували з використанням вірусу SIV hCEF-sogLux (1 × 10⁶ ЕТТ) або вірусу SIV1 hCEF-sohAAT (2 × 10⁶ ЕТТ), розведеного в середовищі до кінцевого об'єму, який дорівнює 1000 мкл і інкубували протягом 4 годин. Після інкубації, середовище замінювали і зберігали при -20°C до подальшого аналізу.

Експресію люциферази Gaussia визначали, як описано вище (Приклад 16). Як показано на Фіг. 17A, високі рівні експресії секретованого репортерного білка люциферази Gaussia супроводжували опосередковане лентивірусом перенесення генів у зрізи легені людини. Експресію AAT (який також називається в даному описі як A1AT) визначали, використовуючи сендвіч-ELISA (Abcam, Cambridge, UK), виконуваний відповідно до рекомендацій виготовлювача. 50 мкл зразка аналізували в двох повторях, і вимірювали на зчитувальному пристрої для мікропланшетів при 450 нм. Як показано на Фіг. 17B, високі рівні експресії альфа-1-антитрипсину (AAT/A1AT) супроводжували опосередковане лентивірусом перенесення генів у зрізи легені людини.

Приклад 18: In vivo уведення A1AT і лентивірусних векторів гена-репортера lux у ніс миші

Трансдукція мишачої легені:

Самиць мишей C57BL/6 (Charles River, UK) анестезували ізофлураном і їм давали 100 мкл вірусу за допомогою назальної інстиляції, як описано в Xenariou S et al, Gene Ther 2007 May;14(9): 768-75. Тваринам давали між 1 і 5 доз і спостерігали щодня на предмет ознак токсичності.

У випадку люциферази Gaussia, самиць мишей C57BL/6 анестезували на день 0, використовуючи ізофлуран, і їм давали однократну дозу 100 мкл вірусу SIV1 hCEF-soGLux (1×10⁶ ЕТТ) за допомогою назальної інстиляції. Контрольних тварин обробляли DMEM (середовище для культур тканин), основним компонентом препарату вірусу, використововуваного в дослідженні.

У випадку A1AT, самиць мишей C57bl/6 ($n=5$ на групу) обробляли 3 дозами SIV1 hCEF-sohAAT при 10-денних інтервалах (100 мкл на дозу, 6,8×10⁷ ЕТТ; загальна доза 2,4×10⁸ ЕТТ). Контрольних тварин інстилювали з використанням 100 мкл стерілізованого PBS (основного компонента завантаження отриманого лентивірусу, використововуваного в дослідженні) у кожній відмітці дозування.

Через 10 днів після третьої дози, тварин умертвляли і гомогенат тканини легені, рідину бронхоальвеолярного лаважу і сироватку аналізували на предмет експресії AAT.

Додатково, досліджували довгострокову експресію A1AT. У дні 1-5 експерименту, мишей C57bl/6 обробляли 100 мкл SIV1 hCEF-sohAAT за допомогою назальної інстиляції (5 доз по 4 × 10⁵ ЕТТ, тобто всього 2×10⁶ ЕТТ на тварину). Контрольних тварин інстилювали 100 мкл DMEM (середовище для культури тканини), основного компонента завантаження отриманого лентивірусу, використововуваного в дослідженні. Тварин умертвляли при різних оцінках часу після трансдукції, і гомогенат тканини легені, рідину бронхоальвеолярного лаважу рідину і сироватку аналізували на предмет експресії AAT.

Збирання мишачої тканини:

Мишей умертвляли при вказаних часових оцінках після трансдукції. Кров збирали за допомогою пункції лівого шлуночка, і центрифугували при 760 g протягом 10 хвилин для одержання сироватки. Сироватку послідовно заморожували при -80 °C.

Бронхоальвеолярний лаваж (BAL) виконували вирізанням ший, введенням канюлі в трахею і закріпленням її на місці за допомогою провідника шовної нитки. 500 мкл PBS інстилювали в легеню, і аспірували три рази для одержання ретельного промивання епітеліальної вистілки. Зразок негайно швидко заморожували в рідкому азоті, і зберігали при -80°C для подальшого аналізу.

Легені потім вирізали і швидко заморожували в рідкому азоті, і послідовно гомогенізували в пробірках з лізуючим матриксом D (MP Biomedicals), центрифугували в машині FastPrep (ThermoFisher Scientific, Waltham, MA, USA) при 4 м/с протягом 45 секунд, і зберігали при -80°C для подальшого аналізу.

Експресію AAT (A1AT) і люциферази Gaussia (Gluc) визначали, як описано в Прикладах 16 і 17 вище. *In vivo* трансдукція клітин мишачих дихальних шляхів з використанням однократної дози лентивірусного вектора гена-репортера lux Приклада 15 приводила до довгострокової експресії (щонайменше 12 місяців) секретованого репортерного білка люциферази Gaussia, у гомогенаті легень (Фіг. 18A), рідині бронхоальвеолярного лаважу (BAL, Фіг. 18B) і сироватці (Фіг. 18C).

Високі рівні експресії A1AT спостерігали в гомогенаті легень, BAL і сироватці після опосередкованого лентивірусом перенесення гена AAT (A1AT) *in vivo* (Фіг. 19), зі спостережуваним більше ніж 100-кратним збільшенням експресії AAT (A1AT) у гомогенаті легень і BAL у порівнянні з відповідними негативними (PBS) контролями. Значне збільшення (щонайменше на один порядок величини) експресії AAT (A1AT) також спостерігали в сироватці.

Додатково, тривалу експресію (щонайменше протягом 90 днів) альфа-1-антитрипсину спостерігали в гомогенаті легень (Фіг. 20A), BAL (Фіг. 20B) і сироватці (Фіг. 20C) після опосередкованого лентивірусом перенесення гена AAT (A1AT) *in vivo*.

Приклад 19: Аналіз сечі

Мишей C57bl/6 ($n=5$ на групу) обробляли 3 дозами SIV1 hCEF-sohAAT при 10-денних інтервалах (100 мкл на дозу, $6,8 \times 10^7$ ЕТТ; сумарна доза $2,4 \times 10^8$ ЕТТ). Контрольних тварин інстилювали з використанням 100 мкл стерилізованого PBS (основного компонента отриманого завантаження лентивірусу, використовуваного в дослідженні) при кожній відмітці дозування.

Через 10 днів після третьої дози, тварин умертвляли і гомогенат тканини легень, рідину бронхоальвеолярного лаважу і сироватку аналізували на предмет експресії A1AT.

Аналіз сечі (Abcam, Cambridge, UK) виконували відповідно до інструкцій виготовлювача.

По-перше, одержували серійні розведення зразків мишачої сироватки і рідини BAL і аналізували їх, щоб визначити відповідні розведення для застосування в подальших експериментах.

По-друге, аналізували відповідні зразки сироватки і рідини BAL від окремих мишей ($n=14$), щоб розрахувати кратність відмінності між концентрацією сечі в сироватці і рідині BAL, еквівалентну ефекту розведення BAL на бронхоальвеолярну лаважну рідину (по Rennard SI et al, J Appl Physiol (1985). 1986 Feb; 60(2):532-8). Середнє розведення BAL було 41-кратним (інтервал 24-88).

Приймаючи до уваги цей ефект розведення, розраховували концентрацію AAT (A1AT) у бронхоальвеолярній лаважній рідині. Конкретно, концентрацію AAT у рідині бронхоальвеолярного лаважу помножували на фактор розведення, щоб одержати оцінку "справжньої" концентрації AAT у бронхоальвеолярній лаважній рідині.

"Захисний" цільовий рівень AAT (A1AT) у бронхоальвеолярній лаважній рідині (ELF, тобто рідини, яка вистилає повітряні шляхи і повітряний простір у легенях) складав 70 мкг/мл (у порівнянні з "нормальним" цільовим рівнем AAT (A1AT) у ELF, який дорівнює 200 мкг/мл). Як показано на Фіг. 21, терапевтичні рівні альфа-1-антитрипсину в бронхоальвеолярній лаважній рідині після опосередкованого лентивірусом перенесення генів AAT (A1AT) *in vivo*.

Приклад 20: Генерація лентивірусних векторів для FVIII

Чотири різні лентивірусні вектори FVIII одержували, використовуючи кістяк SIV і 5-плазмідний метод, описаний вище в Прикладах 2 і 3 (для лентивірусу CFTR) і 15 (для лентивірусу A1AT). Промотор-трансгенні плазміди мали SEQ ID NO: 11-14, відповідно.

Послідовність SIV була ідентичною послідовностям конструкцій CFTR (Приклади 2 і 3) за винятком промотору і кДНК. Промотор людського цитомегаловірусу (CMV) або тканиноспецифічний промотор/енхансер hCEFI застосовували, як вказано (Фіг. 22), для спрямування експресії трансгенів FVIII.

SIV-F/HN-FVIII-N6-ко містив людську кДНК FVIII дикого типу, з якої домен BDD був видалений за допомогою делеції і замінений на кодон-оптимізований 226-амінокислотний фрагмент 6N-глікозилування.

SIV-FVIII-V3 містить людську кДНК FVIII дикого типу, з якої 226-амінокислотний сайт глікозилювання був видалений за допомогою делеції і замінений на 17-амінокислотний пептид, який експресує тріади 6N-глікозилювання усередині В-домену (McIntosh et al., Blood 2013 121(17); 3335-3344).

5 Приклад 21: Кількісна оцінка антигену hFVIII і рівнів активності на моделях *in vivo* і *in vitro*

Рівні людського антигену FVIII у мишачої моделі кількісно визначали за допомогою ферментозв'язаного імуносорбентного аналізу (ELISA) відповідно до протоколу виготовлювача. Коротко, плазму, BAL і легень аналізували на предмет присутності антигену FVIII, використовуючи Asserachrom (FVIII:Ag) Elisa (Stago Diagnostics, France).

10 Зразки розбавляли 1:2 і інкубували в 96-ямковому планшеті, покритому мишачими моноклональними антитілами проти фрагмента фактора VIII протягом 2 годин при кімнатній температурі. Після промивання, у планшет додавали вторинні проти мишачих антитіла, зв'язані з пероксидазою, і інкубацію виконували протягом 2 годин при КТ. hFVIII: рівні Ag визначали спектрофотометрично при 450 нм, використовуючи субстрат TMB (дані не показані).

15 Ще один аналіз ELISA застосовували для оцінки активності FVIII на моделі HEK293T *in vitro* (FVIII:C, Affinity Biological, Canada). Супернатанти збирали через 48 і 72 годин після трансдукції HEK293T з використанням SIV-F/HN-FVIII-N6 або SIV-F/HN-FVIII-N3. Активність FVIII оцінювали відповідно до наступних інструкцій виготовлювача, використовуючи 50 мкл супернатантів, аналізовані в двох повторях. Як негативний контроль тестували супернатант від необроблених клітин HEK293T. Активність hFVIII розраховували по стандартній кривій, генерованій з використанням ряду розведень нормальної людської об'єднаної плазми (13-ий Британський Стандарт для концентрату Фактора VIII коагуляції крові, Людський; NIBSC).

20 Клітини HEK293T трансдукували з використанням двох різних завантажень vGM142 (Завантаження $1.5,9 \times 10^8$ ЕТТ/мл і Завантаження $2.2,8 \times 10^8$ ЕТТ/мл). Клітини HEK293T трансдукували з використанням Завантаження 1 вектора vGM142 (Фіг. 23А) і Завантаження 2 вектора vGM142 (Фіг. 23В) при 3 різних MOI (MOI1;10;100), і збирали через 48 і 72 годин після трансдукції.

Як зрозуміло видно на Фіг. 23, збільшення активності FVIII спостерігали зі збільшенням MOI як для Завантаження 1, так і Завантаження 2 vGM142 як через 48 годин, так і через 72 години після трансдукції. Крім того, активність FVIII збільшувалася в інтервалі від 48 годин до 72 годин для кожного тестованого MOI.

Приклад 22: Уведення лентивірусних векторів FVIII і гена-репортера lux *in vivo* у ніс миші

Трансдукція в мишачу легеню:

35 Усі методики за участю тварин виконували відповідно до умов і обмеження UK Home Office Project і персональних ліцензованих правил відповідно до Animal Scientific Procedure Act (1986).

Самиць мишей C57BL/6 дикого типу у віці 6-8 тижнів (Charles River, UK) анестезували, використовуючи ізофлюран, і їм давали 100 мкл вірусу в забуференому фосфатом фізіологічному розчині по Дульбекко (D-PBS), як описано раніше (Griesenbach et al., 2012), і оцінювали присутність антигену FVIII.

40 У двох експериментах (Група 1 і 2) миші одержували 3 дози (через день) SIV-F/HN-FVIII-N6 (vGM142) і їхній умертвляли через 10 днів після першої дози. Група 1 (n=4) була оброблена загальною дозою вектора, яка дорівнює $1,4 \times 10^6$ ЕТТ/миша. Групу 2 (n=3) обробляли загальною дозою вектора, яка дорівнює $1,57 \times 10^8$ ЕТТ/миша.

45 В одному експерименті (Група 3) мишей обробляли 12 дозами (через день) SIV-F/HN-FVIII-N6 (vGM142) і умертвляли через 28 днів після першої дози. Групу 3 (n=4) обробляли загальною дозою вектора, яка дорівнює $3,36 \times 10^8$ ЕТТ/миша.

50 Плазму, рідину BAL і легень збирали (як описані в Прикладі 14). Коротко, мишей умертвляли у вказані часові оцінки після трансдукції. Кров потім збирали із серця в пробірки для збору з антикоагулянтом, 3,2 цитрати натрію, перед центрифугуванням при 2000-2500 x g для одержання плазми. Рідину BAL збирали, застосовуючи 3 послідовні інстиляції PBS (500 мл) у мишачі легень при кімнатній температурі. Супернатанти зберігали при - 80 °C. Легень збирали і зберігали при - 80 °C перед гомогенізацією тканини.

55 Потім оцінювали наявність експресії FVIII. Рівні FVIII оцінювали в гомогенатах тканин легень (Фіг. 24А). Рідину BAL (Фіг. 24В) і плазму (Фіг. 24С) збирали окремо в 3 незалежних експериментах через 10 і/або 28 днів після обробки SIV-F/HN-FVIII-N6. Аналіз виконували з використанням однофакторного дисперсійного аналізу Anova (GraphPad Prism) з множиною порівнянь між групами обробки (****p<0,0001).

60 Як зрозуміло видно на Фіг. 24А, усі три групи обробки виконували спостережуване збільшення рівнів hFVIII у тканині легень в порівнянні з відповідним контролем (D-PBS). 28-денна обробка Групи 3 приводила до значного збільшення експресії hFVIII у порівнянні з 10-

денними обробками Груп 1 і 2. Аналогічні результати спостерігали для зразків BAL рідини (Фіг. 24В), незважаючи на те, що в цих зразках також було значне збільшення рівнів hFVIII у Групи 2 у порівнянні з Групою 1. Обробка Групи 3 приводила до значного збільшення рівнів hFVIII у плазмі (Фіг. 24С).

5 Приклад 23: Спосіб виробництва лентивірусних векторів відповідно до винаходу
Клітини HEK293 вирощують у суспензії, у середовищі для експресії Freestyle (хімічно визначеному, яке не містить компонентів тваринного походження і білків), і виконують підрахунок клітин. Визначають концентрацію глюкози і виконують титрування до ~35 мМ. Готують суміш для трансфекції пДНК/PEIPro™, і клітини трансфікують при 0,33 мг пДНК/1Е9 клітин.

10 Підрахунок клітин виконують знову і додають додаткове середовище для експресії Freestyle. Повторно визначають концентрацію глюкози і виконують титрування до ~35 мМ. Можна додати 5 мкл/мл Бензонази® і виконують тристадійне лінійне очищення вірусу. Додають Бензоназу® з наступним додаванням TrypLE Select™. Вірус охолоджують до 0 °С і зберігають на вологості льоду для всіх наступних стадій. Після відфільтровування якої-небудь невірусної речовини у формі частинок (фільтр mPES 0,45 мкм), вірус завантажують на Mustang® Q XT (3 мл мембрани/л очищеного вірусу) з наступним промиванням 0,15М NaCl Tris pH 7,5 і елюванням 1,0 М NaCl Tris pH 7,5. Фракцію вірусу збирають і розбавляють до 0,1-0,2 від первинного об'єму середовищем Freestyle. TrypLE Select™ може бути доданий тепер, якщо його не додавали раніше, і Бензоназа® також може бути додана на цій стадії на додаток до або замість доданої раніше.

20 Виконують кругову фільтрацію з тангенціальним потоком (TFF) (УФ до ~0,1-0,05 первинного об'єму=HV; ДФ ультраконцентрату x5 HV проти буфера для лікарської форми; УФ до ~0,001-0,002 від первинного об'єму), і ультраконцентрат збирають. Друга стадія TFF може виконуватися, і може застосовуватися більш дрібний по розміру блок TFF для ДФ і/або кінцевої УФ. Додаткові стадії можуть включати змішаний режим/SEC і стерильну фільтрацію через 0,45 мкм або 0,2 мкм.

25 Фіг. 11 зображує одержання й очищення F/HN-векторів SIV. F/HN-вектори SIV одержували за допомогою 5-плазмідної (пДНК) PEI-опосередкованої транз'єнтної трансфекції клітин 293Т, вирощених у суспензії в масштабі 1 л у біореакторах WAVE з контрольованим рН (GE), застосовуючи масштабовані способи винаходу. Вектори очищали за допомогою глибинної/остаточної фільтрації (GE/Pall), забруднювальні нуклеїнові кислоти видаляли з використанням Бензонази® (Merck), вектори активували за допомогою TrypLE Select™ (Life Technology), очищали і концентрували за допомогою аніонообмінної мембранної хроматографії (Pall) і тангенціальної потокової фільтрації (Spectrum). Усі ємності, контейнери і колонки для процесів відповідали вимогам cGMP для одноразового застосування. Усі реагенти, за винятком плазмідної ДНК, відповідали вимогам cGMP до продуктів, які не містять компонентів тваринного походження. Показано дані для різних конфігурацій вектора (промотор трансгена, трансген, статус інтегрази). Фізичні і функціональні титри визначали, використовуючи К-ПЛР.

40 Див. результати даного ілюстративного способу за винаходом, обговорювані при описі Фігури 11, приведеному вище.

Ключ до SEQ ID NO

SEQ ID NO: 1 Плазміда, визначена на Фіг. 1А (пДНК1 pGM326)
SEQ ID NO: 2 Плазміда, визначена на Фіг. 1В (пДНК2а pGM297)
45 SEQ ID NO: 3 Плазміда, визначена на Фіг. 1С (пДНК2b pGM299)
SEQ ID NO: 4 Плазміда, визначена на Фіг. 1D (пДНК3а pGM301)
SEQ ID NO: 5 Плазміда, визначена на Фіг. 1Е (пДНК3b pGM303)
SEQ ID NO: 6 Ілюстрований промотор hCEF
SEQ ID NO: 7 Ілюстрований трансген CFTR (soCFTR2)
50 SEQ ID NO: 8 Ілюстрований компонент WPRE (mWPRE)
SEQ ID NO: 9 Плазміда F/HN-SIV-hCEF-so1AT, визначена на Фіг. 15 (пДНК1 pGM407)
SEQ ID NO: 10 Плазміда F/HN-SIV-hCEF-sogLux, визначена на Фіг. 15 (пДНК1 pGM358)
SEQ ID NO: 11 Плазміда F/HN-SIV-CMV-HFVIII-V3, визначена на Фіг. 22С (пДНК1 pGM411)
SEQ ID NO: 12 Плазміда F/HN-SIV-hCEF-HFVIII-V3, визначена на Фіг. 22D (пДНК1 pGM413)
55 SEQ ID NO: 13 F/HN-SIV-CMV-HFVIII-N6-коплазміда, визначена на Фіг. 22Е (пДНК1 pGM412)
SEQ ID NO: 14 F/HN-SIV-hCEF-HFVIII-N6-коплазміда, визначена на Фіг. 22F (пДНК1 pGM414)
SEQ ID NO: 15 Ілюстрований трансген A1AT
SEQ ID NO: 16 Ілюстрований трансген FVIII (N6)
SEQ ID NO: 17 Ілюстрований промотор CMV
60 SEQ ID NO: 18 Праймер для побудови pCAGGS-Fct4

SEQ ID NO: 19 Праймер для побудови рCAGGS-Fct4
 SEQ ID NO: 20 Праймер для побудови рCAGGS-Fct4
 SEQ ID NO: 21 Праймер для побудови рCAGGS-Fct4
 SEQ ID NO: 22 Праймер для побудови рCAGGS-Fct4
 5 SEQ ID NO: 23 Праймер для побудови рCAGGS-Fct4
 SEQ ID NO: 24 Праймер для побудови рCAGGS-SIVct+HN
 SEQ ID NO: 25 Праймер для побудови рCAGGS-SIVct+HN
 SEQ ID NO: 26 Ланцюг, комплементарний до ілюстрованого трансгену A1AT
 SEQ ID NO: 27 Ілюстрований пептид A1A1
 10 SEQ ID NO: 28 Ланцюг, комплементарний до ілюстрованого трансгену FVIII (N6)
 SEQ ID NO: 29 Ілюстрований пептид FVIII (N6)
 SEQ ID NO: 30 Ілюстрований трансген FVIII (V3)
 SEQ ID NO: 31 Ланцюг, комплементарний до ілюстрованого трансгену FVIII (V3)
 SEQ ID NO: 32 Ілюстрований пептид FVIII (V3)
 15 SEQ ID NO: 33 Ланцюг, комплементарний до ілюстрованого промотору CMV
 Послідовності
 SEQ ID NO: 1
 1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT AAATCAATAT
 61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT ATATTGGCTC
 20 121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
 AGTAATCAAT
 181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
 TTACGGTAAA
 241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
 25 TGACGTATGT
 301 TCCCATAGTA ACGCCAATAG GGACTIONTCCA TTGACGTCAA TGGGTGGAGT
 ATTTACGGTA
 361 AACTGCCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
 CTATTGACGT
 30 421 CAATGACGGT AAATGGCCCG CCTGGCATTG TGGCCAGTAC ATGACCTTAC
 GGGACTTTCC
 481 TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
 GGTTTTGGCA
 541 GTACACCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC
 35 TCCACCCCAT
 601 TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA
 AATGTCGTAA
 661 CAACTGCGAT CGCCCGCCCC GTTGACGCAA ATGGGCGGTA GGCGTGTACG
 GTGGGAGGTC
 40 721 TATATAAGCA GAGCTCGCTG GCTTGTAACT CAGTCTCTTA CTAGGAGACC
 AGCTTGAGCC
 781 TGGGTGTTCTG CTGGTTAGCC TAACCTGGTT GGCCACCAGG GGTAAGGACT
 CCTTGGCTTA
 841 GAAAGCTAAT AACTTGCCT GCATTAGAGC TTATCTGAGT CAAGTGTCTT
 45 CATTGACGCC
 901 TCACTCTCTT GAACGGGAAT CTTCTTACT GGGTTCTCTC TCTGACCCAG
 GCGAGAGAAA
 961 CTCCAGCAGT GGCGCCCGAA CAGGGACTTG AGTGAGAGTG TAGGCACGTA
 CAGCTGAGAA
 50 1021 GCGTCGGAC GCGAAGGAAG CGCGGGGTGC GACGCGACCA AGAAGGAGAC
 TTGGTGAGTA
 1081 GGCTTCTCGA GTGCCGGGAA AAAGCTCGAG CCTAGTTAGA GGACTIONGAG
 AGGCCGTAGC
 1141 CGTAACTACT CTTGGGCAAG TAGGGCAGGC GGTGGGTACG CAATGGGGGC
 55 GGCTACCTCA
 1201 GCACTAAATA GGAGACAATT AGACCAATTT GAGAAAATAC GACTTCGCCC
 GAACGGAAAG
 1261 AAAAAGTACC AAATTAAACA TTTAATATGG GCAGGCAAGG AGATGGAGCG
 CTTCGGCCTC

	1321	CATGAGAGGT	TGTTGGAGAC	AGAGGAGGGG	TGTAAAAGAA	TCATAGAAGT
		CCTCTACCCC				
	1381	CTAGAACCAA	CAGGATCGGA	GGGCTTAAAA	AGTCTGTTCA	ATCTTGTGTG
		CGTGCTATAT				
5	1441	TGCTTGCACA	AGGAACAGAA	AGTGAAAGAC	ACAGAGGAAG	CAGTAGCAAC
		AGTAAGACAA				
	1501	CACTGCCATC	TAGTGGAAAA	AGAAAAAAGT	GCAACAGAGA	CATCTAGTGG
		ACAAAAGAAA				
10	1561	AATGACAAGG	GAATAGCAGC	GCCACCTGGT	GGCAGTCAGA	ATTTTCCAGC
		GCAACAACAA				
	1621	GGAAATGCCT	GGGTACATGT	ACCCTTGTCA	CCGCGCACCT	TAAATGCGTG
		GGTAAAAGCA				
	1681	GTAGAGGAGA	AAAAATTTGG	AGCAGAAATA	GTACCCATTT	TTTTGTTTCA
		AGCCCTATCG				
15	1741	AATTCCCGTT	TGTGCTAGGG	TTCTTAGGCT	TCTTGGGGGC	TGCTGGAACT
		GCAATGGGAG				
	1801	CAGCGGCGAC	AGCCCTGACG	GTCCAGTCTC	AGCATTTGCT	TGCTGGGATA
		CTGCAGCAGC				
20	1861	AGAAGAATCT	GCTGGCGGCT	GTGGAGGCTC	AACAGCAGAT	GTTGAAGCTG
		ACCATTTGGG				
	1921	GTGTTAAAAA	CCTCAATGCC	CGCGTCACAG	CCCTTGAGAA	GTACCTAGAG
		GATCAGGCAC				
	1981	GACTAAACTC	CTGGGGGTGC	GCATGGAAAC	AAGTATGTCA	TACCACAGTG
		GAGTGGCCCT				
25	2041	GGACAAATCG	GACTCCGGAT	TGGCAAAATA	TGACTTGGTT	GGAGTGGGAA
		AGACAAATAG				
	2101	CTGATTTGGA	AAGCAACATT	ACGAGACAAT	TAGTGAAGGC	TAGAGAACAA
		GAGGAAAAGA				
30	2161	ATCTAGATGC	CTATCAGAAG	TTAACTAGTT	GGTCAGATTT	CTGGTCTTGG
		TTCGATTTCT				
	2221	CAAAATGGCT	TAACATTTTA	AAAATGGGAT	TTTtagtaAT	AGTAGGAATA
		ATAGGGTTAA				
	2281	GATTACTTTA	CACAGTATAT	GGATGTATAG	TGAGGGTTAG	GCAGGGATAT
		GTTCCCTCTAT				
35	2341	CTCCACAGAT	CCATATCCGC	GGCAATTTTA	AAAGAAAGGG	AGGAATAGGG
		GGACAGACTT				
	2401	CAGCAGAGAG	ACTAATTAAT	ATAATAACAA	CACAATTAGA	AATACAACAT
		TTACAAACCA				
40	2461	AAATTCAAAA	AATTTTAAAT	TTTAGAGCCG	CGGAGATCTG	TTACATAACT
		TATGGTAAAT				
	2521	GGCCTGCCTG	GCTGACTGCC	CAATGACCCC	TGCCCAATGA	TGTCAATAAT
		GATGTATGTT				
	2581	CCCATGTAAT	GCCAATAGGG	ACTTTCCATT	GATGTCAATG	GGTGGAGTAT
		TTATGGTAAC				
45	2641	TGCCCACTTG	GCAGTACATC	AAGTGTATCA	TATGCCAAGT	ATGCCCCCTA
		TTGATGTCAA				
	2701	TGATGGTAAA	TGGCCTGCCT	GGCATTATGC	CCAGTACATG	ACCTTATGGG
		ACTTTCTTAC				
50	2761	TTGGCAGTAC	ATCTATGTAT	TAGTCATTGC	TATTACCATG	GGAATTCACT
		AGTGGAGAAG				
	2821	AGCATGCTTG	AGGGCTGAGT	GCCCCTCAGT	GGGCAGAGAG	CACATGGCCC
		ACAGTCCCTG				
	2881	AGAAGTTGGG	GGGAGGGGTG	GGCAATTGAA	CTGGTGCCTA	GAGAAGGTGG
		GGCTTGGGTA				
55	2941	AACTGGGAAA	GTGATGTGGT	GTACTGGCTC	CACCTTTTTT	CCCAGGGTGG
		GGGAGAACCA				
	3001	TATATAAGTG	CAGTAGTCTC	TGTGAACATT	CAAGCTTCTG	CCTTCTCCCT
		CCTGTGAGTT				
60	3061	TGCTAGCCAC	CATGCAGAGA	AGCCCTCTGG	AGAAGGCCTC	TGTGGTGAGC
		AAGCTGTTCT				

	3121	TCAGCTGGAC	CAGGCCCATC	CTGAGGAAGG	GCTACAGGCA	GAGACTGGAG
		CTGTCTGACA				
	3181	TCTACCAGAT	CCCCTCTGTG	GACTCTGCTG	ACAACCTGTC	TGAGAAGCTG
		GAGAGGGAGT				
5	3241	GGGATAGAGA	GCTGGCCAGC	AAGAAGAACC	CCAAGCTGAT	CAATGCCCTG
		AGGAGATGCT				
	3301	TCTTCTGGAG	ATTCATGTTC	TATGGCATCT	TCCTGTACCT	GGGGGAAGTG
		ACCAAGGCTG				
10	3361	TGCAGCCTCT	GCTGCTGGGC	AGAATCATTG	CCAGCTATGA	CCCTGACAAC
		AAGGAGGAGA				
	3421	GGAGCATTGC	CATCTACCTG	GGCATTGGCC	TGTGCCTGCT	GTTCAATTGTG
		AGGACCTGTC				
	3481	TGCTGCACCC	TGCCATCTTT	GGCCTGCACC	ACATTGGCAT	GCAGATGAGG
		ATTGCCATGT				
15	3541	TCAGCCTGAT	CTACAAGAAA	ACCCTGAAGC	TGTCCAGCAG	AGTGCTGGAC
		AAGATCAGCA				
	3601	TTGGCCAGCT	GGTGAGCCTG	CTGAGCAACA	ACCTGAACAA	GTTTGATGAG
		GGCCTGGCCC				
	3661	TGGCCCACTT	TGTGTGGATT	GGCCCTCTGC	AGGTGGCCCT	GCTGATGGGC
20		CTGATTTGGG				
	3721	AGCTGCTGCA	GGCCTCTGCC	TTTTGTGGCC	TGGGCTTCCT	GATTGTGCTG
		GCCCTGTTTC				
	3781	AGGCTGGCCT	GGGCAGGATG	ATGATGAAGT	ACAGGGACCA	GAGGGCAGGC
		AAGATCAGTG				
25	3841	AGAGGCTGGT	GATCACCTCT	GAGATGATTG	AGAACATCCA	GTCTGTGAAG
		GCCTACTGTT				
	3901	GGGAGGAAGC	TATGGAGAAG	ATGATTGAAA	ACCTGAGGCA	GACAGAGCTG
		AAGCTGACCA				
	3961	GGAAGGCTGC	CTATGTGAGA	TACTTCAACA	GCTCTGCCTT	CTTCTTCTCT
30		GGCTTCTTTG				
	4021	TGGTGTTTCT	GTCTGTGCTG	CCCTATGCCC	TGATCAAGGG	GATCATCCTG
		AGAAAGATTT				
	4081	TCACCACCAT	CAGCTTCTGC	ATTGTGCTGA	GGATGGCTGT	GACCAGACAG
		TTCCCCTGGG				
35	4141	CTGTGCAGAC	CTGGTATGAC	AGCCTGGGGG	CCATCAACAA	GATCCAGGAC
		TTCCTGCAGA				
	4201	AGCAGGAGTA	CAAGACCCTG	GAGTACAACC	TGACCACCAC	AGAAGTGGTG
		ATGGAGAATG				
	4261	TGACAGCCTT	CTGGGAGGAG	GGCTTTGGGG	AGCTGTTTGA	GAAGGCCAAG
40		CAGAACAACA				
	4321	ACAACAGAAA	GACCAGCAAT	GGGGATGACT	CCCTGTTCTT	CTCCAACTTC
		TCCCTGCTGG				
	4381	GCACACCTGT	GCTGAAGGAC	ATCAACTTCA	AGATTGAGAG	GGGGCAGCTG
		CTGGCTGTGG				
45	4441	CTGGATCTAC	AGGGGCTGGC	AAGACCAGCC	TGCTGATGAT	GATCATGGGG
		GAGCTGGAGC				
	4501	CTTCTGAGGG	CAAGATCAAG	CACTCTGGCA	GGATCAGCTT	TTGCAGCCAG
		TTCAGCTGGA				
	4561	TCATGCCTGG	CACCATCAAG	GAGAACATCA	TCTTTGGAGT	GAGCTATGAT
50		GAGTACAGAT				
	4621	ACAGGAGTGT	GATCAAGGCC	TGCCAGCTGG	AGGAGGACAT	CAGCAAGTTT
		GCTGAGAAGG				
	4681	ACAACATTGT	GCTGGGGGAG	GGAGGCATTA	CACTGTCTGG	GGGCCAGAGA
		GCCAGAATCA				
55	4741	GCCTGGCCAG	GGCTGTGTAC	AAGGATGCTG	ACCTGTACCT	GCTGGACTCC
		CCCTTTGGCT				
	4801	ACCTGGATGT	GCTGACAGAG	AAGGAGATTT	TTGAGAGCTG	TGTGTGCAAG
		CTGATGGCCA				
	4861	ACAAGACCAG	AATCCTGGTG	ACCAGCAAGA	TGGAGCACCT	GAAGAAGGCT
60		GACAAGATCC				

	4921	TGATCCTGCA	TGAGGGCAGC	AGCTACTTCT	ATGGGACCTT	CTCTGAGCTG
		CAGAACCTGC				
	4981	AGCCTGACTT	CAGCTCTAAG	CTGATGGGCT	GTGACAGCTT	TGACCAGTTC
		TCTGCTGAGA				
5	5041	GGAGGAACAG	CATCCTGACA	GAGACCCTGC	ACAGATTGAG	CCTGGAGGGA
		GATGCCCCTG				
	5101	TGAGCTGGAC	AGAGACCAAG	AAGCAGAGCT	TCAAGCAGAC	AGGGGAGTTT
		GGGGAGAAGA				
	5161	GGAAGAACTC	CATCCTGAAC	CCCATCAACA	GCATCAGGAA	G TTCAGCATT
10		GTGCAGAAAA				
	5221	CCCCCCTGCA	GATGAATGGC	ATTGAGGAAG	ATTCTGATGA	GCCCCCTGGAG
		AGGAGACTGA				
	5281	GCCTGGTGCC	TGATTCTGAG	CAGGGAGAGG	CCATCCTGCC	TAGGATCTCT
		GTGATCAGCA				
15	5341	CAGGCCCTAC	ACTGCAGGCC	AGAAGGAGGC	AGTCTGTGCT	GAACCTGATG
		ACCCACTCTG				
	5401	TGAACCAGGG	CCAGAACATC	CACAGGAAAA	CCACAGCCTC	CACCAGGAAA
		GTGAGCCTGG				
	5461	CCCCTCAGGC	CAATCTGACA	GAGCTGGACA	TCTACAGCAG	GAGGCTGTCT
20		CAGGAGACAG				
	5521	GCCTGGAGAT	TTCTGAGGAG	ATCAATGAGG	AGGACCTGAA	AGAGTGCTTC
		TTTGATGACA				
	5581	TGGAGAGCAT	CCCTGCTGTG	ACCACCTGGA	ACACCTACCT	GAGATACATC
		ACAGTGCACA				
25	5641	AGAGCCTGAT	CTTTGTGCTG	ATCTGGTGCC	TGGTGATCTT	CCTGGCTGAA
		GTGGCTGCCT				
	5701	CTCTGGTGGT	GCTGTGGCTG	CTGGGAAACA	CCCCACTGCA	GGACAAGGGC
		AACAGCACCC				
	5761	ACAGCAGGAA	CAACAGCTAT	GCTGTGATCA	TCACCTCCAC	CTCCAGCTAC
30		TATGTGTTCT				
	5821	ACATCTATGT	GGGAGTGGCT	GATACCCTGC	TGGCTATGGG	CTTCTTTAGA
		GGCCTGCCCC				
	5881	TGGTGCACAC	ACTGATCACA	GTGAGCAAGA	TCCTCCACCA	CAAGATGCTG
		CACTCTGTGC				
35	5941	TGCAGGCTCC	TATGAGCACC	CTGAATACCC	TGAAGGCTGG	GGGCATCCTG
		AACAGATTCT				
	6001	CCAAGGATAT	TGCCATCCTG	GATGACCTGC	TGCCTCTCAC	CATCTTTGAC
		TTCATCCAGC				
	6061	TGCTGCTGAT	TGTGATTGGG	GCCATTGCTG	TGGTGGCAGT	GCTGCAGCCC
40		TACATCTTTG				
	6121	TGGCCACAGT	GCCTGTGATT	GTGGCCTTCA	TCATGCTGAG	GGCCTACTTT
		CTGCAGACCT				
	6181	CCCAGCAGCT	GAAGCAGCTG	GAGTCTGAGG	GCAGAAGCCC	CATCTTCACC
		CACCTGGTGA				
45	6241	CAAGCCTGAA	GGGCCTGTGG	ACCCTGAGAG	CCTTTGGCAG	GCAGCCCTAC
		TTTGAGACCC				
	6301	TGTTCCACAA	GGCCCTGAAC	CTGCACACAG	CCAACTGGTT	CCTCTACCTG
		TCCACCCTGA				
	6361	GATGGTTCCA	GATGAGAATT	GAGATGATCT	TTGTCATCTT	CTTCATTGCT
50		GTGACCTTCA				
	6421	TCAGCATTCT	GACCACAGGA	GAGGGAGAGG	GCAGAGTGGG	CATTATCCTG
		ACCCTGGCCA				
	6481	TGAACATCAT	GAGCACACTG	CAGTGGGCAG	TGAACAGCAG	CATTGATGTG
		GACAGCCTGA				
55	6541	TGAGGAGTGT	GAGCAGAGTG	TTCAAGTTCA	TTGATATGCC	CACAGAGGGC
		AAGCCTACCA				
	6601	AGAGCACCAA	GCCCTACAAG	AATGGCCAGC	TGAGCAAAGT	GATGATCATT
		GAGAACAGCC				
	6661	ATGTGAAGAA	GGATGATATC	TGGCCCAGTG	GAGGCCAGAT	GACAGTGAAG
60		GACCTGACAG				

	6721	CCAAGTACAC	AGAGGGGGGGC	AATGCTATCC	TGGAGAACAT	CTCCTTCAGC
		ATCTCCCCTG				
	6781	GCCAGAGAGT	GGGACTGCTG	GGAAGAACAG	GCTCTGGCAA	GTCTACCCTG
		CTGTCTGCCT				
5	6841	TCCTGAGGCT	GCTGAACACA	GAGGGAGAGA	TCCAGATTGA	TGGAGTGTCC
		TGGGACAGCA				
	6901	TCACACTGCA	GCAGTGGAGG	AAGGCCTTTG	GTGTGATCCC	CCAGAAAGTG
		TTCATCTTCA				
	6961	GTGGCACCTT	CAGGAAGAAC	CTGGACCCCT	ATGAGCAGTG	GTCTGACCAG
10		GAGATTTGGA				
	7021	AAGTGGCTGA	TGAAGTGGGC	CTGAGAAGTG	TGATTGAGCA	GTTCCCTGGC
		AAGCTGGACT				
	7081	TTGTCCTGGT	GGATGGGGGC	TGTGTGCTGA	GCCATGGCCA	CAAGCAGCTG
		ATGTGCCTGG				
15	7141	CCAGATCAGT	GCTGAGCAAG	GCCAAGATCC	TGCTGCTGGA	TGAGCCTTCT
		GCCACCTGG				
	7201	ATCCTGTGAC	CTACCAGATC	ATCAGGAGGA	CCCTCAAGCA	GGCCTTTGCT
		GA CTGCACAG				
	7261	TCATCCTGTG	TGAGCACAGG	ATTGAGGCCA	TGCTGGAGTG	CCAGCAGTTC
20		CTGGTGATTG				
	7321	AGGAGAACAA	AGTGAGGCAG	TATGACAGCA	TCCAGAAGCT	GCTGAATGAG
		AGGAGCCTGT				
	7381	TCAGGCAGGC	CATCAGCCCC	TCTGATAGAG	TGAAGCTGTT	CCCCCACAGG
		AACAGCTCCA				
25	7441	AGTGCAAGAG	CAAGCCCCAG	ATTGCTGCCC	TGAAGGAGGA	GACAGAGGAG
		GAAGTGCAGG				
	7501	ACACCAGGCT	GTGAGGGCCC	AATCAACCTC	TGGATTACAA	AATTTGTGAA
		AGATTGACTG				
	7561	GTATTCTTAA	CTATGTTGCT	CCTTTTACGC	TATGTGGATA	CGCTGCTTTA
30		ATGCCTTTGT				
	7621	ATCATGCTAT	TGCTTCCCGT	ATGGCTTTCA	TTTTCTCCTC	CTTGTATAAA
		TCCTGTTGCT				
	7681	TGTCTCTTTA	TGAGGAGTTG	TGGCCCGTTG	TCAGGCAACG	TGGCGTGGTG
		TGCACTGTGT				
35	7741	TTGCTGACGC	AACCCCCACT	GGTTGGGGCA	TTGCCACCAC	CTGTCAGCTC
		CTTTCCGGGA				
	7801	CTTTCGCTTT	CCCCCTCCCT	ATTGCCACGG	CGGAACTCAT	CGCCGCCTGC
		CTTGCCCGCT				
	7861	GCTGGACAGG	GGCTCGGCTG	TTGGGCACTG	ACAATTCCGT	GGTGTGTGCG
40		GGGAAATCAT				
	7921	CGTCCTTTCC	TTGGCTGCTC	GCCTGTGTTG	CCACCTGGAT	TCTGCGCGGG
		ACGTCCTTCT				
	7981	GCTACGTCCC	TTCGGCCCTC	AATCCAGCGG	ACCTTCCTTC	CCGCGGCCTG
		CTGCCGGCTC				
45	8041	TGCGGCCTCT	TCCGCGTCTT	CGCCTTCGCC	CTCAGACGAG	TCGGATCTCC
		CTTTGGGCCG				
	8101	CCTCCCCGCA	AGCTTCGCAC	TTTTTAAAAG	AAAAGGGAGG	ACTGGATGGG
		ATTTATTA CT				
	8161	CCGATAGGAC	GCTGGCTTGT	AACTCAGTCT	CTTACTAGGA	GACCAGCTTG
50		AGCCTGGGTG				
	8221	TTGCTGGT	AGCCTAACCT	GGTTGGCCAC	CAGGGGTAAG	GA CTCTTG
		CTTAGAAAGC				
	8281	TAATAAACTT	GCCTGCATTA	GAGCTCTTAC	GCGTCCCGGG	CTCGAGATCC
		GCATCTCAAT				
55	8341	TAGTCAGCAA	CCATAGTCCC	GCCCCTAACT	CCGCCCATCC	CGCCCCTAAC
		TCCGCCAGT				
	8401	TCCGCCATT	CTCCGCCCCA	TGGCTGACTA	ATTTTTTTTA	TTTATGCAGA
		GGCCGAGGCC				
	8461	GCCTCGGCCT	CTGAGCTATT	CCAGAAGTAG	TGAGGAGGCT	TTTTTGGAGG
60		CCTAGGCTTT				

	8521	TGCAAAAAGC	TAACCTGTTT	ATTGCAGCTT	ATAATGGTTA	CAAATAAAGC
		AATAGCATCA				
	8581	CAAATTTTAC	AAATAAAGCA	TTTTTTTAC	TGCATTCTAG	TTGTGGTTTG
		TCCAAACTCA				
5	8641	TCAATGTATC	TTATCATGTC	TGTCCGCTTC	CTCGCTCACT	GACTCGCTGC
		GCTCGGTCGT				
	8701	TCGGCTGCGG	CGAGCGGTAT	CAGCTCACTC	AAAGGCGGTA	ATACGGTTAT
		CCACAGAATC				
	8761	AGGGGATAAC	GCAGGAAAGA	ACATGTGAGC	AAAAGGCCAG	CAAAAGGCCA
10		GGAACCGTAA				
	8821	AAAGGCCGCG	TTGCTGGCGT	TTTTCCATAG	GCTCCGCCCC	CCTGACGAGC
		ATCACAAAAA				
	8881	TCGACGCTCA	AGTCAGAGGT	GGCGAAACCC	GACAGGACTA	TAAAGATACC
		AGGCGTTTCC				
15	8941	CCCTGGAAGC	TCCCTCGTGC	GCTCTCCTGT	TCCGACCCTG	CCGCTTACCG
		GATACCTGTC				
	9001	CGCCTTTCTC	CCTTCGGGAA	GCGTGGCGCT	TTCTCATAGC	TCACGCTGTA
		GGTATCTCAG				
	9061	TTCGGTGTAG	GTCGTTCGCT	CCAAGCTGGG	CTGTGTGCAC	GAACCCCCCG
20		TTCAGCCCGA				
	9121	CCGCTGCGCC	TTATCCGGTA	ACTATCGTCT	TGAGTCCAAC	CCGGTAAGAC
		ACGACTTATC				
	9181	GCCACTGGCA	GCAGCCACTG	GTAACAGGAT	TAGCAGAGCG	AGGTATGTAG
		GCGGTGCTAC				
25	9241	AGAGTTCTTG	AAGTGGTGGC	CTAACTACGG	CTACACTAGA	AGAACAGTAT
		TTGGTATCTG				
	9301	CGCTCTGCTG	AAGCCAGTTA	CCTTCGGAAA	AAGAGTTGGT	AGCTCTTGAT
		CCGGCAAACA				
	9361	AACCACCGCT	GGTAGCGGTG	GTTTTTTTGT	TTGCAAGCAG	CAGATTACGC
30		GCAGAAAAAA				
	9421	AGGATCTCAA	GAAGATCCTT	TGATCTTTTC	TACGGGGTCT	GACGCTCAGT
		GGAACGAAAA				
	9481	CTCACGTAA	GGGATTTTGG	TCATGAGATT	ATCAAAAAGG	ATCTTCACCT
		AGATCCTTTT				
35	9541	AAATTAATAA	TGAAGTTTTA	AATCAATCTA	AAGTATATAT	GAGTAAACTT
		GGTCTGACAG				
	9601	TTAGAAAAAC	TCATCGAGCA	TCAAATGAAA	CTGCAATTTA	TTCATATCAG
		GATTATCAAT				
	9661	ACCATATTTT	TGAAAAAGCC	GTTTCTGTAA	TGAAGGAGAA	AACTCACCGA
40		GGCAGTTCCA				
	9721	TAGGATGGCA	AGATCCTGGT	ATCGGTCTGC	GATTCCGACT	CGTCCAACAT
		CAATACAACC				
	9781	TATTAATTTT	CCCTCGTCAA	AAATAAGGTT	ATCAAGTGAG	AAATCACCAT
		GAGTGACGAC				
45	9841	TGAATCCGGT	GAGAATGGCA	ACAGCTTATG	CATTTCTTTC	CAGACTTGTT
		CAACAGGCCA				
	9901	GCCATTACGC	TCGTCATCAA	AATCACTCGC	ATCAACCAAA	CCGTTATTCA
		TTCGTGATTG				
	9961	CGCCTGAGCG	AGACGAAATA	CGCGATCGCT	GTAAAAGGA	CAATTACAAA
50		CAGGAATCGA				
	10021	ATGCAACCGG	CGCAGGAACA	CTGCCAGCGC	ATCAACAATA	TTTTCACCTG
		AATCAGGATA				
	10081	TTCTTCTAAT	ACCTGGAATG	CTGTTTTTCC	GGGGATCGCA	GTGGTGAGTA
		ACCATGCATC				
55	10141	ATCAGGAGTA	CGGATAAAAT	GCTTGATGGT	CGGAAGAGGC	ATAAATTCCG
		TCAGCCAGTT				
	10201	TAGTCTGACC	ATCTCATCTG	TAACATCATT	GGCAACGCTA	CCTTTGCCAT
		GTTTCAGAAA				
	10261	CAACTCTGGC	GCATCGGGCT	TCCCATACAA	TCGATAGATT	GTGCGACCTG
60		ATTGCCCGAC				

```

10321 ATTATCGCGA   GCCCATTTAT   ACCCATATAA   ATCAGCATCC   ATGTTGGAAT
TTAATCGCGG
10381 CCTAGAGCAA   GACGTTTCCC   GTTGAATATG   GCTCATAACA   CCCCTTGTAT
TACTGTTTAT
5  10441 GTAAGCAGAC   AGTTTTATTG   TTCATGATGA   TATATTTTTA   TCTTGTGCAA
TGTAACATCA
10501 GAGATTTTGA   GACACAACAA   TTGGTCGACG   GATCC
SEQ ID NO:2
1  GCTCGAGACT   AGTGACTTGG   TGAGTAGGCT   TCGAGCCTAG   TTAGAGGACT
10 AGGAGAGGCC
61  GTAGCCGTAA   CTA CTCTGGG   CAAGTAGGGC   AGGCGGTGGG   TACGCAATGG
GGGCGGCTAC
121 CTCAGCACTA   AATAGGAGAC   AATTAGACCA   ATTTGAGAAA   ATACGACTTC
GCCCGAACGG
15 181  AAAGAAAAAG   TACCAAATTA   AACATTTAAT   ATGGGCAGGC   AAGGAGATGG
AGCGCTTCGG
241 CCTCCATGAG   AGGTTGTTGG   AGACAGAGGA   GGGGTGTAAA   AGAATCATAG
AAGTCCTCTA
301 CCCCTAGAA   CCAACAGGAT   CGGAGGGGCTT   AAAAAGTCTG   TTCAATCTTG
20 TGTGCGTACT
361 ATATTGCTTG   CACAAGGAAC   AGAAAGTGAA   AGACACAGAG   GAAGCAGTAG
CAACAGTAAG
421 ACAACACTGC   CATCTAGTGG   AAAAAGAAAA   AAGTGCAACA   GAGACATCTA
GTGGACAAAA
25 481  GAAAAATGAC   AAGGGAATAG   CAGCGCCACC   TGGTGGCAGT   CAGAATTTTC
CAGCGCAACA
541 ACAAGGAAAT   GCCTGGGTAC   ATGTACCCTT   GTCACCGCGC   ACCTTAAATG
CGTGGGTAAA
601 AGCAGTAGAG   GAGAAAAAAT   TTGGAGCAGA   AATAGTACCC   ATGTTTCAAG
30 CCCTATCAGA
661 AGGCTGCACA   CCCTATGACA   TTAATCAGAT   GCTTAATGTG   CTAGGAGATC
ATCAAGGGGG
721 ATTACAAATA   GTGAAAGAGA   TCATTAATGA   AGAAGCAGCC   CAGTGGGATG
TAACACACCC
35 781  ACTACCCGCA   GGACCCCTAC   CAGCAGGACA   GCTCAGGGAC   CCTCGCGGCT
CAGATATAGC
841 AGGGACCACC   AGCTCAGTAC   AAGAACAGTT   AGAATGGATC   TATACTGCTA
ACCCCGGGT
901 AGATGTAGGT   GCCATCTACC   GGAGATGGAT   TATTCTAGGA   CTTCAAAAGT
40 GTGTCAAAAT
961 GTACAACCCA   GTATCAGTCC   TAGACATTAG   GCAGGGACCT   AAAGAGCCCT
TCAAGGATTA
1021 TGTGGACAGA   TTTTACAAGG   CAATTAGAGC   AGAACAAGCC   TCAGGGGAAG
TGAAACAATG
45 1081 GATGACAGAA   TCATTACTCA   TTCAAATGTC   TAATCCAGAT   TGTAAGGTCA
TCCTGAAGGG
1141 CCTAGGAATG   CACCCACACC   TTGAAGAAAT   GTTAACGGCT   TGTCAGGGGG
TAGGAGGCCC
1201 AAGCTACAAA   GCAAAAGTAA   TGGCAGAAAT   GATGCAGACC   ATGCAAAATC
50 AAAACATGGT
1261 GCAGCAGGGA   GGTCCAAAAA   GACAAAGACC   CCCACTAAGA   TGTTATAATT
GTGGAAAATT
1321 TGGCCATATG   CAAAGACAAT   GTCCGGAACC   AAGGAAAACA   AAATGTCTAA
AGTGTGGAAT
55 1381 ATTGGGACAC   CTAGCAAAAG   ACTGCAGGGG   ACAGGTGAAT   TTTTtagGGT
ATGGACGGTG
1441 GATGGGGGCA   AAACCGAGAA   ATTTTCCCGC   CGCTACTCTT   GGAGCGGAAC
CGAGTGCGCC
1501 TCCTCCACCG   AGCGGCACCA   CCCCATACGA   CCCAGCAAAG   AAGCTCCTGC
60 AGCAATATGC

```

	1561	AGAGAAAGGG	AAACAACCTGA	GGGAGCAAAA	GAGGAATCCA	CCGGCAATGA
		ATCCGGATTG				
	1621	GACCGAGGGA	TATTCTTTGA	ACTCCCTCTT	TGGAGAAGAC	CAATAAAGAC
		AGTGTATATA				
5	1681	GAAGGGGTCC	CCATTAAGGC	ACTGCTAGAC	ACAGGGGCAG	ATGACACCAT
		AATTAAAGAA				
	1741	AATGATTTAC	AATTATCAGG	TCCATGGAGA	CCCAAAATTA	TAGGGGGCAT
		AGGAGGAGGC				
	1801	CTTAATGTAA	AAGAATATAA	CGACAGGGAA	GTAAAAATAG	AAGATAAAAT
10		TTTGAGAGGA				
	1861	ACAATATTGT	TAGGAGCAAC	TCCCATTAAT	ATAATAGGTA	GAAATTTGCT
		GGCCCCGGA				
	1921	GGTGCCCGGT	TAGTAATGGG	ACAATTATCA	GAAAAAATTC	CTGTCACACC
		TGTCAAATTG				
15	1981	AAGGAAGGGG	CTCGGGGACC	CTGTGTAAGA	CAATGGCCTC	TCTCTAAAGA
		GAAGATTGAA				
	2041	GCTTTACAGG	AAATATGTTC	CCAATTAGAG	CAGGAAGGAA	AAATCAGTAG
		AGTAGGAGGA				
	2101	GAAAATGCAT	ACAATACCCC	AATATTTTGC	ATAAGAAGA	AGGACAAATC
20		CCAGTGAGG				
	2161	ATGCTAGTAG	ACTTTAGAGA	GTAAATAAG	GCAACCCAAG	ATTTCTTTGA
		AGTGCAATTA				
	2221	GGGATACCCC	ACCCAGCAGG	ATTAAGAAAG	ATGAGACAGA	TAACAGTTTT
		AGATGTAGGA				
25	2281	GACGCCTATT	ATTCCATACC	ATTGGATCCA	AATTTTAGGA	AATATACTGC
		TTTTACTATT				
	2341	CCCACAGTGA	ATAATCAGGG	ACCCGGGATT	AGGTATCAAT	TCAACTGTCT
		CCCACAAGGG				
	2401	TGGAAAGGAT	CTCCTACAAT	CTTCCAAAAT	ACAGCAGCAT	CCATTTTGGA
30		GGAGATAAAA				
	2461	AGAACTTGC	CAGCACTAAC	CATTGTACAA	TACATGGATG	ATTTATGGGT
		AGGTTCTCAA				
	2521	GAAAATGAAC	ACACCCATGA	CAAATTAGTA	GAACAGTTAA	GAACAAAATT
		ACAAGCCTGG				
35	2581	GGCTTAGAAA	CCCCAGAAAA	GAAGGTGCAA	AAAGAACCAC	CTTATGAGTG
		GATGGGATAC				
	2641	AACTTTGGC	CTCACAAATG	GGAACATAAGC	AGAATACAAC	TGGAGGAAAA
		AGATGAATGG				
	2701	ACTGTCAATG	ACATCCAGAA	GTTAGTTGGG	AAACTAAATT	GGGCAGCACA
40		ATTGTATCCA				
	2761	GGTCTTAGGA	CCAAGAATAT	ATGCAAGTTA	ATTAGAGGAA	AGAAAAATCT
		GTTAGAGCTA				
	2821	GTGACTTGGA	CACCTGAGGC	AGAAGCTGAA	TATGCAGAAA	ATGCAGAGAT
		TCTTAAAACA				
45	2881	GAACAGGAAG	GAACCTATTA	CAAACCAGGA	ATACCTATTA	GGGCAGCAGT
		ACAGAAATTG				
	2941	GAAGGAGGAC	AGTGGAGTTA	CCAATTCAAA	CAAGAAGGAC	AAGTCTTGAA
		AGTAGGAAAA				
	3001	TACACCAAGC	AAAAGAACAC	CCATACAAAT	GAACCTCGCA	CATTAGCTGG
50		TTTAGTGCAG				
	3061	AAGATTTGCA	AAGAAGCTCT	AGTTATTTGG	GGGATATTAC	CAGTTCTAGA
		ACTCCCGATA				
	3121	GAAAGAGAGG	TATGGGAACA	ATGGTGGGCG	GATTACTGGC	AGGTAAGCTG
		GATTCCCGAA				
55	3181	TGGGATTTTG	TCAGCACCCC	ACCTTTGCTC	AAACTATGGT	ACACATTAAC
		AAAAGAACCC				
	3241	ATACCCAAGG	AGGACGTTTA	CTATGTAGAT	GGAGCATGCA	ACAGAAATTC
		AAAAGAAGGA				
	3301	AAAGCAGGAT	ACATCTCACA	ATACGGAAAA	CAGAGAGTAG	AAACATTAGA
60		AAACACTACC				

	3361	AATCAGCAAG	CAGAATTAAC	AGCTATAAAA	ATGGCTTTGG	AAGACAGTGG
		GCCTAATGTG				
	3421	AACATAGTAA	CAGACTCTCA	ATATGCAATG	GGAATTTTGA	CAGCACAACC
		CACACAAAGT				
5	3481	GATTCACCAT	TAGTAGAGCA	AATTATAGCC	TTAATGATAC	AAAAGCAACA
		AATATATTTG				
	3541	CAGTGGGTAC	CAGCACATAA	AGGAATAGGA	GGAAATGAGG	AGATAGATAA
		ATTAGTGAGT				
10	3601	AAAGGCATTA	GAAGAGTTTT	ATTCTTAGAA	AAAATAGAAG	AAGCTCAAGA
		AGAGCATGAA				
	3661	AGATATCATA	ATAATTGGAA	AAACCTAGCA	GATACATATG	GGCTTCCACA
		AATAGTAGCA				
	3721	AAAGAGATAG	TGGCCATGTG	TCCAAAATGT	CAGATAAAGG	GAGAACCAGT
		GCATGGACAA				
15	3781	GTGGATGCCT	CACCTGGAAC	ATGGCAGATG	GATTGTACTC	ATCTAGAAGG
		AAAAGTAGTC				
	3841	ATAGTTGCGG	TCCATGTAGC	CAGTGGATTG	ATAGAAGCAG	AAGTCATACC
		TAGGGAAACA				
20	3901	GGAAAAGAAA	CGGCAAAGTT	TCTATTAATA	ATACTGAGTA	GATGGCCTAT
		AACACAGTTA				
	3961	CACACAGACA	ATGGGCCTAA	CTTTACCTCC	CAAGAAGTGG	CAGCAATATG
		TTGGTGGGGA				
	4021	AAAATTGAAC	ATACAACAGG	TATACCATAT	AACCCCCAAT	CTCAAGGATC
		AATAGAAAAGC				
25	4081	ATGAACAAAC	AATTAAGA	GATAATTGGG	AAAATAAGAG	ATGATTGCCA
		ATATACAGAG				
	4141	ACAGCAGTAC	TGATGGCTTG	CCATATTCAC	AATTTTAAAA	GAAAGGGAGG
		AATAGGGGGA				
30	4201	CAGACTTCAG	CAGAGAGACT	AATTAATATA	ATAACAACAC	AATTAGAAAT
		ACAACATTTA				
	4261	CAAACCAAAA	TTCAAAAAAT	TTTAAATTTT	AGAGTCTACT	ACAGAGAAGG
		GAGAGACCCT				
	4321	GTGTGGAAAG	GACCAGCACA	ATTAATCTGG	AAAGGGGAAG	GAGCAGTGGT
		CCTCAAGGAC				
35	4381	GGAAGTGACC	TAAAGGTTGT	ACCAAGAAGG	AAAGCTAAAA	TTATTAAGGA
		TTATGAACCC				
	4441	AAACAAAGAG	TGGGTAATGA	GGGTGACGTG	GAAGGTACCA	GGGGATCTGA
		TAATAAATG				
40	4501	GCAGGGAATA	GTCAGATATT	GGATGAGACA	AAGAAATTTG	AAATGGAACT
		ATTATATGCA				
	4561	TCAGCTGGCG	GCCGCGAATT	CACTAGTGAT	TCCCGTTTGT	GCTAGGGTTC
		TTAGGCTTCT				
	4621	TGGGGGCTGC	TGGAAGTACA	ATGGGAGCAG	CGGCGACAGC	CCTGACGGTC
		CAGTCTCAGC				
45	4681	ATTTGCTTGC	TGGGATACTG	CAGCAGCAGA	AGAATCTGCT	GGCGGCTGTG
		GAGGCTCAAC				
	4741	AGCAGATGTT	GAAGCTGACC	ATTTGGGGTG	TTAAAAACCT	CAATGCCCGC
		GTCACAGCCC				
50	4801	TTGAGAAGTA	CCTAGAGGAT	CAGGCACGAC	TAAACTCCTG	GGGGTGCGCA
		TGGAACAAG				
	4861	TATGTCATAC	CACAGTGGAG	TGGCCCTGGA	CAAATCGGAC	TCCGGATTGG
		CAAAATATGA				
	4921	CTTGTTGGA	GTGGGAAAGA	CAAATAGCTG	ATTTGGAAAG	CAACATTACG
		AGACAATTAG				
55	4981	TGAAGGCTAG	AGAACAAGAG	GAAAAGAATC	TAGATGCCTA	TCAGAAGTTA
		ACTAGTTGGT				
	5041	CAGATTCTG	GTCTTGTTTC	GATTTCTCAA	AATGGCTTAA	CATTTTAAAA
		ATGGGATTTT				
60	5101	TAGTAATAGT	AGGAATAATA	GGGTTAAGAT	TACTTTACAC	AGTATATGGA
		TGTATAGTGA				

	5161	GGGTTAGGCA	GGGATATGTT	CCTCTATCTC	CACAGATCCA	TATCCAATCG
		AATTCCCGCG				
	5221	GCCGCAATTC	ACTCCTCAGG	TGCAGGCTGC	CTATCAGAAG	GTGGTGGCTG
		GTGTGGCCAA				
5	5281	TGCCCTGGCT	CACAAATACC	ACTGAGATCT	TTTTCCCTCT	GCCAAAAATT
		ATGGGGACAT				
	5341	CATGAAGCCC	CTTGAGCATC	TGACTTCTGG	CTAATAAAGG	AAATTTATTT
		TCATTGCAAT				
10	5401	AGTGTGTTGG	AATTTTTTGT	GTCTCTCACT	CGGAAGGACA	TATGGGAGGG
		CAAATCATTT				
	5461	AAAACATCAG	AATGAGTATT	TGGTTTAGAG	TTTGGCAACA	TATGCCCATA
		TGCTGGCTGC				
	5521	CATGAACAAA	GGTTGGCTAT	AAAGAGGTCA	TCAGTATATG	AAACAGCCCC
		CTGCTGTCCA				
15	5581	TTCCTTATTC	CATAGAAAAG	CCTTGACTTG	AGGTTAGATT	TTTTTTATAT
		TTTGTTTTGT				
	5641	GTTATTTTTT	TCTTTAACAT	CCCTAAAATT	TTCCTTACAT	GTTTTACTAG
		CCAGATTTTT				
	5701	CCTCCTCTCC	TGACTACTCC	CAGTCATAGC	TGTCCCTCTT	CTCTTATGGA
		GATCCCTCGA				
20	5761	CCTGCAGCCC	AAGCTTGCGC	TAATCATGGT	CATAGCTGTT	TCCTGTGTGA
		AATTGTTATC				
	5821	CGCTCACAAT	TCCACACAAC	ATACGAGCCG	GAAGCATAAA	GTGTAAAGCC
		TGGGGTGCCT				
	5881	AATGAGTGAG	CTAACTCACA	TTAATTGCGT	TGCGCTCACT	GCCCGCTTTC
25		CAGTCGGGAA				
	5941	ACCTGTCGTG	CCAGCGGATC	CGCATCTCAA	TTAGTCAGCA	ACCATAGTCC
		CGCCCCTAAC				
	6001	TCCGCCCATC	CCGCCCTTAA	CTCCGCCCAG	TTCCGCCCAT	TCTCCGCCCC
		ATGGCTGACT				
30	6061	AATTTTTTTT	ATTTATGCAG	AGGCCGAGGC	CGCCTCGGCC	TCTGAGCTAT
		TCCAGAAGTA				
	6121	GTGAGGAGGC	TTTTTTGGAG	GCCTAGGCTT	TTGCAAAAAG	CTAACTTGTT
		TATTGCAGCT				
35	6181	TATAATGGTT	ACAAATAAAG	CAATAGCATC	ACAAATTTCA	CAAATAAAGC
		ATTTTTTTCA				
	6241	CTGCATTCTA	GTTGTGGTTT	GTCCAAACTC	ATCAATGTAT	CTTATCATGT
		CTGTCCGCTT				
	6301	CCTCGCTCAC	TGACTCGCTG	CGCTCGGTCTG	TTCGGCTGCG	GCGAGCGGTA
		TCAGCTCACT				
40	6361	CAAAGGCGGT	AATACGGTTA	TCCACAGAAT	CAGGGGATAA	CGCAGGAAAG
		AACATGTGAG				
	6421	CAAAAGGCCA	GCAAAAGGCC	AGGAACCGTA	AAAAGGCCGC	GTTGCTGGCG
		TTTTTCCATA				
	6481	GGCTCCGCCC	CCCTGACGAG	CATCACAAAA	ATCGACGCTC	AAGTCAGAGG
45		TGGCGAAACC				
	6541	CGACAGGACT	ATAAAGATAC	CAGGCGTTTC	CCCCTGGAAG	CTCCCTCGTG
		CGCTCTCCTG				
	6601	TTCCGACCCT	GCCGCTTACC	GGATACCTGT	CCGCCTTTCT	CCCTTCGGGA
		AGCGTGCGC				
50	6661	TTTCTCATAG	CTCACGCTGT	AGGTATCTCA	GTTGCGTGTA	GGTCGTTTCG
		TCCAAGCTGG				
	6721	GCTGTGTGCA	CGAACCCCCC	GTTGAGCCCG	ACCGCTGCGC	CTTATCCGGT
		AACTATCGTC				
55	6781	TTGAGTCCAA	CCCGGTAAGA	CACGACTTAT	CGCCACTGGC	AGCAGCCACT
		GGTAACAGGA				
	6841	TTAGCAGAGC	GAGGTATGTA	GGCGGTGCTA	CAGAGTTCTT	GAAGTGGTGG
		CCTAACTACG				
	6901	GCTACACTAG	AAGAACAGTA	TTTGGTATCT	GCGCTCTGCT	GAAGCCAGTT
		ACCTTCGGAA				

	6961	AAAGAGTTGG	TAGCTCTTGA	TCCGGCAAAC	AAACCACCGC	TGGTAGCGGT
		GGTTTTTTTG				
	7021	TTTGCAAGCA	GCAGATTACG	CGCAGAAAAA	AAGGATCTCA	AGAAGATCCT
		TTGATCTTTT				
5	7081	CTACGGGGTC	TGACGCTCAG	TGGAACGAAA	ACTCACGTTA	AGGGATTTTG
		GTCATGAGAT				
	7141	TATCAAAAAG	GATCTTCACC	TAGATCCTTT	TAAATTAAAA	ATGAAGTTTT
		AAATCAATCT				
	7201	AAAGTATATA	TGAGTAAACT	TGGTCTGACA	GTTAGAAAAA	CTCATCGAGC
10		ATCAAATGAA				
	7261	ACTGCAATTT	ATTCATATCA	GGATTATCAA	TACCATATTT	TTGAAAAAGC
		CGTTTCTGTA				
	7321	ATGAAGGAGA	AAACTCACCG	AGGCAGTTCC	ATAGGATGGC	AAGATCCTGG
		TATCGGTCTG				
15	7381	CGATTCCGAC	TCGTCCAACA	TCAATACAAC	CTATTAATTT	CCCCTCGTCA
		AAAATAAGGT				
	7441	TATCAAGTGA	GAAATCACCA	TGAGTGACGA	CTGAATCCGG	TGAGAATGGC
		AACAGCTTAT				
	7501	GCATTTCTTT	CCAGACTTGT	TCAACAGGCC	AGCCATTACG	CTCGTCATCA
20		AAATCACTCG				
	7561	CATCAACCAA	ACCGTTATTC	ATTCGTGATT	GCGCCTGAGC	GAGACGAAAT
		ACGCGATCGC				
	7621	TGTTAAAAGG	ACAATTACAA	ACAGGAATCG	AATGCAACCG	GCGCAGGAAC
		ACTGCCAGCG				
25	7681	CATCAACAAT	ATTTTCACCT	GAATCAGGAT	ATTCTTCTAA	TACCTGGAAT
		GCTGTTTTTC				
	7741	CGGGGATCGC	AGTGGTGAGT	AACCATGCAT	CATCAGGAGT	ACGGATAAAA
		TGCTTGATGG				
	7801	TCGGAAGAGG	CATAAATTCC	GTCAGCCAGT	TTAGTCTGAC	CATCTCATCT
30		GTAACATCAT				
	7861	TGGCAACGCT	ACCTTTGCCA	TGTTTCAGAA	ACAACTCTGG	CGCATCGGGC
		TTCCCATACA				
	7921	ATCGATAGAT	TGTCGCACCT	GATTGCCCGA	CATTATCGCG	AGCCCATTTA
		TACCCATATA				
35	7981	AATCAGCATC	CATGTTGGAA	TTTAATCGCG	GCCTAGAGCA	AGACGTTTCC
		CGTTGAATAT				
	8041	GGCTCATAAC	ACCCCTTGTA	TTACTGTTTA	TGTAAGCAGA	CAGTTTTTATT
		GTTCATGATG				
	8101	ATATATTTTT	ATCTTG TGCA	ATGTAACATC	AGAGATTTTG	AGACACAACA
40		ATTGTCGACA				
	8161	TTGATTATTG	ACTAGTTATT	AATAGTAATC	AATTACGGGG	TCATTAGTTC
		ATAGCCCATA				
	8221	TATGGAGTTC	CGCGTTACAT	AACTTACGGT	AAATGGCCCCG	CCTGGCTGAC
		CGCCCAACGA				
45	8281	CCCCCGCCCA	TTGACGTCAA	TAATGACGTA	TGTTCCCATATA	GTAACGCCAA
		TAGGGACTTT				
	8341	CCATTGACGT	CAATGGGTGG	AGTATTTACG	GTAAACTGCC	CACTTGGCAG
		TACATCAAGT				
	8401	GTATCATATG	CCAAGTACGC	CCCCTATTGA	CGTCAATGAC	GGTAAATGGC
50		CCGCCTGGCA				
	8461	TTATGCCCAG	TACATGACCT	TATGGGACTT	TCCTACTTGG	CAGTACATCT
		ACGTATTAGT				
	8521	CATCGCTATT	ACCATGGTCG	AGGTGAGCCC	CACGTTCTGC	TTCACTCTCC
		CCATCTCCCC				
55	8581	CCCCTCCCCA	CCCCCAATTT	TGTATTTATT	TATTTTTTAA	TTATTTTGTG
		CAGCGATGGG				
	8641	GGCGGGGGGG	GGGGGGGGGC	GCGCGCCAGG	CGGGGCGGGG	CGGGGCGAGG
		GGCGGGGGCG				
	8701	GGCGAGGCGG	AGAGGTGCGG	CGGCAGCCAA	TCAGAGCGGC	GCGCTCCGAA
60		AGTTTCCTTT				

```

      8761 TATGGCGAGG CGGCGGCGGC GCGGCCCTA TAAAAAGCGA AGCGCGCGGC
GGGCGGGAGT
      8821 CGCTGCGCGC TGCCTTCGCC CCGTGCCCCG CTCCGCCGCC GCCTCGCGCC
GCCCGCCCCG
5      8881 GCTCTGACTG ACCGCGTTAC TCCCACAGGT GAGCGGGCGG GACGGCCCTT
CTCCTCCGGG
      8941 CTGTAATTAG CGCTTGTTT AATGACGGCT TGTTCCTTT CTGTGGCTGC
GTGAAAGCCT
      9001 TGAGGGGCTC CGGGAGGGCC CTTTGTGCGG GGGGAGCGGC TCGGGGGGTG
10     CGTGCGTGTG
      9061 TGTGTGCGTG GGGAGCGCCG CGTGCGGCTC CGCGCTGCCC GCGGGCTGTG
AGCGCTGGCG
      9121 GCGCGGCGCG GGGCTTTGTG CGCTCCGCAG TGTGCGCGAG GGGAGCGCGG
CCGGGGGCGG
15     9181 TGCCCCGCGG TGCGGGGGGG GCTGCGAGGG GAACAAAGGC TCGTGCGGG
GTGTGTGCGT
      9241 GGGGGGGTGA GCAGGGGGTG TGGGCGCGTC GGTCGGGCTG CAACCCCCC
TGCACCCCC
      9301 TCCCCGAGTT GCTGAGCACG GCCCGGCTTC GGTGCGGGG CTCCGTACGG
20     GCGGTGGCGC
      9361 GGGGCTCGCC GTGCCGGGCG GGGGTGGCG GCAGGTGGGG GTGCCGGGCG
GGGCGGGGCC
      9421 GCCTCGGGCC GGGGAGGGCT CGGGGAGGG GCGCGGCGGC CCCCAGAGCG
CCGGCGGCTG
25     9481 TCGAGGCGCG GCGAGCCGCA GCCATTGCCT TTTATGGTAA TCGTGCGAGA
GGGCGCAGGG
      9541 ACTTCCTTTG TCCCAAATCT GTGCGGAGCC GAAATCTGGG AGGCGCCGCC
GCACCCCCTC
      9601 TAGCGGGCGC GGGGCGAAGC GGTGCGGCGC CGGCAGGAAG GAAATGGGCG
30     GGGAGGGCCT
      9661 TCGTGCGTCG CCGCGCCGCC GTCCCCTTCT CCCTCTCCAG CCTCGGGGCT
GTCCGCGGGG
      9721 GGACGGCTGC CTTGCGGGGG GACGGGGCAG GCGGGGGTTC GGCTTCTGGC
GTGTGACCGG
35     9781 CGGCTCTAGA GCCTCTGCTA ACCATGTTCA TGCCTTCTTC TTTTCCTAC
AGCTCCTGGG
      9841 CAACGTGCTG GTTATTGTGC TGTCTCATCA TTTTGGCAA GAATT
SEQ ID NO:3
      1 TCAATATTGG CCATTAGCCA TATTATTCAT TGGTTATATA GCATAAATCA ATATTGGCTA
40     61 TTGGCCATTG CATACTTGT ATCTATATCA TAATATGTAC ATTTATATTG GCTCATGTCC
121 AATATGACCG CCATGTTGGC ATTGATTATT GACTAGTTAT TAATAGTAAT
CAATTACGGG
      181 GTCATTAGTT CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTACGG
TAAATGGCCC
45     241 GCCTGGCTGA CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT
ATGTTCCCAT
      301 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
GGTAAACTGC
      361 CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTCCG CCCCTATTG
50     ACGTCAATGA
      421 CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTACGGGACT
TTCCTACTTG
      481 GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT
GGCAGTACAC
55     541 CAATGGGCGT GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC
CCATTGACGT
      601 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC
GTAATAACCC
      661 CGCCCCGTTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA
60     TAAGCAGAGC

```

	721	TCGTTTAGTG	AACCGTCAGA	TCACTAGAAG	CTTTATTGCG	GTAGTTTATC
		ACAGTTAAAT				
	781	TGCTAACGCA	GTCAGTGCTT	CTGACACAAC	AGTCTCGAAC	TTAAGCTGCA
		GAAGTTGGTC				
5	841	GTGAGGCACT	GGGCAGGTAA	GTATCAAGGT	TACAAGACAG	GTTTAAGGAG
		ACCAATAGAA				
	901	ACTGGGCTTG	TCGAGACAGA	GAAGACTCTT	GCGTTTCTGA	TAGGCACCTA
		TTGGTCTTAC				
	961	TGACATCCAC	TTTGCCTTTC	TCTCCACAGG	TGTCCACTCC	CAGTTCAATT
10		ACAGCTCTTA				
	1021	AGGCTAGAGT	ACTTAATACG	ACTCACTATA	GGCTAGCCTC	GAGAATTCTGA
		TTATGCCCTT				
	1081	AGGACCAGAA	GAAAGAAGAT	TGCTTCGCTT	GATTTGGCTC	CTTTACAGCA
		CCAATCCATA				
15	1141	TCCACCAAGT	GGGGAAGGGA	CGGCCAGACA	ACGCCGACGA	GCCAGGAGAA
		GGTGGAGACA				
	1201	ACAGCAGGAT	CAAATTAGAG	TCTTGGTAGA	AAGACTCCAA	GAGCAGGTGT
		ATGCAGTTGA				
	1261	CCGCCTGGCT	GACGAGGCTC	AACACTTGGC	TATACAACAG	TTGCCTGACC
20		CTCCTCATTC				
	1321	AGCTTAGAAT	CACTAGTGAA	TTCACGCGTG	GTACCTCTAG	AGTCGACCCG
		GGCGGCCGCT				
	1381	TCGAGCAGAC	ATGATAAGAT	ACATTGATGA	GTTTGGACAA	ACCACAACCTA
		GAATGCAGTG				
25	1441	AAAAAAATGC	TTTATTTGTG	AAATTTGTGA	TGCTATTGCT	TTATTTGTAA
		CCATTATAAG				
	1501	CTGCAATAAA	CAAGTTAACA	ACAACAATTG	CATTCATTTT	ATGTTTCAGG
		TTCAGGGGGA				
	1561	GATGTGGGAG	GTTTTTTAAA	GCAAGTAAAA	CCTCTACAAA	TGTGGTAAAA
30		TCGATAAGGA				
	1621	TCCGTCGACC	AATTGTTGTG	TCTCAAAATC	TCTGATGTTA	CATTGCACAA
		GATAAAAATA				
	1681	TATCATCATG	AACAATAAAA	CTGTCTGCTT	ACATAAACAG	TAATACAAGG
		GGTGTTATGA				
35	1741	GCCATATTCA	ACGGGAAACG	TCTTGCTCTA	GGCCGCGATT	AAATTCCAAC
		ATGGATGCTG				
	1801	ATTTATATGG	GTATAAATGG	GCTCGCGATA	ATGTCGGGCA	ATCAGGTGCG
		ACAATCTATC				
	1861	GATTGTATGG	GAAGCCCGAT	GCGCCAGAGT	TGTTTCTGAA	ACATGGCAAA
40		GGTAGCGTTG				
	1921	CCAATGATGT	TACAGATGAG	ATGGTCAGAC	TAAACTGGCT	GACGGAATTT
		ATGCCTCTTC				
	1981	CGACCATCAA	GCATTTTATC	CGTACTCCTG	ATGATGCATG	GTTACTCACC
		ACTGCGATCC				
45	2041	CCGGA AAAAC	AGCATTCCAG	GTATTAGAAG	AATATCCTGA	TTCAGGTGAA
		AATATTGTTG				
	2101	ATGCGCTGGC	AGTGTTCCCTG	CGCCGGTTGC	ATTCGATTCC	TGTTTGTAAT
		TGTCCTTTTA				
	2161	ACAGCGATCG	CGTATTTTCG	CTCGCTCAGG	CGCAATCACG	AATGAATAAC
50		GGTTTGGTTG				
	2221	ATGCGAGTGA	TTTTGATGAC	GAGCGTAATG	GCTGGCCTGT	TGAACAAGTC
		TGGAAGAAA				
	2281	TGCATAAGCT	GTTGCCATTC	TCACCGGATT	CAGTCGTCAC	TCATGGTGAT
		TTCTCACTTG				
55	2341	ATAACCTTAT	TTTTGACGAG	GGGAAATTAA	TAGGTTGTAT	TGATGTTGGA
		CGAGTCGGAA				
	2401	TCGCAGACCG	ATACCAGGAT	CTTGCCATCC	TATGGAACCTG	CCTCGGTGAG
		TTTTCTCCTT				
	2461	CATTACAGAA	ACGGCTTTTT	CAAAAATATG	GTATTGATAA	TCCTGATATG
60		AATAAATTGC				

	2521	AGTTTCATTT	GATGCTCGAT	GAGTTTTTCT	AACTGTCAGA	CCAAGTTTAC
		TCATATATAC				
	2581	TTTAGATTGA	TTTAAAACTT	CATTTTTAAT	TTAAAAGGAT	CTAGGTGAAG
		ATCCTTTTTG				
5	2641	ATAATCTCAT	GACCAAAATC	CCTTAACGTG	AGTTTTTCGT	CCACTGAGCG
		TCAGACCCCG				
	2701	TAGAAAAGAT	CAAAGGATCT	TCTTGAGATC	CTTTTTTCT	GCGCGTAATC
		TGCTGCTTGC				
10	2761	AAACAAAAAA	ACCACCGCTA	CCAGCGGTGG	TTTGTTTGCC	GGATCAAGAG
		CTACCAACTC				
	2821	TTTTTCCGAA	GGTAACTGGC	TTCAGCAGAG	CGCAGATACC	AAATACTGTT
		CTTCTAGTGT				
	2881	AGCCGTAGTT	AGGCCACCAC	TTCAAGAACT	CTGTAGCACC	GCCTACATAC
		CTCGCTCTGC				
15	2941	TAATCCTGTT	ACCAGTGGCT	GCTGCCAGTG	GCGATAAGTC	GTGTCTTACC
		GGGTTGGA				
	3001	CAAGACGATA	GTTACCGGAT	AAGGCGCAGC	GGTCGGGCTG	AACGGGGGGT
		TCGTGCACAC				
20	3061	AGCCAGCTT	GGAGCGAACG	ACCTACACCG	AACTGAGATA	CCTACAGCGT
		GAGCTATGAG				
	3121	AAAGCGCCAC	GCTTCCCGAA	GGGAGAAAGG	CGGACAGGTA	TCCGGTAAGC
		GGCAGGGTCG				
	3181	GAACAGGAGA	GCGCACGAGG	GAGCTTCCAG	GGGGAACGC	CTGGTATCTT
		TATAGTCCTG				
25	3241	TCGGGTTTCG	CCACCTCTGA	CTTGAGCGTC	GATTTTTGTG	ATGCTCGTCA
		GGGGGGCGGA				
	3301	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTTACGGTT	CCTGGCCTTT
		TGCTGGCCTT				
30	3361	TTGCTCACAT	GGCTCGACAG	ATCT		
		SEQ ID NO:4				
	1	ATTGATTATT	GACTAGTTAT	TAATAGTAAT	CAATTACGGG	GTCATTAGTT
		CATAGCCCAT				
	61	ATATGGAGTT	CCGCGTTACA	TAACTTACGG	TAAATGGCCC	GCCTGGCTGA
		CCGCCAACG				
35	121	ACCCCCGCCC	ATTGACGTCA	ATAATGACGT	ATGTTCCCAT	AGTAACGCCA
		ATAGGACTT				
	181	TCCATTGACG	TCAATGGGTG	GAGTATTTAC	GGTAAACTGC	CCACTTGGCA
		GTACATCAAG				
	241	TGTATCATAT	GCCAAGTACG	CCCCCTATTG	ACGTCAATGA	CGGTAAATGG
		CCCGCCTGGC				
40	301	ATTATGCCCA	GTACATGACC	TTATGGGACT	TTCTACTTGG	GCAGTACATC
		TACGTATTAG				
	361	TCATCGCTAT	TACCATGGTC	GAGGTGAGCC	CCACGTTCTG	CTTCACTCTC
		CCCATCTCCC				
45	421	CCCCCTCCCC	ACCCCAATT	TTGTATTTAT	TTATTTTTTA	ATTATTTTGT
		GCAGCGATGG				
	481	GGGCGGGGGG	GGGGGGGGGG	CGCGCGCCAG	GCGGGGCGGG	GCGGGGCAAG
		GGGCGGGGCG				
	541	GGGCGAGGCG	GAAAGGTGCG	GCGGCAGCCA	ATCAAAGCGG	CGCGCTCCGA
		AAGTTTCCTT				
50	601	TTATGGCGAG	GCGGCGGCGG	CGGCGGCCCT	ATAAAAAGCG	AAGCGCGCGG
		CGGGCGGGAG				
	661	TCGCTGCGCG	CTGCCTTCGC	CCCGTGCCCC	GCTCCGCGCG	CGCCTCGCGC
		CGCCCGCCCC				
55	721	GGCTCTGACT	GACCGCGTTA	CTCCACAGG	TGAGCGGGCG	GGACGGCCCT
		TCTCCTCCGG				
	781	GCTGTAATTA	GCGCTTG GTT	TAATGACGGC	TTGTTTCTTT	TCTGTGGCTG
		CGTGAAAGCC				
	841	TTGAGGGGCT	CCGGGAGGGC	CCTTTGTGCG	GGGGGAGCGG	CTCGGGGGGT
		GCGTGCGTGT				

	901	GTGTGTGCGT	GGGGAGCGCC	GCGTGCGGCT	CCGCGCTGCC	CGGCGGCTGT
		GAGCGCTGCG				
	961	GGCGCGGCGC	GGGGCTTTGT	GCGCTCCGCA	GTGTGCGCGA	GGGGAGCGCG
		GCCGGGGGCG				
5	1021	GTGCCCCGCG	GTGCGGGGGG	GGCTGCGAGG	GGAACAAAGG	CTGCGTGCGG
		GGTGTGTGCG				
	1081	TGGGGGGGTG	AGCAGGGGGT	GTGGGCGCGT	CGGTCGGGCT	GCAACCCCCC
		CTGCACCCCC				
10	1141	CTCCCCGAGT	TGCTGAGCAC	GGCCCCGGCTT	CGGGTGCGGG	GCTCCGTACG
		GGGCGTGCGG				
	1201	CGGGGCTCGC	CGTGCCGGGC	GGGGGGTGGC	GGCAGGTGGG	GGTGCCGGGC
		GGGGCGGGGC				
	1261	CGCCTCGGGC	CGGGGAGGGC	TCGGGGGAGG	GGCGCGGGCG	CCCCCGGAGC
		GCCGGCGGCT				
15	1321	GTCGAGGCGC	GGCGAGCCGC	AGCCATTGCC	TTTTATGGTA	ATCGTGCGAG
		AGGGCGCAGG				
	1381	GACTTCCTTT	GTCCCAAATC	TGTGCGGAGC	CGAAATCTGG	GAGGCGCCGC
		CGCACCCCCCT				
	1441	CTAGCGGGCG	CGGGGCGAAG	CGGTGCGGCG	CCGGCAGGAA	GGAAATGGGC
20		GGGGAGGGCC				
	1501	TTCGTGCGTC	GCCGCGCCGC	CGTCCCCTTC	TCCCTCTCCA	GCCTCGGGGC
		TGTCCGCGGG				
	1561	GGGACGGCTG	CCTTCGGGGG	GGACGGGGCA	GGGCGGGGTT	CGGCTTCTGG
		CGTGTGACCG				
25	1621	GCGGCTCTAG	AGCCTCTGCT	AACCATGTTC	ATGCCTTCTT	CTTTTTCTTA
		CAGCTCCTGG				
	1681	GCAACGTGCT	GGTATTGTG	CTGTCTCATC	ATTTTGGCAA	AGAATTCGAT
		TGCCATGGCA				
30	1741	ACATATATCC	AGAGAGTACA	GTGCATCTCA	ACATCACTAC	TGGTTGTTCT
		CACCACATTG				
	1801	GTCTCGTGTC	AGATTCCCAG	GGATAGGCTC	TCTAACATAG	GGGTCATAGT
		CGATGAAGGG				
	1861	AAATCACTGA	AGATAGCTGG	ATCCCACGAA	TCGAGGTACA	TAGTACTGAG
		TCTAGTTCCG				
35	1921	GGGGTAGACT	TTGAGAATGG	GTGCGGAACA	GCCCAGGTTA	TCCAGTACAA
		GAGCCTACTG				
	1981	AACAGGCTGT	TAATCCCATT	GAGGGATGCC	TTAGATCTTC	AGGAGGCTCT
		GATAACTGTC				
40	2041	ACCAATGATA	CGACACAAAA	TGCCGGTGCT	CCCCAGTCGA	GATTCTTCGG
		TGCTGTGATT				
	2101	GGTACTATCG	CACTTGAGGT	GGCGACATCA	GCACAAATCA	CCGCAGGGAT
		TGCACTAGCC				
	2161	GAAGCGAGGG	AGGCCAAAAG	AGACATAGCG	CTCATCAAAG	AATCGATGAC
		AAAAACACAC				
45	2221	AAGTCTATAG	AACTGCTGCA	AAACGCTGTG	GGGGAACAAA	TTCTTGCTCT
		AAAGACACTC				
	2281	CAGGATTTCTG	TGAATGATGA	GATCAAACCC	GCAATAAGCG	AATTAGGCTG
		TGAGACTGCT				
50	2341	GCCTTAAGAC	TGGGTATAAA	ATTGACACAG	CATTACTCCG	AGCTGTTAAC
		TGCGTTCCGGC				
	2401	TCGAATTTCTG	GAACCATCGG	AGAGAAGAGC	CTCACGCTGC	AGGCGCTGTC
		TTCACTTTAC				
	2461	TCTGCTAACA	TTACTGAGAT	TATGACCACA	ATCAGGACAG	GGCAGTCTAA
		CATCTATGAT				
55	2521	GTCATTTATA	CAGAACAGAT	CAAAGGAACG	GTGATAGATG	TGGATCTAGA
		GAGATACATG				
	2581	GTCACCCTGT	CTGTGAAGAT	CCCTATTCTT	TCTGAAGTCC	CAGGTGTGCT
		CATACACAAG				
60	2641	GCATCATCTA	TTTCTTACAA	CATAGACGGG	GAGGAATGGT	ATGTGACTGT
		CCCCAGCCAT				

	2701	ATACTCAGTC	GTGCTTCTTT	CTTAGGGGGT	GCAGACATAA	CCGATTGTGT
		TGAGTCCAGA				
	2761	TTGACCTATA	TATGCCCCAG	GGATCCCGCA	CAACTGATAC	CTGACAGCCA
		GCAAAAGTGT				
5	2821	ATCCTGGGGG	ACACAACAAG	GTGTCCTGTC	ACAAAAGTTG	TGGACAGCCT
		TATCCCCAAG				
	2881	TTTGCTTTTG	TGAATGGGGG	CGTTGTTGCT	AACTGCATAG	CATCCACATG
		TACCTGCGGG				
	2941	ACAGGCCGAA	GACCAATCAG	TCAGGATCGC	TCTAAAGGTG	TAGTATTCCT
10		AACCCATGAC				
	3001	AACTGTGGTC	TTATAGGTGT	CAATGGGGTA	GAATTGTATG	CTAACCGGAG
		AGGGCACGAT				
	3061	GCCACTTGGG	GGGTCCAGAA	CTTGACAGTC	GGTCCTGCAA	TTGCTATCAG
		ACCCGTTGAT				
15	3121	ATTTCTCTCA	ACCTTGCTGA	TGCTACGAAT	TTCTTGCAAG	ACTCTAAGGC
		TGAGCTTGAG				
	3181	AAAGCACGGA	AAATCCTCTC	GGAGGTAGGT	AGATGGTACA	ACTCAAGAGA
		GA CTGTGATT				
	3241	ACGATCATAG	TAGTTATGGT	CGTAATATTG	GTGGTCATTA	TAGTGATCAT
20		CATCGTGCTT				
	3301	TATAGACTCA	GAAGGTGAAA	TCACTAGTGA	ATTCACCTCT	CAGGTGCAGG
		CTGCCTATCA				
	3361	GAAGGTGGTG	GCTGGTGTGG	CCAATGCCCT	GGCTCACAAA	TACCACTGAG
		ATCTTTTTCC				
25	3421	CTCTGCCAAA	AATTATGGGG	ACATCATGAA	GCCCCTTGAG	CATCTGACTT
		CTGGCTAATA				
	3481	AAGGAAATTT	ATTTTCATTG	CAATAGTGTG	TTGGAATTTT	TTGTGTCTCT
		CACTCGGAAG				
	3541	GACATATGGG	AGGGCAAATC	ATTTAAAACA	TCAGAATGAG	TATTTGGTTT
30		AGAGTTTGGC				
	3601	AACATATGCC	CATATGCTGG	CTGCCATGAA	CAAAGGTTGG	CTATAAAGAG
		GTCATCAGTA				
	3661	TATGAAACAG	CCCCCTGCTG	TCCATTCCTT	ATTCCATAGA	AAAGCCTTGA
		CTTGAGGTTA				
35	3721	GATTTTTTTT ATATTTTGT TGTGTTATT TTTTCTTTA				ACATCCCTAA AATTTTCCTT
		3781	AGATGTTTAA	CTAGCCAGAT	TTTTCCTCCT	CTCCTGACTA CTCCCAGTCA
			TAGCTGTCCC			
	3841	TCTTCTCTTA	TGGAGATCCC	TCGACCTGCA	GCCCAAGCTT	GGCGTAATCA
		TGGTCATAGC				
40	3901	TGTTTCCTGT	GTGAAATTGT	TATCCGCTCA	CAATTCCACA	CAACATACGA
		GCCGGAAGCA				
	3961	TAAAGTGTA	AGCCTGGGGT	GCCTAATGAG	TGAGCTAACT	CACATTAATT
		GCGTTGCGCT				
	4021	CACTGCCCGC	TTTCCAGTCG	GGAAACCTGT	CGTGCCAGCG	GATCCGCATC
45		TCAATTAGTC				
	4081	AGCAACCATA	GTCCCGCCCC	TAACTCCGCC	CATCCCGCCC	CTAACTCCGC
		CCAGTTCCGC				
	4141	CCATTCTCCG	CCCCATGGCT	GACTAATTTT	TTTTATTTAT	GCAGAGGCCG
		AGGCCGCTC				
50	4201	GGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	AGGCTTTTTT	GGAGGCCTAG
		GCTTTTGCAA				
	4261	AAAGCTAACT	TGTTTATTGC	AGCTTATAAT	GGTTACAAAT	AAAGCAATAG
		CATCACAAAT				
	4321	TTCACAAATA	AAGCATTTTT	TTCACTGCAT	TCTAGTTGTG	GTTTGTCCAA
55		ACTCATCAAT				
	4381	GTATCTTATC	ATGTCTGTCC	GCTTCCTCGC	TCACTGACTC	GCTGCGCTCG
		GTCGTTCCGC				
	4441	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAATACG	GTTATCCACA
		GAATCAGGGG				

	4501	ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	GCCAGCAAAA	GGCCAGGAAC
		CGTAAAAAGG				
	4561	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCCTGA	CGAGCATCAC
		AAAAATCGAC				
5	4621	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG
		TTTCCCCCTG				
	4681	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	CCCTGCCGCT	TACCGGATAC
		CTGTCCGCCT				
10	4741	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC	ATAGCTCACG	CTGTAGGTAT
		CTCAGTTCGG				
	4801	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCA
		CCCGACCGCT				
	4861	GCGCCTTATC	CGGTAACAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC
		TTATCGCCAC				
15	4921	TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT
		GCTACAGAGT				
	4981	TCTTGAAGTG	GTGGCCTAAC	TACGGCTACA	CTAGAAGAAC	AGTATTTGGT
		ATCTGCGCTC				
20	5041	TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC
		AAACAAACCA				
	5101	CCGCTGGTAG	CGGTGGTTTT	TTTGTTTGCA	AGCAGCAGAT	TACGCGCAGA
		AAAAAAGGAT				
	5161	CTCAAGAAGA	TCCTTTGATC	TTTTCTACGG	GGTCTGACGC	TCAGTGGAAC
		GAAAACTCAC				
25	5221	GTTAAGGGAT	TTTGGTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC
		CTTTTAAATT				
	5281	AAAAATGAAG	TTTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTTGGTCT
		GACAGTTAGA				
30	5341	AAAACATCATC	GAGCATCAAA	TGAAACTGCA	ATTTATTCAT	ATCAGGATTA
		TCAATACCAT				
	5401	ATTTTTGAAA	AAGCCGTTTC	TGTAATGAAG	GAGAAAACTC	ACCGAGGCAG
		TTCCATAGGA				
	5461	TGGCAAGATC	CTGGTATCGG	TCTGCGATTG	CGACTCGTCC	AACATCAATA
		CAACCTATTA				
35	5521	ATTTCCCTC	GTCAAAAATA	AGGTTATCAA	GTGAGAAATC	ACCATGAGTG
		ACGACTGAAT				
	5581	CCGGTGAGAA	TGGCAACAGC	TTATGCATTT	CTTTCCAGAC	TTGTTCAACA
		GGCCAGCCAT				
40	5641	TACGCTCGTC	ATCAAAATCA	CTCGCATCAA	CCAAACCGTT	ATTCATTCTG
		GATTGCGCCT				
	5701	GAGCGAGACG	AAATACGCGA	TCGCTGTTAA	AAGGACAATT	ACAAACAGGA
		ATCGAATGCA				
	5761	ACCGGCGCAG	GAACACTGCC	AGCGCATCAA	CAATATTTTC	ACCTGAATCA
		GGATATTCTT				
45	5821	CTAATACCTG	GAATGCTGTT	TTTCCGGGGA	TCGCAGTGGT	GAGTAACCAT
		GCATCATCAG				
	5881	GAGTACGGAT	AAAATGCTTG	ATGGTCGGAA	GAGGCATAAA	TTCCGTCAGC
		CAGTTTAGTC				
50	5941	TGACCATCTC	ATCTGTAACA	TCATTGGCAA	CGCTACCTTT	GCCATGTTTC
		AGAAACAAC				
	6001	CTGGCGCATC	GGGCTTCCCA	TACAATCGAT	AGATTGTGCG	ACCTGATTGC
		CCGACATTAT				
	6061	CGCGAGCCCA	TTTATACCCA	TATAAATCAG	CATCCATGTT	GGAATTTAAT
		CGCGGCCTAG				
55	6121	AGCAAGACGT	TTCCCGTTGA	ATATGGCTCA	TAACACCCCT	TGTATTACTG
		TTTATGTAAG				
	6181	CAGACAGTTT	TATTGTTCAT	GATGATATAT	TTTTATCTTG	TGCAATGTAA
		CATCAGAGAT				
60	6241	TTTGAGACAC	AACAATTGGT	CGAC		
		SEQ ID NO:5				


```

1 ATTGATTATT GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT CATAGCCCAT
61 ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA
CCGCCCAACG
121 ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA
5 ATAGGGACTT
181 TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC CCACTTGGCA
GTACATCAAG
241 TGTATCATAT GCCAAGTACG CCCCTATTG ACGTCAATGA CGGTAAATGG
CCCGCCTGGC
10 301 ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC
TACGTATTAG
361 TCATCGCTAT TACCATGGTC GAGGTGAGCC CCACGTTCTG CTTCACTCTC
CCCATCTCCC
421 CCCCTCCCC ACCCCCAATT TTGTATTTAT TTATTTTTTA ATTATTTTGT
15 GCAGCGATGG
481 GGGCGGGGGG GGGGGGGGGG CGCGCGCCAG GCGGGGCGGG GCGGGGCGAG
GGGCGGGGCG
541 GGGCGAGGCG GAAAGGTGCG GCGGCAGCCA ATCAAAGCGG CGCGCTCCGA
AAGTTTCCTT
20 601 TTATGGCGAG GCGGCGGCGG CGGCGGCCCT ATAAAAAGCG AAGCGCGCGG
CGGGCGGGAG
661 TCGCTGCGCG CTGCCTTCGC CCCGTGCCCC GCTCCGCCGC CGCCTCGCGC
CGCCCGCCCC
721 GGCTCTGACT GACCGCGTTA CTCCCACAGG TGAGCGGGCG GGACGGCCCT
25 TCTCCTCCGG
781 GCTGTAATTA GCGCTTGGTT TAATGACGGC TTGTTTCTTT TCTGTGGCTG
CGTGAAAGCC
841 TTGAGGGGCT CCGGGAGGGC CCTTTGTGCG GGGGGAGCGG CTCGGGGGGT
GCGTGCGTGT
30 901 GTGTGTGCGT GGGGAGCGCC GCGTGCGGCT CCGCGCTGCC CGGCGGCTGT
GAGCGCTGCG
961 GGCGCGGCGC GGGGCTTTGT GCGCTCCGCA GTGTGCGCGA GGGGAGCGCG
GCCGGGGGCG
1021 GTGCCCCGCG GTGCGGGGGG GGCTGCGAGG GGAACAAAGG CTGCGTGCGG
35 GGTGTGTGCG
1081 TGGGGGGGTG AGCAGGGGGT GTGGGCGCGT CGGTCGGGCT GCAACCCCCC
CTGCACCCCC
1141 CTCCCCGAGT TGCTGAGCAC GGCCCGGCTT CGGGTGCGGG GCTCCGTACG
GGGCGTGCGG
40 1201 CGGGGCTCGC CGTGCCGGGC GGGGGGTGGC GGCAGGTGGG GGTGCCGGGC
GGGGCGGGGC
1261 CGCCTCGGGC CGGGGAGGGC TCGGGGGAGG GGC GCGGCGG CCCC CGGAGC
GCCGCGGCT
1321 GTCGAGGCGC GGCGAGCCGC AGCCATTGCC TTTTATGGTA ATCGTGCGAG
45 AGGGCGCAGG
1381 GACTTCCTTT GTCCCAAATC TGTGCGGAGC CGAAATCTGG GAGGCGCCGC
CGCACCCCTT
1441 CTAGCGGGCG CGGGGCGAAG CGGTGCGGCG CCGGCAGGAA GGAAATGGGC
GGGGAGGGCC
50 1501 TTCGTGCGTC GCCGCGCCGC CGTCCCCTTC TCCCTCTCCA GCCTCGGGGC
TGTCCGCGGG
1561 GGGACGGGGC AGGGCGGGGT TCGGCTTCTG GCGTGTGACC GGCGGCTCTA
GAGCCTCTGC
1621 TAACCATGTT CATGCCTTCT TCTTTTCTCT ACAGCTCCTG GGCAACGTGC
55 TGGTTATTGT
1681 GCTGTCTCAT CATTTTGGCA AAGAATTCCT CGAGCATGTG GTCTGAGTTA
AAAATCAGGA
1741 GCAACGACGG AGGTGAAGGA CCAGAGGACG CCAACGACCC CCGGGGAAAG
GGGGTGCAAC

```

	1801	ACATCCATAT	CCAGCCATCT	CTACCTGTTT	ATGGACAGAG	GGTTAGGGAT
		GGTGATAGGG				
	1861	GCAAACGTGA	CTCGTACTGG	TCTACTTCTC	CTAGTGGTAG	CACCACAAAA
		CCAGCATCAG				
5	1921	GTTGGGAGAG	GTCAAGTAAA	GCCGACACAT	GGTTGCTGAT	TCTCTCATTC
		ACCCAGTGGG				
	1981	CTTTGTCAAT	TGCCACAGTG	ATCATCTGTA	TCATAATTTC	TGCTAGACAA
		GGGTATAGTA				
	2041	TGAAAGAGTA	CTCAATGACT	GTAGAGGCAT	TGAACATGAG	CAGCAGGGAG
10		GTGAAAGAGT				
	2101	CACTTACCAG	TCTAATAAGG	CAAGAGGTTA	TAGCAAGGGC	TGTCAACATT
		CAGAGCTCTG				
	2161	TGCAAACCGG	AATCCCAGTC	TTGTTGAACA	AAAACAGCAG	GGATGTCATC
		CAGATGATTG				
15	2221	ATAAGTCGTG	CAGCAGACAA	GAGCTCACTC	AGCACTGTGA	GAGTACGATC
		GCAGTCCACC				
	2281	ATGCCGATGG	AATTGCCCCA	CTTGAGCCAC	ATAGTTTCTG	GAGATGCCCT
		GTCGGAGAAC				
	2341	CGTATCTTAG	CTCAGATCCT	GAAATCTCAT	TGCTGCCTGG	TCCGAGCTTG
20		TTATCTGGTT				
	2401	CTACAACGAT	CTCTGGATGT	GTTAGGCTCC	CTTCACTCTC	AATTGGCGAG
		GCAATCTATG				
	2461	CCTATTCATC	AAATCTCATT	ACACAAGGTT	GTGCTGACAT	AGGGAAATCA
		TATCAGGTCC				
25	2521	TGCAGCTAGG	GTACATATCA	CTCAATTCAG	ATATGTTCCC	TGATCTTAAC
		CCCGTAGTGT				
	2581	CCCACACTTA	TGACATCAAC	GACAATCGGA	AATCATGCTC	TGTGGTGGCA
		ACCGGGACTA				
	2641	GGGGTTATCA	GCTTTGCTCC	ATGCCGACTG	TAGACGAAAG	AACCGACTAC
30		TCTAGTGATG				
	2701	GTATTGAGGA	TCTGGTCCTT	GATGTCCTGG	ATCTCAAAGG	GAGAACTAAG
		TCTCACCGGT				
	2761	ATCGCAACAG	CGAGGTAGAT	CTTGATCACC	CGTTCTCTGC	ACTATACCCC
		AGTGTAGGCA				
35	2821	ACGGCATTGC	AACAGAAGGC	TCATTGATAT	TTCTTGGGTA	TGGTGGACTA
		ACCACCCCTC				
	2881	TGCAGGGTGA	TACAAAATGT	AGGACCCAAG	GATGCCAACA	GGTGTGCGAA
		GACACATGCA				
	2941	ATGAGGCTCT	GAAAATTACA	TGGCTAGGAG	GGAAACAGGT	GGTCAGCGTG
40		ATCATCCAGG				
	3001	TCAATGACTA	TCTCTCAGAG	AGGCCAAAGA	TAAGAGTCAC	AACCATTCCA
		ATCACTCAAA				
	3061	ACTATCTCGG	GGCGGAAGGT	AGATTATTAA	AATTGGGTGA	TCGGGTGTAC
		ATCTATACAA				
45	3121	GATCATCAGG	CTGGCACTCT	CAACTGCAGA	TAGGAGTACT	TGATGTCAGC
		CACCCTTTGA				
	3181	CTATCAACTG	GACACCTCAT	GAAGCCTTGT	CTAGACCAGG	AAATAAAGAG
		TGCAATTGGT				
	3241	ACAATAAGTG	TCCGAAGGAA	TGCATATCAG	GCGTATACAC	TGATGCTTAT
50		CCATTGTCCC				
	3301	CTGATGCAGC	TAACGTCGCT	ACCGTCACGC	TATATGCCAA	TACATCGCGT
		GTCAACCCAA				
	3361	CAATCATGTA	TTCTAACACT	ACTAACATTA	TAAATATGTT	AAGGATAAAG
		GATGTTCAAT				
55	3421	TAGAGGCTGC	ATATACCACG	ACATCGTGTA	TCACGCATTT	TGGTAAAGGC
		TACTGCTTTC				
	3481	ACATCATCGA	GATCAATCAG	AAGAGCCTGA	ATACCTTACA	GCCGATGCTC
		TTTAAGACTA				
	3541	GCATCCCTAA	ATTATGCAAG	GCCGAGTCTT	AAGCGGCCGC	GCATGCGAAT
60		TCACTCCTCA				

	3601	GGTGCAGGCT	GCCTATCAGA	AGGTGGTGGC	TGGTGTGGCC	AATGCCCTGG
		CTCACAAATA				
	3661	CCACTGAGAT	CTTTTTCCCT	CTGCCAAAAA	TTATGGGGAC	ATCATGAAGC
		CCCTTGAGCA				
5	3721	TCTGACTTCT	GGCTAATAAA	GGAAATTTAT	TTTCATTGCA	ATAGTGTGTT
		GGAATTTTTT				
	3781	GTGTCTCTCA	CTCGGAAGGA	CATATGGGAG	GGCAAATCAT	TTAAAACATC
		AGAATGAGTA				
	3841	TTTGGTTTAG	AGTTTGGCAA	CATATGCCCA	TATGCTGGCT	GCCATGAACA
10		AAGGTTGGCT				
	3901	ATAAAGAGGT	CATCAGTATA	TGAAACAGCC	CCCTGCTGTC	TATTCCTTAT
		TCCATAGAAA				
	3961	AGCCTTGACT	TGAGGTTAGA	TTTTTTTTAT	ATTTTGT	TTTCTTTAAC
	4021	ATCCCTAAAA	TTTTCTTAC	ATGTTTTACT	AGCCAGATTT	TTCCTCCTCT
15		CCTGACTACT				
	4081	CCCAGTCATA	GCTGTCCCTC	TTCTCTTATG	GAGATCCCTC	GACCTGCAGC
		CCAAGCTTGG				
	4141	CGTAATCATG	GTCATAGCTG	TTTCCTGTGT	GAAATTGTTA	TCCGCTCACA
		ATTCCACACA				
20	4201	ACATACGAGC	CGGAAGCATA	AAGTGTAAG	CCTGGGGTGC	CTAATGAGTG
		AGCTAACTCA				
	4261	CATTAATTGC	GTTGCGCTCA	CTGCCCGCTT	TCCAGTCGGG	AAACCTGTCTG
		TGCCAGCGGA				
	4321	TCCGCATCTC	AATTAGTCAG	CAACCATAGT	CCCGCCCCTA	ACTCCGCCCA
25		TCCCGCCCT				
	4381	AACTCCGCCC	AGTTCCGCCC	ATTCTCCGCC	CCATGGCTGA	CTAATTTTTT
		TTATTTATGC				
	4441	AGAGGCCGAG	GCCGCCTCGG	CCTCTGAGCT	ATTCCAGAAG	TAGTGAGGAG
		GCTTTTTTGG				
30	4501	AGGCCTAGGC	TTTTGCAAAA	AGCTAACTTG	TTTATTGCAG	CTTATAATGG
		TTACAAATAA				
	4561	AGCAATAGCA	TCACAAATTT	CACAAATAAA	GCATTTTTTT	CACTGCATTC
		TAGTTGTGGT				
	4621	TTGTCCAAAC	TCATCAATGT	ATCTTATCAT	GTCTGTCCGC	TTCTCGCTC
35		ACTGACTCGC				
	4681	TGCGCTCGGT	CGTTCGGCTG	CGGCGAGCGG	TATCAGCTCA	CTCAAAGGCG
		GTAATACGGT				
	4741	TATCCACAGA	ATCAGGGGAT	AACGCAGGAA	AGAACATGTG	AGCAAAAGGC
		CAGCAAAAGG				
40	4801	CCAGGAACCG	TAAAAAGGCC	GC GTTGCTGG	CGTTTTTCCA	TAGGCTCCGC
		CCCCCTGACG				
	4861	AGCATCACAA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA	CCCGACAGGA
		CTATAAAGAT				
	4921	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG	TGCGCTCTCC	TGTTCCGACC
45		CTGCCGCTTA				
	4981	CCGGATACCT	GTCCGCCTTT	CTCCCTTCGG	GAAGCGTGGC	GCTTTCTCAT
		AGCTCACGCT				
	5041	GTAGGTATCT	CAGTTCGGTG	TAGGTCGTTC	GCTCCAAGCT	GGGCTGTGTG
		CACGAACCCC				
50	5101	CGTTTCAGCC	CGACCGCTGC	GCCTTATCCG	GTAACATATCG	TCTTGAGTCC
		AACCCG GTAA				
	5161	GACACGACTT	ATCGCCACTG	GCAGCAGCCA	CTGGTAACAG	GATTAGCAGA
		GCGAGGTATG				
	5221	TAGGCGGTGC	TACAGAGTTC	TTGAAGTGGT	GGCCTAACTA	CGGCTACACT
55		AGAAGAACAG				
	5281	TATTTGGTAT	CTGCGCTCTG	CTGAAGCCAG	TTACCTTCGG	AAAAAGAGTT
		GGTAGCTCTT				
	5341	GATCCGGCAA	ACAAACCACC	GCTGGTAGCG	GTGGTTTTTT	TGTTTGCAAG
		CAGCAGATTA				

```

5401 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
TCTGACGCTC
5461 AGTGGAACGA AAAC TCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
AGGATCTTCA
5 5521 CCTAGATCCT TTAAATTA AAATGAAGTT TTAAATCAAT CTAAAGTATA
TATGAGTAAA
5581 CTTGGTCTGA CAGTTAGAAA AACTCATCGA GCATCAAATG AAAC TGCAAT
TTATTCATAT
5641 CAGGATTATC AATACCATAT TTTTGAAAAA GCCGTTTCTG TAATGAAGGA
10 GAAAACTCAC
5701 CGAGGCAGTT CCATAGGATG GCAAGATCCT GGTATCGGTC TGCGATTCCG
ACTCGTCCAA
5761 CATCAATACA ACCTATTAAT TTCCCCTCGT CAAAAATAAG GTTATCAAGT
GAGAAATCAC
15 5821 CATGAGTGAC GACTGAATCC GGTGAGAATG GCAACAGCTT ATGCATTTCT
TTCCAGACTT
5881 GTTCAACAGG CCAGCCATTA CGCTCGTCAT CAAATCACT CGCATCAACC
AAACCGTTAT
5941 TCATTCGTGA TTGCGCCTGA GCGAGACGAA ATACGCGATC GCTGTTAAAA
20 GGACAATTAC
6001 AACAGGAAT CGAATGCAAC CGGCGCAGGA AACTGCCAG CGCATCAACA
ATATTTTCAC
6061 CTGAATCAGG ATATTCTTCT AATACCTGGA ATGCTGTTTT TCCGGGGATC
GCAGTGGTGA
25 6121 GTAACCATGC ATCATCAGGA GTACGGATAA AATGCTTGAT GGTGGAAGA
GGCATAAATT
6181 CCGTCAGCCA GTTTAGTCTG ACCATCTCAT CTGTAACATC ATTGGCAACG
CTACCTTTGC
6241 CATGTTTCAG AAACAACCTCT GCGCATCGG GCTTCCCAT AATCGATAG
30 ATTGTCGCAC
6301 CTGATTGCCC GACATTATCG CGAGCCCATT TATACCCATA TAAATCAGCA
TCCATGTTGG
6361 AATTTAATCG CGGCCTAGAG CAAGACGTTT CCCGTTGAAT ATGGCTCATA
ACACCCCTTG
35 6421 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTTCATGA TGATATATTT
TTATCTTGTG
6481 CAATGTAACA TCAGAGATTT TGAGACACAA CAATTGGTCG AC
SEQ ID NO:6
1 AGATCTGTGA CATAACTTAT GGTAATGGC CTGCCTGGCT GACTGCCCAA
40 TGACCCCTGC
61 CCAATGATGT CAATAATGAT GTATGTTCCC ATGTAATGCC AATAGGGACT TTCCATTGAT
121 GTCAATGGGT GGAGTATTTA TGGTAACTGC CCACTTGGCA GTACATCAAG
TGTATCATAT
181 GCCAAGTATG CCCCCTATTG ATGTCAATGA TGGTAAATGG CCTGCCTGGC
45 ATTATGCCCA
241 GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC TATGTATTAG
TCATTGCTAT
301 TACCATGGGA ATTCACTAGT GGAGAAGAGC ATGCTTGAGG GCTGAGTGCC
CCTCAGTGGG
50 361 CAGAGAGCAC ATGGCCCACA GTCCCTGAGA AGTTGGGGGG AGGGGTGGGC
AATTGAACTG
421 GTGCCTAGAG AAGGTGGGGC TTGGGTAAAC TGGGAAAGTG ATGTGGTGTA
CTGGCTCCAC
481 CTTTTTCCCC AGGGTGGGGG AGAACCATAT ATAAGTGCAG TAGTCTCTGT
55 GAACATTCAA
541 GCTTCTGCCT TCTCCCTCCT GTGAGTTTGC TAGC
SEQ ID NO:7
1 GCTAGCCACC ATGCAGAGAA GCCCTCTGGA GAAGGCCTCT GTGGTGAGCA
AGCTGTTCTT

```

	61	CAGCTGGACC	AGGCCCATCC	TGAGGAAGGG	CTACAGGCAG	AGACTGGAGC
		TGTCTGACAT				
	121	CTACCAGATC	CCCTCTGTGG	ACTCTGCTGA	CAACCTGTCT	GAGAAGCTGG
		AGAGGGAGTG				
5	181	GGATAGAGAG	CTGGCCAGCA	AGAAGAACCC	CAAGCTGATC	AATGCCCTGA
		GGAGATGCTT				
	241	CTTCTGGAGA	TTCATGTTCT	ATGGCATCTT	CCTGTACCTG	GGGGAAGTGA
		CCAAGGCTGT				
10	301	GCAGCCTCTG	CTGCTGGGCA	GAATCATTGC	CAGCTATGAC	CCTGACAACA
		AGGAGGAGAG				
	361	GAGCATTGCC	ATCTACCTGG	GCATTGGCCT	GTGCCTGCTG	TTCATTGTGA
		GGACCCTGCT				
	421	GCTGCACCCT	GCCATCTTTG	GCCTGCACCA	CATTGGCATG	CAGATGAGGA
		TTGCCATGTT				
15	481	CAGCCTGATC	TACAAGAAAA	CCCTGAAGCT	GTCCAGCAGA	GTGCTGGACA
		AGATCAGCAT				
	541	TGGCCAGCTG	GTGAGCCTGC	TGAGCAACAA	CCTGAACAAG	TTTGATGAGG
		GCCTGGCCCT				
20	601	GGCCCACTTT	GTGTGGATTG	CCCCTCTGCA	GGTGGCCCTG	CTGATGGGCC
		TGATTTGGGA				
	661	GCTGCTGCAG	GCCTCTGCCT	TTTGTGGCCT	GGGCTTCCTG	ATTGTGCTGG
		CCCTGTTTCA				
	721	GGCTGGCCTG	GGCAGGATGA	TGATGAAGTA	CAGGGACCAG	AGGGCAGGCA
		AGATCAGTGA				
25	781	GAGGCTGGTG	ATCACCTCTG	AGATGATTGA	GAACATCCAG	TCTGTGAAGG
		CCTACTGTTG				
	841	GGAGGAAGCT	ATGGAGAAGA	TGATTGAAAA	CCTGAGGCAG	ACAGAGCTGA
		AGCTGACCAG				
30	901	GAAGGCTGCC	TATGTGAGAT	ACTTCAACAG	CTCTGCCTTC	TTCTTCTCTG
		GCTTCTTTGT				
	961	GGTGTTCTG	TCTGTGCTGC	CCTATGCCCT	GATCAAGGGG	ATCATCCTGA
		GAAAGATTTT				
	1021	CACCACCATC	AGCTTCTGCA	TTGTGCTGAG	GATGGCTGTG	ACCAGACAGT
		TCCCCTGGGC				
35	1081	TGTGCAGACC	TGGTATGACA	GCCTGGGGGC	CATCAACAAG	ATCCAGGACT
		TCCTGCAGAA				
	1141	GCAGGAGTAC	AAGACCCTGG	AGTACAACCT	GACCACCACA	GAAGTGGTGA
		TGGAGAATGT				
40	1201	GACAGCCTTC	TGGGAGGAGG	GCTTTGGGGA	GCTGTTTGAG	AAGGCCAAGC
		AGAACAACAA				
	1261	CAACAGAAAG	ACCAGCAATG	GGGATGACTC	CCTGTTCTTC	TCCAATTCT
		CCCTGCTGGG				
	1321	CACACCTGTG	CTGAAGGACA	TCAACTTCAA	GATTGAGAGG	GGGCAGCTGC
		TGGCTGTGGC				
45	1381	TGGATCTACA	GGGGCTGGCA	AGACCAGCCT	GCTGATGATG	ATCATGGGGG
		AGCTGGAGCC				
	1441	TTCTGAGGGC	AAGATCAAGC	ACTCTGGCAG	GATCAGCTTT	TGCAGCCAGT
		TCAGCTGGAT				
50	1501	CATGCCTGGC	ACCATCAAGG	AGAACATCAT	CTTTGGAGTG	AGCTATGATG
		AGTACAGATA				
	1561	CAGGAGTGTG	ATCAAGGCCT	GCCAGCTGGA	GGAGGACATC	AGCAAGTTTG
		CTGAGAAGGA				
	1621	CAACATTGTG	CTGGGGGAGG	GAGGCATTAC	ACTGTCTGGG	GGCCAGAGAG
		CCAGAATCAG				
55	1681	CCTGGCCAGG	GCTGTGTACA	AGGATGCTGA	CCTGTACCTG	CTGGACTCCC
		CCTTTGGCTA				
	1741	CCTGGATGTG	CTGACAGAGA	AGGAGATTTT	TGAGAGCTGT	GTGTGCAAGC
		TGATGGCCAA				
60	1801	CAAGACCAGA	ATCCTGGTGA	CCAGCAAGAT	GGAGCACCTG	AAGAAGGCTG
		ACAAGATCCT				

	1861	GATCCTGCAT	GAGGGCAGCA	GCTACTTCTA	TGGGACCTTC	TCTGAGCTGC
		AGAACCTGCA				
	1921	GCCTGACTTC	AGCTCTAAGC	TGATGGGCTG	TGACAGCTTT	GACCAGTTCT
		CTGCTGAGAG				
5	1981	GAGGAACAGC	ATCCTGACAG	AGACCCTGCA	CAGATTCAGC	CTGGAGGGAG
		ATGCCCCTGT				
	2041	GAGCTGGACA	GAGACCAAGA	AGCAGAGCTT	CAAGCAGACA	GGGGAGTTTG
		GGGAGAAGAG				
	2101	GAAGAACTCC	ATCCTGAACC	CCATCAACAG	CATCAGGAAG	TTCAGCATTG
10		TGCAGAAAAC				
	2161	CCCCCTGCAG	ATGAATGGCA	TTGAGGAAGA	TTCTGATGAG	CCCCTGGAGA
		GGAGACTGAG				
	2221	CCTGGTGCCT	GATTCTGAGC	AGGGAGAGGC	CATCCTGCCT	AGGATCTCTG
		TGATCAGCAC				
15	2281	AGGCCCTACA	CTGCAGGCCA	GAAGGAGGCA	GTCTGTGCTG	AACCTGATGA
		CCCACTCTGT				
	2341	GAACCAGGGC	CAGAACATCC	ACAGGAAAAC	CACAGCCTCC	ACCAGGAAAG
		TGAGCCTGGC				
	2401	CCCTCAGGCC	AATCTGACAG	AGCTGGACAT	CTACAGCAGG	AGGCTGTCTC
20		AGGAGACAGG				
	2461	CCTGGAGATT	TCTGAGGAGA	TCAATGAGGA	GGACCTGAAA	GAGTGCTTCT
		TTGATGACAT				
	2521	GGAGAGCATC	CCTGCTGTGA	CCACCTGGAA	CACCTACCTG	AGATACATCA
		CAGTGCACAA				
25	2581	GAGCCTGATC	TTTGTGCTGA	TCTGGTGCCT	GGTGATCTTC	CTGGCTGAAG
		TGGCTGCCTC				
	2641	TCTGGTGGTG	CTGTGGCTGC	TGGGAAACAC	CCCACTGCAG	GACAAGGGCA
		ACAGCACCCA				
	2701	CAGCAGGAAC	AACAGCTATG	CTGTGATCAT	CACCTCCACC	TCCAGCTACT
30		ATGTGTTCTA				
	2761	CATCTATGTG	GGAGTGGCTG	ATACCCTGCT	GGCTATGGGC	TTCTTTAGAG
		GCCTGCCCT				
	2821	GGTGCACACA	CTGATCACAG	TGAGCAAGAT	CCTCCACCAC	AAGATGCTGC
		ACTCTGTGCT				
35	2881	GCAGGCTCCT	ATGAGCACCC	TGAATACCCT	GAAGGCTGGG	GGCATCCTGA
		ACAGATTCTC				
	2941	CAAGGATATT	GCCATCCTGG	ATGACCTGCT	GCCTCTCACC	ATCTTTGACT
		TCATCCAGCT				
	3001	GCTGCTGATT	GTGATTGGGG	CCATTGCTGT	GGTGGCAGTG	CTGCAGCCCT
40		ACATCTTTGT				
	3061	GGCCACAGTG	CCTGTGATTG	TGGCCTTCAT	CATGCTGAGG	GCCTACTTTC
		TGCAGACCTC				
	3121	CCAGCAGCTG	AAGCAGCTGG	AGTCTGAGGG	CAGAAGCCCC	ATCTTCACCC
		ACCTGGTGAC				
45	3181	AAGCCTGAAG	GGCCTGTGGA	CCCTGAGAGC	CTTTGGCAGG	CAGCCCTACT
		TTGAGACCCT				
	3241	GTTCCACAAG	GCCCTGAACC	TGCACACAGC	CAACTGGTTC	CTCTACCTGT
		CCACCCTGAG				
	3301	ATGGTTCCAG	ATGAGAATTG	AGATGATCTT	TGTCATCTTC	TTCATTGCTG
50		TGACCTTCAT				
	3361	CAGCATTCTG	ACCACAGGAG	AGGGAGAGGG	CAGAGTGGGC	ATTATCCTGA
		CCCTGGCCAT				
	3421	GAACATCATG	AGCACACTGC	AGTGGGCAGT	GAACAGCAGC	ATTGATGTGG
		ACAGCCTGAT				
55	3481	GAGGAGTGTG	AGCAGAGTGT	TCAAGTTCAT	TGATATGCCC	ACAGAGGGCA
		AGCCTACCAA				
	3541	GAGCACCAAG	CCCTACAAGA	ATGGCCAGCT	GAGCAAAGTG	ATGATCATTG
		AGAACAGCCA				
60	3601	TGTGAAGAAG	GATGATATCT	GGCCCAGTGG	AGGCCAGATG	ACAGTGAAGG
		ACCTGACAGC				

```

3661 CAAGTACACA GAGGGGGGGCA ATGCTATCCT GGAGAACATC TCCTTCAGCA
TCTCCCTGG
3721 CCAGAGAGTG GGAAGTGGTGG GAAGAACAGG CTCTGGCAAG TCTACCCTGC
TGTCTGCCTT
5 3781 CCTGAGGCTG CTGAACACAG AGGGAGAGAT CCAGATTGAT GGAGTGTCTT
GGGACAGCAT
3841 CACACTGCAG CAGTGGAGGA AGGCCTTTGG TGTGATCCCC CAGAAAGTGT
TCATCTTCAG
3901 TGGCACCTTC AGGAAGAACC TGGACCCCTA TGAGCAGTGG TCTGACCAGG
10 AGATTTGGAA
3961 AGTGGCTGAT GAAGTGGGCC TGAGAAGTGT GATTGAGCAG TTCCCTGGCA
AGCTGGACTT
4021 TGTCTCTGGT GATGGGGGCT GTGTGCTGAG CCATGGCCAC AAGCAGCTGA
TGTGCCTGGC
15 4081 CAGATCAGTG CTGAGCAAGG CCAAGATCCT GCTGCTGGAT GAGCCTTCTG
CCCACCTGGA
4141 TCCTGTGACC TACCAGATCA TCAGGAGGAC CCTCAAGCAG GCCTTTGCTG
ACTGCACAGT
4201 CATCCTGTGT GAGCACAGGA TTGAGGCCAT GCTGGAGTGC CAGCAGTTCC
20 TGGTGATTGA
4261 GGAGAACAAA GTGAGGCAGT ATGACAGCAT CCAGAAGCTG CTGAATGAGA
GGAGCCTGTT
4321 CAGGCAGGCC ATCAGCCCCT CTGATAGAGT GAAGCTGTTC CCCCACAGGA
ACAGCTCCAA
25 4381 GTGCAAGAGC AAGCCCCAGA TTGCTGCCCT GAAGGAGGAG ACAGAGGAGG
AAGTGCAGGA
4441 CACCAGGCTG TGAGGGCCCC
SEQ ID NO:8
1 GGGCCCAATC AACCTCTGGA TTACAAAATT TGTGAAAGAT TGAAGTGGTAT TCTTAACTAT
30 61 GTTGCTCCTT TTACGCTATG TGGATACGCT GCTTTAATGC CTTTGTATCA TGCTATTGCT
121 TCCCGTATGG CTTTCATTTT CTCCTCCTTG TATAAATCCT GGTTGCTGTC
TCTTTATGAG
181 GAGTTGTGGC CCGTTGTCAG GCAACGTGGC GTGGTGTGCA CTGTGTTTGC
TGACGCAACC
35 241 CCCACTGGTT GGGGCATTGC CACCACCTGT CAGCTCCTTT CCGGGACTTT
CGCTTTCCCC
301 CTCCCTATTG CCACGGCGGA ACTCATCGCC GCCTGCCTTG CCCGCTGCTG
GACAGGGGCT
361 CGGCTGTTGG GCACTGACAA TTCCGTGGTG TTGTCGGGGA AATCATCGTC
40 CTTTCCTTGG
421 CTGCTCGCCT GTGTTGCCAC CTGGATTCTG CGCGGGACGT CCTTCTGCTA
CGTCCCTTCG
481 GCCCTCAATC CAGCGGACCT TCCTTCCCGC GGCCTGCTGC CGGCTCTGCG
GCCTCTTCCG
45 541 CGTCTTCGCC TTCGCCCTCA GACGAGTCGG ATCTCCCTTT GGGCCGCCTC
CCCGCAAGCT
SEQ ID NO:9
1 GGTACCTCAA TATTGGCCAT TAGCCATATT ATTCATTGGT TATATAGCAT AAATCAATAT
50 61 TGGCTATTGG CCATTGCATA CGTTGTATCT ATATCATAAT ATGTACATTT ATATTGGCTC
121 ATGTCCAATA TGACCGCCAT GTTGGCATTG ATTATTGACT AGTTATTAAT
AGTAATCAAT
181 TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
TTACGGTAAA
241 TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA
55 TGACGTATGT
301 TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT
ATTTACGGTA
361 AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTCCGCCCC
CTATTGACGT

```

	421	CAATGACGGT	AAATGGCCCG	CCTGGCATT	TGCCCAGTAC	ATGACCTTAC
		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
5	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
		TCCACCCCAT				
	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
10	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
	721	TATATAAGCA	GAGCTCGCTG	GCTTGTAAC	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTCT	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
15	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCTT
		CATTGACGCC				
	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
20	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTAGGAG
		AGGCCGTAGC				
25	1141	CGTAACTACT	CTTGGGCAAG	TAGGGCAGGC	GGTGGGTACG	CAATGGGGGC
		GGCTACCTCA				
	1201	GCACTAAATA	GGAGACAATT	AGACCAATTT	GAGAAAATAC	GACTTCGCCC
		GAACGGAAAG				
30	1261	AAAAAGTACC	AAATTAAACA	TTTAATATGG	GCAGGCAAGG	AGATGGAGCG
		CTTCGGCCTC				
	1321	CATGAGAGGT	TGTTGGAGAC	AGAGGAGGGG	TGTAAAAGAA	TCATAGAAGT
		CCTCTACCCC				
	1381	CTAGAACCAA	CAGGATCGGA	GGGCTTAAAA	AGTCTGTTCA	ATCTTGTGTG
		CGTGCTATAT				
35	1441	TGCTTGCACA	AGGAACAGAA	AGTGAAAGAC	ACAGAGGAAG	CAGTAGCAAC
		AGTAAGACAA				
	1501	CACTGCCATC	TAGTGGAAAA	AGAAAAAAGT	GCAACAGAGA	CATCTAGTGG
		ACAAAAGAAA				
40	1561	AATGACAAGG	GAATAGCAGC	GCCACCTGGT	GGCAGTCAGA	ATTTTCCAGC
		GCAACAACAA				
	1621	GGAAATGCCT	GGGTACATGT	ACCCTTGTCA	CCGCGCACCT	TAAATGCGTG
		GGTAAAAGCA				
	1681	GTAGAGGAGA	AAAAATTTGG	AGCAGAAATA	GTACCCATTT	TTTTGTTTCA
		AGCCCTATCG				
45	1741	AATTCCCGTT	TGTGCTAGGG	TTCTTAGGCT	TCTTGGGGGC	TGCTGGAAC
		GCAATGGGAG				
	1801	CAGCGGCGAC	AGCCCTGACG	GTCCAGTCTC	AGCATTTGCT	TGCTGGGATA
		CTGCAGCAGC				
50	1861	AGAAGAATCT	GCTGGCGGCT	GTGGAGGCTC	AACAGCAGAT	GTTGAAGCTG
		ACCATTTGGG				
	1921	GTGTTAAAAA	CCTCAATGCC	CGCGTCACAG	CCCTTGAGAA	GTACCTAGAG
		GATCAGGCAC				
	1981	GACTAACTC	CTGGGGGTGC	GCATGGAAAC	AAGTATGTCA	TACCACAGTG
		GAGTGGCCCT				
55	2041	GGACAAATCG	GACTCCGGAT	TGGCAAATA	TGACTTGGTT	GGAGTGGGAA
		AGACAAATAG				
	2101	CTGATTTGGA	AAGCAACATT	ACGAGACAAT	TAGTGAAGGC	TAGAGAACAA
		GAGGAAAAGA				
60	2161	ATCTAGATGC	CTATCAGAAG	TTAACTAGTT	GGTCAGATTT	CTGGTCTTGG
		TTCGATTTCT				

	2221	CAAAATGGCT	TAACATTTTA	AAAATGGGAT	TTTtagtaAT	AGTAGGAATA
		ATAGGGTTAA				
	2281	GATTACTTTA	CACAGTATAT	GGATGTATAG	TGAGGGTTAG	GCAGGGATAT
		GTTCTCTAT				
5	2341	CTCCACAGAT	CCATATCCGC	GGCAATTTTA	AAAGAAAGGG	AGGAATAGGG
		GGACAGACTT				
	2401	CAGCAGAGAG	ACTAATTAAT	ATAATAACAA	CACAATTAGA	AATACAACAT
		TTACAAACCA				
	2461	AAATTCAAAA	AATTTTAAAT	TTTAGAGCCG	CGGAGATCTG	TTACATAACT
10		TATGGTAAAT				
	2521	GGCCTGCCTG	GCTGACTGCC	CAATGACCCC	TGCCCAATGA	TGTCAATAAT
		GATGTATGTT				
	2581	CCCATGTAAT	GCCAATAGGG	ACTTTCCATT	GATGTCAATG	GGTGGAGTAT
		TTATGGTAAC				
15	2641	TGCCCACTTG	GCAGTACATC	AAGTGTATCA	TATGCCAAGT	ATGCCCCCTA
		TTGATGTCAA				
	2701	TGATGGTAAA	TGGCCTGCCT	GGCATTATGC	CCAGTACATG	ACCTTATGGG
		ACTTTCCTAC				
	2761	TTGGCAGTAC	ATCTATGTAT	TAGTCATTGC	TATTACCATG	GGAATTCACT
20		AGTGGAGAAG				
	2821	AGCATGCTTG	AGGGCTGAGT	GCCCCTCAGT	GGGCAGAGAG	CACATGGCCC
		ACAGTCCCTG				
	2881	AGAAGTTGGG	GGGAGGGGTG	GGCAATTGAA	CTGGTGCCTA	GAGAAGGTGG
		GGCTTGGGTA				
25	2941	AACTGGGAAA	GTGATGTGGT	GTACTGGCTC	CACCTTTTTC	CCCAGGGTGG
		GGGAGAACCA				
	3001	TATATAAGTG	CAGTAGTCTC	TGTGAACATT	CAAGCTTCTG	CCTTCTCCCT
		CCTGTGAGTT				
	3061	TGCTAGCCAC	CATGCCCAGC	TCTGTGTCTT	GGGGCATTCT	GCTGCTGGCT
30		GGCCTGTGCT				
	3121	GTCTGGTGCC	TGTGTCCCTG	GCTGAGGACC	CTCAGGGGGA	TGCTGCCCAG
		AAAACAGACA				
	3181	CCTCCCACCA	TGACCAGGAC	CACCCCACCT	TCAACAAGAT	CACCCCCAAC
		CTGGCAGAGT				
35	3241	TTGCCTTCAG	CCTGTACAGA	CAGCTGGCCC	ACCAGAGCAA	CAGCACCAAC
		ATCTTTTCA				
	3301	GCCCTGTGTC	CATTGCCACA	GCCTTTGCCA	TGCTGAGCCT	GGGCACCAAG
		GCTGACACCC				
	3361	ATGATGAGAT	CCTGGAAGGC	CTGAACCTCA	ACCTGACAGA	GATCCCTGAG
40		GCCCAGATCC				
	3421	ATGAGGGCTT	CCAGGAACTG	CTGAGAACCC	TGAACCAGCC	AGACAGCCAG
		CTGCAGCTGA				
	3481	CAACAGGCAA	TGGGCTGTTC	CTGTCTGAGG	GCCTGAAGCT	GGTGGACAAG
		TTTCTGGAAG				
45	3541	ATGTGAAGAA	GCTGTACCAC	TCTGAGGCCT	TCACAGTGAA	CTTTGGGGAC
		ACAGAAGAGG				
	3601	CCAAGAAACA	GATCAATGAC	TATGTGGAAA	AGGGCACCCA	GGGCAAGATT
		GTGGACCTTG				
	3661	TGAAAGAGCT	GGACAGGGAC	ACTGTGTTTG	CCCTTGTGAA	CTACATCTTC
50		TTCAAGGGCA				
	3721	AGTGGGAGAG	GCCCTTTGAA	GTGAAGGACA	CTGAGGAAGA	GGACTTCCAT
		GTGGACCAAG				
	3781	TGACCACAGT	GAAGGTGCCA	ATGATGAAGA	GA CTGGGGAT	GTTCAATATC
		CAGCACTGCA				
55	3841	AGAACTGAG	CAGCTGGGTG	CTGCTGATGA	AGTACCTGGG	CAATGCTACA
		GCCATATTCT				
	3901	TTCTGCCTGA	TGAGGGCAAG	CTGCAGCACC	TGGA AAATGA	GCTGACCCAT
		GACATCATCA				
	3961	CCAAATTTCT	GGAAAATGAG	GACAGAAGAT	CTGCCAGCCT	GCATCTGCCC
60		AAGCTGAGCA				

	4021 TCACAGGCAC	ATATGACCTG	AAGTCTGTGC	TGGGACAGCT	GGGAATCACC
	AAGGTGTTCA				
	4081 GCAATGGGGC	AGACCTGAGT	GGAGTGACAG	AGGAAGCCCC	TCTGAAGCTG
	TCCAAGGCTG				
5	4141 TGCACAAGGC	AGTGCTGACC	ATTGATGAGA	AGGGCACAGA	GGCTGCTGGG
	GCCATGTTTC				
	4201 TGGAAGCCAT	CCCCATGTCC	ATCCCCCAG	AAGTGAAGTT	CAACAAGCCC
	TTTGTGTTCC				
	4261 TGATGATTGA	GCAGAACACC	AAGAGCCCCC	TGTTTCATGGG	CAAGGTTGTG
10	AACCCACCCC				
	4321 AGAAATGAGG	GCCCAATCAA	CCTCTGGATT	ACAAAATTTG	TGAAAGATTG
	ACTGGTATTC				
	4381 TTAACATATGT	TGCTCCTTTT	ACGCTATGTG	GATACGCTGC	TTAATGCCT
	TTGTATCATG				
15	4441 CTATTGCTTC	CCGTATGGCT	TTCATTTTCT	CCTCCTTGTA	TAAATCCTGG
	TTGCTGTCTC				
	4501 TTTATGAGGA	GTTGTGGCCC	GTTGTCAGGC	AACGTGGCGT	GGTGTGCACT
	GTGTTTGCTG				
	4561 ACGCAACCCC	CACTGGTTGG	GGCATTGCCA	CCACCTGTCA	GCTCCTTTCC
20	GGGACTTTCCG				
	4621 CTTTCCCCCT	CCCTATTGCC	ACGGCGGAAC	TCATCGCCGC	CTGCCTTGCC
	CGCTGCTGGA				
	4681 CAGGGGCTCG	GCTGTTGGGC	ACTGACAATT	CCGTGGTGTT	GTCGGGGAAA
	TCATCGTCCT				
25	4741 TTCCTTGGCT	GCTCGCCTGT	GTTGCCACCT	GGATTCTGCG	CGGGACGTCC
	TTCTGCTACG				
	4801 TCCCTTCGGC	CCTCAATCCA	GCGGACCTTC	CTTCCCGCGG	CCTGCTGCCG
	GCTCTGCGGC				
	4861 CTCTTCGCG	TCTTCGCCTT	CGCCCTCAGA	CGAGTCGGAT	CTCCCTTTGG
30	GCCGCCTCCC				
	4921 CGCAAGCTTC	GCACTTTTTA	AAAGAAAAGG	GAGGACTGGA	TGGGATTTAT
	TACTCCGATA				
	4981 GGACGCTGGC	TTGTAACCTCA	GTCTCTTACT	AGGAGACCAG	CTTGAGCCTG
	GGTGTTTCGT				
35	5041 GGTTAGCCTA	ACCTGGTTGG	CCACCAGGGG	TAAGGACTCC	TTGGCTTAGA
	AAGCTAATAA				
	5101 ACTTGCCTGC	ATTAGAGCTC	TTACGCGTCC	CGGGCTCGAG	ATCCGCATCT
	CAATTAGTCA				
	5161 GCAACCATAG	TCCCGCCCCT	AACTCCGCCC	ATCCCGCCCC	TAACTCCGCC
40	CAGTTCCGCC				
	5221 CATTCTCCGC	CCCATGGCTG	ACTAATTTTT	TTTATTTATG	CAGAGGCCGA
	GGCCGCCTCG				
	5281 GCCTCTGAGC	TATTCCAGAA	GTAGTGAGGA	GGCTTTTTTG	GAGGCCTAGG
	CTTTTGCAAA				
45	5341 AAGCTAACTT	GTTTATTGCA	GCTTATAATG	GTTACAAATA	AAGCAATAGC
	ATCACAAATT				
	5401 TCACAAATAA	AGCATTTTTT	TCACTGCATT	CTAGTTGTGG	TTTGTCCAAA
	CTCATCAATG				
	5461 TATCTTATCA	TGTCTGTCCG	CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG
50	TCGTTCCGGCT				
	5521 GCGGCGAGCG	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG
	AATCAGGGGA				
	5581 TAACGCAGGA	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC
	GTAAAAAGGC				
55	5641 CGCGTTGCTG	GCGTTTTTCC	ATAGGCTCCG	CCCCCTGAC	GAGCATCACA
	AAAATCGACG				
	5701 CTCAAGTCAG	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT
	TTCCCCCTGG				
	5761 AAGCTCCCTC	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC
60	TGTCCGCCTT				

	5821	TCTCCCTTCG	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC
		TCAGTTCGGT				
	5881	GTAGGTCGTT	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC
		CCGACCGCTG				
5	5941	CGCCTTATCC	GGTAACTATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT
		TATCGCCACT				
	6001	GGCAGCAGCC	ACTGGTAACA	GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTG
		CTACAGAGTT				
	6061	CTTGAAGTGG	TGGCCTAACT	ACGGCTACAC	TAGAAGAACA	GTATTTGGTA
10		TCTGCGCTCT				
	6121	GCTGAAGCCA	GTTACCTTCG	GAAAAAGAGT	TGGTAGCTCT	TGATCCGGCA
		AACAAACCAC				
	6181	CGCTGGTAGC	GGTGGTTTTT	TTGTTTGCAA	GCAGCAGATT	ACGCGCAGAA
		AAAAAGGATC				
15	6241	TCAAGAAGAT	CCTTTGATCT	TTTCTACGGG	GTCTGACGCT	CAGTGGAACG
		AAACTCACG				
	6301	TTAAGGGATT	TTGGTCATGA	GATTATCAAA	AAGGATCTTC	ACCTAGATCC
		TTTTAAATTA				
	6361	AAAATGAAGT	TTTAAATCAA	TCTAAAGTAT	ATATGAGTAA	ACTTGGTCTG
20		ACAGTTAGAA				
	6421	AAACTCATCG	AGCATCAAAT	GAAACTGCAA	TTTATTCATA	TCAGGATTAT
		CAATACCATA				
	6481	TTTTTGAAAA	AGCCGTTTCT	GTAATGAAGG	AGAAAACTCA	CCGAGGCAGT
		TCCATAGGAT				
25	6541	GGCAAGATCC	TGGTATCGGT	CTGCGATTCC	GACTCGTCCA	ACATCAATAC
		AACCTATTAA				
	6601	TTTCCCCTCG	TCAAAAATAA	GGTTATCAAG	TGAGAAATCA	CCATGAGTGA
		CGACTGAATC				
	6661	CGGTGAGAAT	GGCAACAGCT	TATGCATTTC	TTTCCAGACT	TGTTCAACAG
30		GCCAGCCATT				
	6721	ACGCTCGTCA	TCAAAATCAC	TCGCATCAAC	CAAACCGTTA	TTCATTCTGT
		ATTGCGCCTG				
	6781	AGCGAGACGA	AATACGCGAT	CGCTGTATAA	AGGACAATTA	CAAACAGGAA
		TCGAATGCAA				
35	6841	CCGGCGCAGG	AACACTGCCA	GCGCATCAAC	AATATTTTCA	CCTGAATCAG
		GATATTCTTC				
	6901	TAATACCTGG	AATGCTGTTT	TTCCGGGGAT	CGCAGTGGTG	AGTAACCATG
		CATCATCAGG				
	6961	AGTACGGATA	AAATGCTTGA	TGGTCGGAAG	AGGCATAAAT	TCCGTCAGCC
40		AGTTTAGTCT				
	7021	GACCATCTCA	TCTGTAACAT	CATTGGCAAC	GCTACCTTTG	CCATGTTTCA
		GAAACAATC				
	7081	TGGCGCATCG	GGCTTCCCAT	ACAATCGATA	GATTGTCGCA	CCTGATTGCC
		CGACATTATC				
45	7141	GCGAGCCCAT	TTATACCCAT	ATAAATCAGC	ATCCATGTTG	GAATTTAATC
		GCGGCCTAGA				
	7201	GCAAGACGTT	TCCCGTTGAA	TATGGCTCAT	AACACCCCTT	GTATTACTGT
		TTATGTAAGC				
	7261	AGACAGTTTT	ATTGTTTCATG	ATGATATATT	TTTATCTTGT	GCAATGTAAC
50		ATCAGAGATT				
	7321	TTGAGACACA	ACAATTGGTC	GACGGATCC		
		SEQ ID NO:10				
	1	GGTACCTCAA	TATTGGCCAT	TAGCCATATT	ATTCATTGGT	TATATAGCAT
		AAATCAATAT				
	61	TGGCTATTGG	CCATTGCATA	CGTTGTATCT	ATATCATAAT	ATGTACATTT
		ATATTGGCTC				
55	121	ATGTCCAATA	TGACCGCCAT	GTTGGCATTG	ATTATTGACT	AGTTATTAAT
		AGTAATCAAT				
	181	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCCGC	GTTACATAAC
		TTACGGTAAA				
	241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA
60		TGACGTATGT				

	301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT
		ATTTACGGTA				
	361	AACTGCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTCCGCCCC
		CTATTGACGT				
5	421	CAATGACGGT	AAATGGCCCG	CCTGGCATT	TGCCCAGTAC	ATGACCTTAC
		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
10		TCCACCCCAT				
	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
15	721	TATATAAGCA	GAGCTCGCTG	GCTTGTA	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTTCG	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCTT
20		CATTGACGCC				
	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
25	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTAGGAG
		AGGCCGTAGC				
	1141	CGTAACTACT	CTTGGGCAAG	TAGGGCAGGC	GGTGGGTACG	CAATGGGGGC
30		GGCTACCTCA				
	1201	GCACTAAATA	GGAGACAATT	AGACCAATTT	GAGAAAATAC	GACTTCGCCC
		GAACGGAAAG				
	1261	AAAAAGTACC	AAATTAAACA	TTTAATATGG	GCAGGCAAGG	AGATGGAGCG
		CTTCGGCCTC				
35	1321	CATGAGAGGT	TGTTGGAGAC	AGAGGAGGGG	TGTAAAAGAA	TCATAGAAGT
		CCTCTACCCC				
	1381	CTAGAACCAA	CAGGATCGGA	GGGCTTAAAA	AGTCTGTTCA	ATCTTGTGTG
		CGTGCTATAT				
	1441	TGCTTGCACA	AGGAACAGAA	AGTGAAAGAC	ACAGAGGAAG	CAGTAGCAAC
40		AGTAAGACAA				
	1501	CACTGCCATC	TAGTGGAAAA	AGAAAAAGT	GCAACAGAGA	CATCTAGTGG
		ACAAAAGAAA				
	1561	AATGACAAGG	GAATAGCAGC	GCCACCTGGT	GGCAGTCAGA	ATTTTCCAGC
		GCAACAACAA				
45	1621	GGAAATGCCT	GGGTACATGT	ACCCTTGTCA	CCGCGCACCT	TAAATGCGTG
		GGTAAAAGCA				
	1681	GTAGAGGAGA	AAAAATTTGG	AGCAGAAATA	GTACCCATTT	TTTTGTTTCA
		AGCCCTATCG				
	1741	AATTCCTGTT	TGTGCTAGGG	TTCTTAGGCT	TCTTGGGGGC	TGCTGGA
50		GCAATGGGAG				
	1801	CAGCGGCGAC	AGCCCTGACG	GTCCAGTCTC	AGCATTTGCT	TGCTGGGATA
		CTGCAGCAGC				
	1861	AGAAGAATCT	GCTGGCGGCT	GTGGAGGCTC	AACAGCAGAT	GTTGAAGCTG
		ACCATTTGGG				
55	1921	GTGTTAAAAA	CCTCAATGCC	CGCGTCACAG	CCCTTGAGAA	GTACCTAGAG
		GATCAGGCAC				
	1981	GACTAACTC	CTGGGGGTGC	GCATGGAAAC	AAGTATGTCA	TACCACAGTG
		GAGTGGCCCT				
	2041	GGACAAATCG	GACTCCGGAT	TGGCAAAATA	TGACTTGGTT	GGAGTGGGAA
60		AGACAAATAG				

	2101	CTGATTTGGA	AAGCAACATT	ACGAGACAAT	TAGTGAAGGC	TAGAGAACAA
		GAGGAAAAGA				
	2161	ATCTAGATGC	CTATCAGAAG	TTAACTAGTT	GGTCAGATTT	CTGGTCTTGG
		TTCGATTTCT				
5	2221	CAAAATGGCT	TAACATTTTA	AAAATGGGAT	TTTTAGTAAT	AGTAGGAATA
		ATAGGGTTAA				
	2281	GATTACTTTA	CACAGTATAT	GGATGTATAG	TGAGGGTTAG	GCAGGGATAT
		GTTCTCTAT				
10	2341	CTCCACAGAT	CCATATCCGC	GGCAATTTTA	AAAGAAAGGG	AGGAATAGGG
		GGACAGACTT				
	2401	CAGCAGAGAG	ACTAATTAAT	ATAATAACAA	CACAATTAGA	AATACAACAT
		TTACAAACCA				
	2461	AAATTCAAAA	AATTTTAAAT	TTTAGAGCCG	CGGAGATCTG	TTACATAACT
		TATGGTAAAT				
15	2521	GGCCTGCCTG	GCTGACTGCC	CAATGACCCC	TGCCCAATGA	TGTCAATAAT
		GATGTATGTT				
	2581	CCCATGTAAT	GCCAATAGGG	ACTTTCCATT	GATGTCAATG	GGTGGAGTAT
		TTATGGTAAC				
20	2641	TGCCCACTTG	GCAGTACATC	AAGTGTATCA	TATGCCAAGT	ATGCCCCCTA
		TTGATGTCAA				
	2701	TGATGGTAAA	TGGCCTGCCT	GGCATTATGC	CCAGTACATG	ACCTTATGGG
		ACTTTCCTAC				
	2761	TTGGCAGTAC	ATCTATGTAT	TAGTCATTGC	TATTACCATG	GGAATTCACT
		AGTGGAGAAG				
25	2821	AGCATGCTTG	AGGGCTGAGT	GCCCCTCAGT	GGGCAGAGAG	CACATGGCCC
		ACAGTCCCTG				
	2881	AGAAGTTGGG	GGGAGGGGTG	GGCAATTGAA	CTGGTGCCTA	GAGAAGGTGG
		GGCTTGGGTA				
30	2941	AACTGGGAAA	GTGATGTGGT	GTACTGGCTC	CACCTTTTTT	CCCAGGGTGG
		GGGAGAACCA				
	3001	TATATAAGTG	CAGTAGTCTC	TGTGAACATT	CAAGCTTCTG	CCTTCTCCCT
		CCTGTGAGTT				
	3061	TGCTAGCCAC	CATGGGAGTG	AAGGTGCTGT	TTGCCCTGAT	CTGCATTGCT
		GTGGCTGAGG				
35	3121	CCAAGCCCAC	AGAGAACAAT	GAGGACTTCA	ACATTGTGGC	TGTGGCCAGC
		AACTTTGCCA				
	3181	CCACAGACCT	GGATGCTGAC	AGGGGCAAGC	TGCCTGGCAA	GAAGCTGCCC
		CTGGAAGTCC				
40	3241	TGAAAGAGAT	GGAAGCCAAT	GCCAGGAAGG	CTGGCTGCAC	AAGAGGCTGT
		CTGATCTGCC				
	3301	TGAGCCACAT	CAAGTGCACC	CCCAAGATGA	AGAAGTTCAT	CCCTGGCAGG
		TGCCACACCT				
	3361	ATGAAGGGGA	CAAAGAGTCT	GCCCAGGGGG	GAATTGGAGA	GGCCATTGTG
		GACATCCCTG				
45	3421	AGATCCCTGG	CTTCAAGGAC	CTGGAACCCA	TGGAACAGTT	CATTGCCCAG
		GTGGACCTGT				
	3481	GTGTGGACTG	CACTACAGGC	TGTCTCAAGG	GCCTGGCCAA	TGTGCAGTGC
		TCTGACCTGC				
50	3541	TGAAGAAGTG	GCTGCCCCAG	AGATGTGCCA	CCTTTGCCAG	CAAGATCCAG
		GGCCAGGTGG				
	3601	ACAAGATCAA	GGGAGCTGGG	GGAGATTGAT	GAGGGCCCCA	TCAACCTCTG
		GATTACAAAA				
	3661	TTTGTGAAAG	ATTGACTGGT	ATTCTTAACT	ATGTTGCTCC	TTTTACGCTA
		TGTGGATACG				
55	3721	CTGCTTTAAT	GCCTTTGTAT	CATGCTATTG	CTTCCCGTAT	GGCTTTCATT
		TTCTCCTCCT				
	3781	TGTATAAATC	CTGGTTGCTG	TCTCTTTATG	AGGAGTTGTG	GCCCGTTGTC
		AGGCAACGTG				
60	3841	GCGTGGTGTG	CACTGTGTTT	GCTGACGCAA	CCCCCACTGG	TTGGGGCATT
		GCCACCACCT				

	3901	GTCAGCTCCT	TTCCGGGACT	TTCGCTTTCC	CCCTCCCTAT	TGCCACGGCG
		GAAGTCATCG				
	3961	CCGCCTGCCT	TGCCCCGCTGC	TGGACAGGGG	CTCGGCTGTT	GGGCACTGAC
		AATTCCGTGG				
5	4021	TGTTGTCGGG	GAAATCATCG	TCCTTTCTTT	GGCTGCTCGC	CTGTGTTGCC
		ACCTGGATTC				
	4081	TGCGCGGGAC	GTCCTTCTGC	TACGTCCCTT	CGGCCCTCAA	TCCAGCGGAC
		CTTCCTTCCC				
10	4141	GCGGCCTGCT	GCCGGCTCTG	CGGCCTCTTC	CGCGTCTTCG	CCTTCGCCCT
		CAGACGAGTC				
	4201	GGATCTCCCT	TTGGGCCGCC	TCCCCGCAAG	CTTCGCACTT	TTTAAAAGAA
		AAGGGAGGAC				
	4261	TGGATGGGAT	TTATTACTCC	GATAGGACGC	TGGCTTGTA	CTCAGTCTCT
		TACTAGGAGA				
15	4321	CCAGCTTGAG	CCTGGGTGTT	CGCTGGTTAG	CCTAACCTGG	TTGGCCACCA
		GGGGTAAGGA				
	4381	CTCCTTGGCT	TAGAAAGCTA	ATAAACTTGC	CTGCATTAGA	GCTCTTACGC
		GTCCCGGGCT				
20	4441	CGAGATCCGC	ATCTCAATTA	GTCAGCAACC	ATAGTCCCGC	CCCTAACTCC
		GCCCATCCCG				
	4501	CCCCTAACTC	CGCCCACTTC	CGCCCATCTC	CCGCCCATG	GCTGACTAAT
		TTTTTTTATT				
	4561	TATGCAGAGG	CCGAGGCCGC	CTCGGCCTCT	GAGCTATTCC	AGAAGTAGTG
		AGGAGGCTTT				
25	4621	TTTGGAGGCC	TAGGCTTTTG	CAAAAAGCTA	ACTTGTTTAT	TGCAGCTTAT
		AATGGTTACA				
	4681	AATAAAGCAA	TAGCATCACA	AATTTACAAA	ATAAAGCATT	TTTTTCACTG
		CATTCTAGTT				
30	4741	GTGGTTTGTC	CAAACATCATC	AATGTATCTT	ATCATGTCTG	TCCGCTTCCT
		CGCTCACTGA				
	4801	CTCGCTGCGC	TCGGTCGTTC	GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA
		AGGCGGTAAT				
	4861	ACGGTTATCC	ACAGAATCAG	GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA
		AAGGCCAGCA				
35	4921	AAAGGCCAGG	AACCGTAAAA	AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC
		TCCGCCCCCG				
	4981	TGACGAGCAT	CACAAAAATC	GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA
		CAGGACTATA				
40	5041	AAGATACCAG	GCGTTTCCCC	CTGGAAGCTC	CCTCGTGCGC	TCTCCTGTTC
		CGACCCTGCC				
	5101	GCTTACCGGA	TACCTGTCCG	CCTTTCTCCC	TTCGGGAAGC	GTGGCGCTTT
		CTCATAGCTC				
	5161	ACGCTGTAGG	TATCTCAGTT	CGGTGTAGGT	CGTTCGCTCC	AAGCTGGGCT
		GTGTGCACGA				
45	5221	ACCCCCCGTT	CAGCCCGACC	GCTGCGCCTT	ATCCGGTAAC	TATCGTCTTG
		AGTCCAACCC				
	5281	GGTAAGACAC	GACTTATCGC	CACTGGCAGC	AGCCACTGGT	AACAGGATTA
		GCAGAGCGAG				
50	5341	GTATGTAGGC	GGTGCTACAG	AGTTCTTGAA	GTGGTGGCCT	AACTACGGCT
		ACACTAGAAG				
	5401	AACAGTATTT	GGTATCTGCG	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA
		GAGTTGGTAG				
	5461	CTCTTGATCC	GGCAAACAAA	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT
		GCAAGCAGCA				
55	5521	GATTACGCGC	AGAAAAAAG	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA
		CGGGGTCTGA				
	5581	CGCTCAGTGG	AACGAAAAC	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT
		CAAAAAGGAT				
60	5641	CTTCACCTAG	ATCCTTTTAA	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA
		GTATATATGA				

	5701	GTAAACTTGG	TCTGACAGTT	AGAAAAACTC	ATCGAGCATC	AAATGAAACT
		GCAATTTATT				
	5761	CATATCAGGA	TTATCAATAC	CATATTTTTG	AAAAAGCCGT	TTCTGTAATG
		AAGGAGAAAA				
5	5821	CTCACCGAGG	CAGTTCCATA	GGATGGCAAG	ATCCTGGTAT	CGGTCTGCGA
		TTCCGACTCG				
	5881	TCCAACATCA	ATACAACCTA	TTAATTTCCC	CTCGTCAAAA	ATAAGGTTAT
		CAAGTGAGAA				
10	5941	ATCACCATGA	GTGACGACTG	AATCCGGTGA	GAATGGCAAC	AGCTTATGCA
		TTTCTTTCCA				
	6001	GACTTGTTCA	ACAGGCCAGC	CATTACGCTC	GTCATCAAAA	TCACTCGCAT
		CAACCAAACC				
	6061	GTTATTCATT	CGTGATTGCG	CCTGAGCGAG	ACGAAATACG	CGATCGCTGT
		TAAAAGGACA				
15	6121	ATTACAAACA	GGAATCGAAT	GCAACCGGCG	CAGGAACACT	GCCAGCGCAT
		CAACAATATT				
	6181	TTCACCTGAA	TCAGGATATT	CTTCTAATAC	CTGGAATGCT	GTTTTTCCGG
		GGATCGCAGT				
20	6241	GGTGAGTAAC	CATGCATCAT	CAGGAGTACG	GATAAAATGC	TTGATGGTCG
		GAAGAGGCAT				
	6301	AAATTCCGTC	AGCCAGTTTA	GTCTGACCAT	CTCATCTGTA	ACATCATTGG
		CAACGCTACC				
	6361	TTTGCCATGT	TTCAGAAACA	ACTCTGGCGC	ATCGGGCTTC	CCATACAATC
		GATAGATTGT				
25	6421	CGCACCTGAT	TGCCCGACAT	TATCGCGAGC	CCATTTATAC	CCATATAAAT
		CAGCATCCAT				
	6481	GTTGGAATTT	AATCGCGGCC	TAGAGCAAGA	CGTTTCCCGT	TGAATATGGC
		TCATAACACC				
30	6541	CCTTGTATTA	CTGTTTATGT	AAGCAGACAG	TTTTATTGTT	CATGATGATA
		TATTTTTATC				
	6601	TTGTGCAATG	TAACATCAGA	GATTTTGAGA	CACAACAATT	GGTCGACGGA
		SEQ ID NO:11				
	1	GGTACCTCAA	TATTGGCCAT	TAGCCATATT	ATTCATTGGT	TATATAGCAT
		AAATCAATAT				
35	61	TGGCTATTGG	CCATTGCATA	CGTTGTATCT	ATATCATAAT	ATGTACATTT
		ATATTGGCTC				
	121	ATGTCCAATA	TGACCGCCAT	GTTGGCATTG	ATTATTGACT	AGTTATTAAT
		AGTAATCAAT				
	181	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCCGC	GTTACATAAC
		TTACGGTAAA				
40	241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA
		TGACGTATGT				
	301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT
		ATTTACGGTA				
	361	AACTGCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTCCGCCCC
		CTATTGACGT				
45	421	CAATGACGGT	AAATGGCCCG	CCTGGCATTG	TGCCCAGTAC	ATGACCTTAC
		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
50	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
		TCCACCCCAT				
	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
55	721	TATATAAGCA	GAGCTCGCTG	GCTTGTA ACT	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTTCG	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
60	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCTT
		CATTGACGCC				

	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
5	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTAGGAG
		AGGCCGTAGC				
10	1141	CGTAACTACT	CTGGGCAAGT	AGGGCAGGCG	GTGGGTACGC	AATGGGGGCG
		GCTACCTCAG				
	1201	CACTAAATAG	GAGACAATTA	GACCAATTTG	AGAAAATACG	ACTTCGCCCCG
		AACGGAAAGA				
	1261	AAAAGTACCA	AATTAAACAT	TTAATATGGG	CAGGCAAGGA	GATGGAGCGC
		TTCGGCCTCC				
15	1321	ATGAGAGGTT	GTTGGAGACA	GAGGAGGGGT	GTAAAAGAAT	CATAGAAGTC
		CTCTACCCCC				
	1381	TAGAACCAAC	AGGATCGGAG	GGCTTAAAAA	GTCTGTTCAA	TCTTGTGTGC
		GTGCTATATT				
	1441	GCTTGCACAA	GGAACAGAAA	GTGAAAGACA	CAGAGGAAGC	AGTAGCAACA
20		GTAAGACAAC				
	1501	ACTGCCATCT	AGTGGAAGAAA	GAAAAAAGTG	CAACAGAGAC	ATCTAGTGGA
		CAAAAGAAAA				
	1561	ATGACAAGGG	AATAGCAGCG	CCACCTGGTG	GCAGTCAGAA	TTTTCCAGCG
		CAACAACAAG				
25	1621	GAAATGCCTG	GGTACATGTA	CCCTTGTCAC	CGCGCACCTT	AAATGCGTGG
		GTAAAAGCAG				
	1681	TAGAGGAGAA	AAAATTTGGA	GCAGAAATAG	TACCCATGTT	TCAAGCCCTA
		TCGAATTCCC				
	1741	GTTTGTGCTA	GGGTTCTTAG	GCTTCTTGGG	GGCTGCTGGA	ACTGCAATGG
30		GAGCAGCGGC				
	1801	GACAGCCCTG	ACGGTCCAGT	CTCAGCATTT	GCTTGCTGGG	ATACTGCAGC
		AGCAGAAGAA				
	1861	TCTGCTGGCG	GCTGTGGAGG	CTCAACAGCA	GATGTTGAAG	CTGACCATTT
		GGGGTGTTAA				
35	1921	AAACCTCAAT	GCCCGCGTCA	CAGCCCTTGA	GAAGTACCTA	GAGGATCAGG
		CACGACTAAA				
	1981	CTCCTGGGGG	TGCGCATGGA	AACAAGTATG	TCATACCACA	GTGGAGTGGC
		CCTGGACAAA				
	2041	TCGGAAGTCCG	GATTGGCAAA	ATATGACTTG	GTTGGAGTGG	GAAAGACAAA
40		TAGCTGATTT				
	2101	GGAAAGCAAC	ATTACGAGAC	AATTAGTGAA	GGCTAGAGAA	CAAGAGGAAA
		AGAATCTAGA				
	2161	TGCCTATCAG	AAGTTAACTA	GTTGGTCAGA	TTTCTGGTCT	TGGTTCGATT
		TCTCAAAATG				
45	2221	GCTTAACATT	TTAAAAATGG	GATTTTTAGT	AATAGTAGGA	ATAATAGGGT
		TAAGATTACT				
	2281	TTACACAGTA	TATGGATGTA	TAGTGAGGGT	TAGGCAGGGA	TATGTTCCCTC
		TATCTCCACA				
	2341	GATCCATATC	CGCGGCAATT	TTAAAAGAAA	GGGAGGAATA	GGGGGACAGA
50		CTTCAGCAGA				
	2401	GAGACTAATT	AATATAATAA	CAACACAATT	AGAAATACAA	CATTTACAAA
		CCAAAATTCA				
	2461	AAAAATTTTA	AATTTTAGAG	CCGCGGAGAT	CTCAATATTG	GCCATTAGCC
		ATATTATTCA				
55	2521	TTGGTTATAT	AGCATAAATC	AATATTGGCT	ATTGGCCATT	GCATACGTTG
		TATCTATATC				
	2581	ATAATATGTA	CATTTATATT	GGCTCATGTC	CAATATGACC	GCCATGTTGG
		CATTGATTAT				
	2641	TGACTAGTTA	TTAATAGTAA	TCAATTACGG	GGTCATTAGT	TCATAGCCCA
60		TATATGGAGT				

	2701	TCCGCGTTAC	ATAACTTACG	GTAAATGGCC	CGCCTGGCTG	ACCGCCCAAC
		GACCCCCGCC				
	2761	CATTGACGTC	AATAATGACG	TATGTTCCCA	TAGTAACGCC	AATAGGGACT
		TTCCATTGAC				
5	2821	GTCAATGGGT	GGAGTATTTA	CGGTAAACTG	CCCACCTGGC	AGTACATCAA
		GTGTATCATA				
	2881	TGCCAAGTCC	GCCCCCTATT	GACGTCAATG	ACGGTAAATG	GCCCGCCTGG
		CATTATGCCC				
10	2941	AGTACATGAC	CTTACGGGAC	TTTCCTACTT	GGCAGTACAT	CTACGTATTA
		GTCATCGCTA				
	3001	TTACCATGGT	GATGCGGTTT	TGGCAGTACA	CCAATGGGCG	TGGATAGCGG
		TTTGACTCAC				
	3061	GGGGATTTCC	AAGTCTCCAC	CCCATTGACG	TCAATGGGAG	TTTGTTTTGG
		CACCAAAATC				
15	3121	AACGGGACTT	TCCAAAATGT	CGTAATAACC	CCGCCCCGTT	GACGCAAATG
		GGCGGTAGGC				
	3181	GTGTACGGTG	GGAGGTCTAT	ATAAGCAGAG	CTCGTTTAGT	GAACCGTCAG
		ATCACTAGAA				
20	3241	GCTTTATTGC	GGTAGTTTAT	CACAGTTAAA	TTGCTAACGC	AGTCAGTGCT
		TCTGACACAA				
	3301	CAGTCTCGAA	CTTAAGCTGC	AGAAGTTGGT	CGTGAGGCAC	TGGGCAGGCT
		AGCCACCAAT				
	3361	GCAGATTGAG	CTGAGCACCT	GCTTCTTCCT	GTGCCTGCTG	AGGTTCTGCT
		TCTCTGCCAC				
25	3421	CAGGAGATAC	TACCTGGGGG	CTGTGGAGCT	GAGCTGGGAC	TACATGCAGT
		CTGACCTGGG				
	3481	GGAGCTGCCT	GTGGATGCCA	GGTCCCCCCC	CAGAGTGCCC	AAGAGCTTCC
		CCTTCAACAC				
30	3541	CTCTGTGGTG	TACAAGAAGA	CCCTGTTTGT	GGAGTTCACT	GACCACCTGT
		TCAACATTGC				
	3601	CAAGCCCAGG	CCCCCCTGGA	TGGGCCTGCT	GGGCCCCACC	ATCCAGGCTG
		AGGTGTATGA				
	3661	CACTGTGGTG	ATCACCCCTGA	AGAACATGGC	CAGCCACCCT	GTGAGCCTGC
		ATGCTGTGGG				
35	3721	GGTGAGCTAC	TGGAAGGCCT	CTGAGGGGGC	TGAGTATGAT	GACCAGACCA
		GCCAGAGGGA				
	3781	GAAGGAGGAT	GACAAGGTGT	TCCCTGGGGG	CAGCCACACC	TATGTGTGGC
		AGGTGCTGAA				
40	3841	GGAGAAATGGC	CCCATGGCCT	CTGACCCCCT	GTGCCTGACC	TACAGCTACC
		TGAGCCATGT				
	3901	GGACCTGGTG	AAGGACCTGA	ACTCTGGCCT	GATTGGGGCC	CTGCTGGTGT
		GCAGGGAGGG				
	3961	CAGCCTGGCC	AAGGAGAAGA	CCCAGACCCT	GCACAAGTTC	ATCCTGCTGT
		TTGCTGTGTT				
45	4021	TGATGAGGGC	AAGAGCTGGC	ACTCTGAAAC	CAAGAACAGC	CTGATGCAGG
		ACAGGGATGC				
	4081	TGCCTCTGCC	AGGGCCTGGC	CCAAGATGCA	CACTGTGAAT	GGCTATGTGA
		ACAGGAGCCT				
50	4141	GCCTGGCCTG	ATTGGCTGCC	ACAGGAAGTC	TGTGTACTGG	CATGTGATTG
		GCATGGGCAC				
	4201	CACCCCTGAG	GTGCACAGCA	TCTTCCTGGA	GGGCCACACC	TTCCTGGTCA
		GGAACCACAG				
	4261	GCAGGCCAGC	CTGGAGATCA	GCCCCATCAC	CTTCCTGACT	GCCCAGACCC
		TGCTGATGGA				
55	4321	CCTGGGCCAG	TTCCTGCTGT	TCTGCCACAT	CAGCAGCCAC	CAGCATGATG
		GCATGGAGGC				
	4381	CTATGTGAAG	GTGGACAGCT	GCCCTGAGGA	GCCCCAGCTG	AGGATGAAGA
		ACAATGAGGA				
60	4441	GGCTGAGGAC	TATGATGATG	ACCTGACTGA	CTCTGAGATG	GATGTGGTGA
		GGTTTGATGA				

	4501	TGACAACAGC	CCCAGCTTCA	TCCAGATCAG	GTCTGTGGCC	AAGAAGCACC
		CCAAGACCTG				
	4561	GGTGCACTAC	ATTGCTGCTG	AGGAGGAGGA	CTGGGACTAT	GCCCCCCTGG
		TGCTGGCCCC				
5	4621	TGATGACAGG	AGCTACAAGA	GCCAGTACCT	GAACAATGGC	CCCCAGAGGA
		TTGGCAGGAA				
	4681	GTACAAGAAG	GTCAGGTTCA	TGGCCTACAC	TGATGAAACC	TTCAAGACCA
		GGGAGGCCAT				
10	4741	CCAGCATGAG	TCTGGCATCC	TGGGCCCCCT	GCTGTATGGG	GAGGTGGGGG
		ACACCCTGCT				
	4801	GATCATCTTC	AAGAACCAGG	CCAGCAGGCC	CTACAACATC	TACCCCCATG
		GCATCACTGA				
	4861	TGTGAGGCCC	CTGTACAGCA	GGAGGCTGCC	CAAGGGGGTG	AAGCACCTGA
		AGGACTTCCC				
15	4921	CATCCTGCCT	GGGGAGATCT	TCAAGTACAA	GTGGACTGTG	ACTGTGGAGG
		ATGGCCCCAC				
	4981	CAAGTCTGAC	CCCAGGTGCC	TGACCAGATA	CTACAGCAGC	TTTGTGAACA
		TGGAGAGGGA				
20	5041	CCTGGCCTCT	GGCCTGATTG	GCCCCCTGCT	GATCTGCTAC	AAGGAGTCTG
		TGGACCAGAG				
	5101	GGGCAACCAG	ATCATGTCTG	ACAAGAGGAA	TGTGATCCTG	TTCTCTGTGT
		TTGATGAGAA				
	5161	CAGGAGCTGG	TACCTGACTG	AGAACATCCA	GAGGTTCTCTG	CCCAACCCTG
		CTGGGGTGCA				
25	5221	GCTGGAGGAC	CCTGAGTTCC	AGGCCAGCAA	CATCATGCAC	AGCATCAATG
		GCTATGTGTT				
	5281	TGACAGCCTG	CAGCTGTCTG	TGTGCCTGCA	TGAGGTGGCC	TACTGGTACA
		TCCTGAGCAT				
30	5341	TGGGGCCCCAG	ACTGACTTCC	TGTCTGTGTT	CTTCTCTGGC	TACACCTTCA
		AGCACAAGAT				
	5401	GGTGTATGAG	GACACCCTGA	CCCTGTTCCC	CTTCTCTGGG	GAGACTGTGT
		TCATGAGCAT				
	5461	GGAGAACCCT	GGCCTGTGGA	TTCTGGGCTG	CCACAACCTCT	GACTTCAGGA
		ACAGGGGCAT				
35	5521	GACTGCCCTG	CTGAAAGTCT	CCAGCTGTGA	CAAGAACACT	GGGGACTACT
		ATGAGGACAG				
	5581	CTATGAGGAC	ATCTCTGCCT	ACCTGCTGAG	CAAGAACAAT	GCCATTGAGC
		CCAGGAGCTT				
40	5641	CAGCCAGAAT	GCCACTAATG	TGTCTAACAA	CAGCAACACC	AGCAATGACA
		GCAATGTGTC				
	5701	TCCCCCAGTG	CTGAAGAGGC	ACCAGAGGGA	GATCACCAGG	ACCACCCTGC
		AGTCTGACCA				
	5761	GGAGGAGATT	GACTATGATG	ACACCATCTC	TGTGGAGATG	AAGAAGGAGG
		ACTTTGACAT				
45	5821	CTACGACGAG	GACGAGAACC	AGAGCCCCAG	GAGCTTCCAG	AAGAAGACCA
		GGCACTACTT				
	5881	CATTGCTGCT	GTGGAGAGGC	TGTGGGACTA	TGGCATGAGC	AGCAGCCCCC
		ATGTGCTGAG				
50	5941	GAACAGGGCC	CAGTCTGGCT	CTGTGCCCCA	GTTCAAGAAG	GTGGTGTTCC
		AGGAGTTCAC				
	6001	TGATGGCAGC	TTCACCCAGC	CCCTGTACAG	AGGGGAGCTG	AATGAGCACC
		TGGGCCTGCT				
	6061	GGGCCCCTAC	ATCAGGGCTG	AGGTGGAGGA	CAACATCATG	GTGACCTTCA
		GGAACCAGGC				
55	6121	CAGCAGGCCC	TACAGCTTCT	ACAGCAGCCT	GATCAGCTAT	GAGGAGGACC
		AGAGGCAGGG				
	6181	GGCTGAGCCC	AGGAAGAACT	TTGTGAAGCC	CAATGAAACC	AAGACCTACT
		TCTGGAAGGT				
60	6241	GCAGCACCAC	ATGGCCCCCA	CCAAGGATGA	GTTTGACTGC	AAGGCCTGGG
		CCTACTTCTC				

	6301	TGATGTGGAC	CTGGAGAAGG	ATGTGCACTC	TGGCCTGATT	GGCCCCCTGC
		TGGTGTGCCA				
	6361	CACCAACACC	CTGAACCCTG	CCCATGGCAG	GCAGGTGACT	GTGCAGGAGT
		TTGCCCTGTT				
5	6421	CTTCACCATC	TTTGATGAAA	CCAAGAGCTG	GTACTTCACT	GAGAACATGG
		AGAGGAACTG				
	6481	CAGGGCCCCC	TGCAACATCC	AGATGGAGGA	CCCCACCTTC	AAGGAGAACT
		ACAGGTTCCA				
	6541	TGCCATCAAT	GGCTACATCA	TGGACACCCT	GCCTGGCCTG	GTGATGGCCC
10		AGGACCAGAG				
	6601	GATCAGGTGG	TACCTGCTGA	GCATGGGCAG	CAATGAGAAC	ATCCACAGCA
		TCCACTTCTC				
	6661	TGGCCATGTG	TTCACTGTGA	GGAAGAAGGA	GGAGTACAAG	ATGGCCCTGT
		ACAACCTGTA				
15	6721	CCCTGGGGTG	TTTGAGACTG	TGGAGATGCT	GCCCAGCAAG	GCTGGCATCT
		GGAGGGTGGA				
	6781	GTGCCTGATT	GGGGAGCACC	TGCATGCTGG	CATGAGCACC	CTGTTCCTGG
		TGTACAGCAA				
	6841	CAAGTGCCAG	ACCCCCCTGG	GCATGGCCTC	TGGCCACATC	AGGGACTTCC
20		AGATCACTGC				
	6901	CTCTGGCCAG	TATGGCCAGT	GGGCCCCCAA	GCTGGCCAGG	CTGCACTACT
		CTGGCAGCAT				
	6961	CAATGCCTGG	AGCACCAAGG	AGCCCTTCAG	CTGGATCAAG	GTGGACCTGC
		TGGCCCCCAT				
25	7021	GATCATCCAT	GGCATCAAGA	CCCAGGGGGC	CAGGCAGAAG	TTCAGCAGCC
		TGTACATCAG				
	7081	CCAGTTCATC	ATCATGTACA	GCCTGGATGG	CAAGAAGTGG	CAGACCTACA
		GGGGCAACAG				
	7141	CACTGGCACC	CTGATGGTGT	TCTTTGGCAA	TGTGGACAGC	TCTGGCATCA
30		AGCACAAACAT				
	7201	CTTCAACCCC	CCCATCATTG	CCAGATACAT	CAGGCTGCAC	CCCACCCACT
		ACAGCATCAG				
	7261	GAGCACCTG	AGGATGGAGC	TGATGGGCTG	TGACCTGAAC	AGCTGCAGCA
		TGCCCCTGGG				
35	7321	CATGGAGAGC	AAGGCCATCT	CTGATGCCCA	GATCACTGCC	AGCAGCTACT
		TCACCAACAT				
	7381	GTTTGCCACC	TGGAGCCCCA	GCAAGGCCAG	GCTGCACCTG	CAGGGCAGGA
		GCAATGCCTG				
	7441	GAGGCCCCAG	GTCAACAACC	CCAAGGAGTG	GCTGCAGGTG	GACTTCCAGA
40		AGACCATGAA				
	7501	GGTGAAGGTG	GTGACCACCC	AGGGGGTGAA	GAGCCTGCTG	ACCAGCATGT
		ATGTGAAGGA				
	7561	GTTCTGATC	AGCAGCAGCC	AGGATGGCCA	CCAGTGGACC	CTGTTCTTCC
		AGAATGGCAA				
45	7621	GGTGAAGGTG	TTCCAGGGCA	ACCAGGACAG	CTTCACCCCT	GTGGTGAACA
		GCCTGGACCC				
	7681	CCCCCTGCTG	ACCAGATACC	TGAGGATTCA	CCCCCAGAGC	TGGGTGCACC
		AGATTGCCCT				
	7741	GAGGATGGAG	GTGCTGGGCT	GTGAGGCCCA	GGACCTGTAC	TGAGCGGCCG
50		CGGGCCCAAT				
	7801	CAACCTCTGG	ATTACAAAAT	TTGTGAAAGA	TTGACTGGTA	TTCTTAACTA
		TGTTGCTCCT				
	7861	TTTACGCTAT	GTGGATACGC	TGCTTTAATG	CCTTTGTATC	ATGCTATTGC
		TTCCCGTATG				
55	7921	GCTTTCATTT	TCTCCTCCTT	GTATAAATCC	TGGTTGCTGT	CTCTTTATGA
		GGAGTTGTGG				
	7981	CCC GTTGTC A	GGCAACGTGG	CGTGGTGTGC	ACTGTGTTTG	CTGACGCAAC
		CCCCACTGGT				
60	8041	TGGGGCATTG	CCACCACCTG	TCAGCTCCTT	TCCGGGACTT	TCGCTTTCCC
		CCTCCCTATT				

	8101	GCCACGGCGG	AACTCATCGC	CGCCTGCCTT	GCCCGCTGCT	GGACAGGGGC
		TCGGCTGTTG				
	8161	GGCACTGACA	ATTCCGTGGT	GTTGTCGGGG	AAATCATCGT	CCTTTCCTTG
		GCTGCTCGCC				
5	8221	TGTGTTGCCA	CCTGGATTCT	GCGCGGGACG	TCCTTCTGCT	ACGTCCCTTC
		GGCCCTCAAT				
	8281	CCAGCGGACC	TTCCTTCCCG	CGGCCTGCTG	CCGGCTCTGC	GGCCTCTTCC
		GCGTCTTCGC				
10	8341	CTTCGCCCTC	AGACGAGTCG	GATCTCCCTT	TGGGCCGCCT	CCCCGCAAGC
		TTCGCACTTT				
	8401	TTAAAAGAAA	AGGGAGGACT	GGATGGGATT	TATTACTCCG	ATAGGACGCT
		GGCTTGTAAC				
	8461	TCAGTCTCTT	ACTAGGAGAC	CAGCTTGAGC	CTGGGTGTTC	GCTGGTTAGC
		CTAACCTGGT				
15	8521	TGGCCACCAG	GGGTAAGGAC	TCCTTGGCTT	AGAAAGCTAA	TAAACTTGCC
		TGCATTAGAG				
	8581	CTCTTACGCG	TCCCGGGCTC	GAGATCCGCA	TCTCAATTAG	TCAGCAACCA
		TAGTCCCGCC				
	8641	CCTAACTCCG	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC	GCCCATTCTC
20		CGCCCCATGG				
	8701	CTGACTAATT	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC	TCGGCCTCTG
		AGCTATTCCA				
	8761	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC	AAAAAGCTAA
		CTTGTTTATT				
25	8821	GCAGCTTATA	ATGGTTACAA	ATAAAGCAAT	AGCATCACAA	ATTCACAAA
		TAAAGCATTT				
	8881	TTTTCACTGC	ATTCTAGTTG	TGGTTTGTCC	AAACTCATCA	ATGTATCTTA
		TCATGTCTGT				
	8941	CCGCTTCCTC	GCTCACTGAC	TCGCTGCGCT	CGGTCGTTCC	GCTGCGGCGA
30		GCGGTATCAG				
	9001	CTCACTCAAA	GGCGGTAATA	CGGTTATCCA	CAGAATCAGG	GGATAACGCA
		GGAAAGAACA				
	9061	TGTGAGCAAA	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA	GGCCGCGTTG
		CTGGCGTTTT				
35	9121	TCCATAGGCT	CCGCCCCCCT	GACGAGCATC	ACAAAAATCG	ACGCTCAAGT
		CAGAGGTGGC				
	9181	GAAACCCGAC	AGGACTATAA	AGATACCAGG	CGTTTCCCCC	TGGAAGCTCC
		CTCGTGCGCT				
	9241	CTCCTGTTCC	GACCCTGCCG	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT
40		TCGGGAAGCG				
	9301	TGGCGCTTTC	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTC	GGTGTAGGTC
		GTTGCTCCA				
	9361	AGCTGGGCTG	TGTGCACGAA	CCCCCGTTTC	AGCCCGACCG	CTGCGCCTTA
		TCCGGTAACT				
45	9421	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGCC	ACTGGCAGCA
		GCCACTGGTA				
	9481	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG	GTGCTACAGA	GTTCTTGAAG
		TGGTGGCCTA				
	9541	ACTACGGCTA	CACTAGAAGA	ACAGTATTTG	GTATCTGCGC	TCTGCTGAAG
50		CCAGTTACCT				
	9601	TCGGA AAAAG	AGTTGGTAGC	TCTTGATCCG	GCAAAACAAAC	CACCGCTGGT
		AGCGGTGGTT				
	9661	TTTTTGTTTG	CAAGCAGCAG	ATTACGCGCA	GAAAAAAAGG	ATCTCAAGAA
		GATCCTTTGA				
55	9721	TCTTTTCTAC	GGGGTCTGAC	GCTCAGTGGA	ACGAAAACCTC	ACGTTAAGGG
		ATTTTGGTCA				
	9781	TGAGATTATC	AAAAAGGATC	TTCACCTAGA	TCCTTTTAAA	TTAAAAATGA
		AGTTTTAAAT				
	9841	CAATCTAAAG	TATATATGAG	TAAACTTGGT	CTGACAGTTA	GAAAAACTCA
60		TCGAGCATCA				

	9901	AATGAAACTG	CAATTTATTC	ATATCAGGAT	TATCAATACC	ATATTTTTGA
		AAAAGCCGTT				
	9961	TCTGTAATGA	AGGAGAAAAC	TCACCGAGGC	AGTTCCATAG	GATGGCAAGA
		TCCTGGTATC				
5	10021	GGTCTGCGAT	TCCGACTCGT	CCAACATCAA	TACAACCTAT	TAATTTCCCC
		TCGTCAAAAA				
	10081	TAAGGTTATC	AAGTGAGAAA	TCACCATGAG	TGACGACTGA	ATCCGGTGAG
		AATGGCAACA				
	10141	GCTTATGCAT	TTCTTTCCAG	ACTTGTTCAA	CAGGCCAGCC	ATTACGCTCG
10		TCATCAAAAT				
	10201	CACTCGCATC	AACCAAACCG	TTATTCATTC	GTGATTGCGC	CTGAGCGAGA
		CGAAATACGC				
	10261	GATCGCTGTT	AAAAGGACAA	TTACAAACAG	GAATCGAATG	CAACCGGCGC
		AGGAACACTG				
15	10321	CCAGCGCATC	AACAATATTT	TCACCTGAAT	CAGGATATTC	TTCTAATACC
		TGGAATGCTG				
	10381	TTTTTCCGGG	GATCGCAGTG	GTGAGTAACC	ATGCATCATC	AGGAGTACGG
		ATAAAATGCT				
	10441	TGATGGTCGG	AAGAGGCATA	AATTCCGTCA	GCCAGTTTAG	TCTGACCATC
20		TCATCTGTAA				
	10501	CATCATTGGC	AACGCTACCT	TTGCCATGTT	TCAGAAACAA	CTCTGGCGCA
		TCGGGCTTCC				
	10561	CATACAATCG	ATAGATTGTC	GCACCTGATT	GCCCGACATT	ATCGCGAGCC
		CATTTATACC				
25	10621	CATATAAATC	AGCATCCATG	TTGGAATTTA	ATCGCGGCCT	AGAGCAAGAC
		GTTTCCCGTT				
	10681	GAATATGGCT	CATAACACCC	CTTGTATTAC	TGTTTATGTA	AGCAGACAGT
		TTTATTGTTC				
	10741	ATGATGATAT	ATTTTTATCT	TGTGCAATGT	AACATCAGAG	ATTTTGAGAC
30		ACAACAATTG				
	10801	GTCGACGGAT CC				
		SEQ ID NO:12				
	1	GGTACCTCAA	TATTGGCCAT	TAGCCATATT	ATTCATTGGT	TATATAGCAT
	61	TGGCTATTGG	CCATTGCATA	CGTTGTATCT	ATATCATAAT	ATGTACATTT
35		ATATTGGCTC				
	121	ATGTCCAATA	TGACCGCCAT	GTTGGCATTG	ATTATTGACT	AGTTATTAAT
		AGTAATCAAT				
	181	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCCGC	GTTACATAAC
		TTACGGTAAA				
	241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA
40		TGACGTATGT				
	301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT
		ATTTACGGTA				
	361	AACTGCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTCCGCCCC
		CTATTGACGT				
45	421	CAATGACGGT	AAATGGCCCG	CCTGGCATTG	TGCCCAGTAC	ATGACCTTAC
		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
50		TCCACCCCAT				
	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
55	721	TATATAAGCA	GAGCTCGCTG	GCTTGTAAC	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTTC	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCTT
60		CATTGACGCC				

	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
5	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTIONAGAG
		AGGCCGTAGC				
10	1141	CGTAACTACT	CTGGGCAAGT	AGGGCAGGCG	GTGGGTACGC	AATGGGGGCG
		GCTACCTCAG				
	1201	CACTAAATAG	GAGACAATTA	GACCAATTTG	AGAAAATACG	ACTTCGCCCCG
		AACGGAAAGA				
	1261	AAAAGTACCA	AATTAAACAT	TTAATATGGG	CAGGCAAGGA	GATGGAGCGC
		TTCGGCCTCC				
15	1321	ATGAGAGGTT	GTTGGAGACA	GAGGAGGGGT	GTAAAAGAAT	CATAGAAGTC
		CTCTACCCCC				
	1381	TAGAACCAAC	AGGATCGGAG	GGCTTAAAAA	GTCTGTTCAA	TCTTGTGTGC
		GTGCTATATT				
20	1441	GCTTGACAAA	GGAACAGAAA	GTGAAAGACA	CAGAGGAAGC	AGTAGCAACA
		GTAAGACAAC				
	1501	ACTGCCATCT	AGTGGAAGAAA	GAAAAAAGTG	CAACAGAGAC	ATCTAGTGGA
		CAAAAGAAAA				
	1561	ATGACAAGGG	AATAGCAGCG	CCACCTGGTG	GCAGTCAGAA	TTTTCCAGCG
		CAACAACAAG				
25	1621	GAAATGCCTG	GGTACATGTA	CCCTTGTAC	CGCGCACCTT	AAATGCGTGG
		GTAAAAGCAG				
	1681	TAGAGGAGAA	AAAATTTGGA	GCAGAAATAG	TACCCATGTT	TCAAGCCCTA
		TCGAATTCCC				
30	1741	GTTTGTGCTA	GGGTTCTTAG	GCTTCTTGGG	GGCTGCTGGA	ACTGCAATGG
		GAGCAGCGGC				
	1801	GACAGCCCTG	ACGGTCCAGT	CTCAGCATTT	GCTTGCTGGG	ATACTGCAGC
		AGCAGAAGAA				
	1861	TCTGCTGGCG	GCTGTGGAGG	CTCAACAGCA	GATGTTGAAG	CTGACCATTT
		GGGGTGTTAA				
35	1921	AAACCTCAAT	GCCCGCGTCA	CAGCCCTTGA	GAAGTACCTA	GAGGATCAGG
		CACGACTAAA				
	1981	CTCCTGGGGG	TGCGCATGGA	AACAAGTATG	TCATACCACA	GTGGAGTGGC
		CCTGGACAAA				
40	2041	TCGGACTIONCCG	GATTGGCAAA	ATATGACTTG	GTTGGAGTGG	GAAAGACAAA
		TAGCTGATTT				
	2101	GGAAAGCAAC	ATTACGAGAC	AATTAGTGAA	GGCTAGAGAA	CAAGAGGAAA
		AGAATCTAGA				
	2161	TGCCTATCAG	AAGTTAACTA	GTTGGTCAGA	TTTCTGGTCT	TGGTTGATT
		TCTCAAAATG				
45	2221	GCTTAACATT	TTAAAAATGG	GATTTTTAGT	AATAGTAGGA	ATAATAGGGT
		TAAGATTACT				
	2281	TTACACAGTA	TATGGATGTA	TAGTGAGGGT	TAGGCAGGGA	TATGTTCTCTC
		TATCTCCACA				
50	2341	GATCCATATC	CGCGGCAATT	TTAAAAGAAA	GGGAGGAATA	GGGGGACAGA
		CTTCAGCAGA				
	2401	GAGACTAATT	AATATAATAA	CAACACAATT	AGAAATACAA	CATTTACAAA
		CCAAAATTCA				
	2461	AAAAATTTTA	AATTTTAGAG	CCGCGGAGAT	CTGTTACATA	ACTTATGGTA
		AATGGCCTGC				
55	2521	CTGGCTGACT	GCCCAATGAC	CCCTGCCCAA	TGATGTCAAT	AATGATGTAT
		GTTCCCATGT				
	2581	AATGCCAATA	GGGACTTTCC	ATTGATGTCA	ATGGGTGGAG	TATTTATGGT
		AACTGCCCCAC				
60	2641	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTATGCCCC	CTATTGATGT
		CAATGATGGT				

	2701	AAATGGCCTG	CCTGGCATT	TGCCCAGTAC	ATGACCTTAT	GGGACTTTCC
		TACTTGGCAG				
	2761	TACATCTATG	TATTAGTCAT	TGCTATTACC	ATGGGAATTC	ACTAGTGGAG
		AAGAGCATGC				
5	2821	TTGAGGGCTG	AGTGCCCCTC	AGTGGGCAGA	GAGCACATGG	CCCACAGTCC
		CTGAGAAAGTT				
	2881	GGGGGGAGGG	GTGGGCAATT	GAAGTGGTGC	CTAGAGAAGG	TGGGGCTTGG
		GTAAACTGGG				
	2941	AAAGTGATGT	GGTGTACTGG	CTCCACCTTT	TTCCCCAGGG	TGGGGGAGAA
10		CCATATATAA				
	3001	GTGCAGTAGT	CTCTGTGAAC	ATTCAAGCTT	CTGCCTTCTC	CCTCCTGTGA
		GTTTGCTAGC				
	3061	CACCAATGCA	GATTGAGCTG	AGCACCTGCT	TCTTCCTGTG	CCTGCTGAGG
		TTCTGCTTCT				
15	3121	CTGCCACCAG	GAGATACTAC	CTGGGGGCTG	TGGAGCTGAG	CTGGGACTAC
		ATGCAGTCTG				
	3181	ACCTGGGGGA	GCTGCCTGTG	GATGCCAGGT	TCCCCCCCAG	AGTGCCCAAG
		AGCTTCCCCT				
	3241	TCAACACCTC	TGTGGTGTAC	AAGAAGACCC	TGTTTGTGGA	GTTCACTGAC
20		CACCTGTTCA				
	3301	ACATTGCCAA	GCCCAGGCC	CCCTGGATGG	GCCTGCTGGG	CCCCACCATC
		CAGGCTGAGG				
	3361	TGTATGACAC	TGTGGTGTATC	ACCCTGAAGA	ACATGGCCAG	CCACCCTGTG
		AGCCTGCATG				
25	3421	CTGTGGGGGT	GAGCTACTGG	AAGGCCTCTG	AGGGGGCTGA	GTATGATGAC
		CAGACCAGCC				
	3481	AGAGGGAGAA	GGAGGATGAC	AAGGTGTTCC	CTGGGGGCAG	CCACACCTAT
		GTGTGGCAGG				
	3541	TGCTGAAGGA	GAATGGCCCC	ATGGCCTCTG	ACCCCCTGTG	CCTGACCTAC
30		AGCTACCTGA				
	3601	GCCATGTGGA	CCTGGTGAAG	GACCTGAACT	CTGGCCTGAT	TGGGGCCCTG
		CTGGTGTGCA				
	3661	GGGAGGGCAG	CCTGGCCAAG	GAGAAGACCC	AGACCCTGCA	CAAGTTCATC
		CTGCTGTTTG				
35	3721	CTGTGTTTGA	TGAGGGCAAG	AGCTGGCACT	CTGAAACCAA	GAACAGCCTG
		ATGCAGGACA				
	3781	GGGATGCTGC	CTCTGCCAGG	GCCTGGCCCA	AGATGCACAC	TGTGAATGGC
		TATGTGAACA				
	3841	GGAGCCTGCC	TGGCCTGATT	GGCTGCCACA	GGAAGTCTGT	GTAAGTGGCAT
40		GTGATTGGCA				
	3901	TGGGCACCAC	CCCTGAGGTG	CACAGCATCT	TCCTGGAGGG	CCACACCTTC
		CTGGTCAGGA				
	3961	ACCACAGGCA	GGCCAGCCTG	GAGATCAGCC	CCATCACCTT	CCTGACTGCC
		CAGACCCTGC				
45	4021	TGATGGACCT	GGGCCAGTTC	CTGCTGTTCT	GCCACATCAG	CAGCCACCAG
		CATGATGGCA				
	4081	TGGAGGCCTA	TGTGAAGGTG	GACAGCTGCC	CTGAGGAGCC	CCAGCTGAGG
		ATGAAGAACA				
	4141	ATGAGGAGGC	TGAGGACTAT	GATGATGACC	TGACTGACTC	TGAGATGGAT
50		GTGGTGAGGT				
	4201	TTGATGATGA	CAACAGCCCC	AGCTTCATCC	AGATCAGGTC	TGTGGCCAAG
		AAGCACCCCA				
	4261	AGACCTGGGT	GCACTACATT	GCTGCTGAGG	AGGAGGACTG	GGACTATGCC
		CCCCTGGTGC				
55	4321	TGGCCCCTGA	TGACAGGAGC	TACAAGAGCC	AGTACCTGAA	CAATGGCCCC
		CAGAGGATTG				
	4381	GCAGGAAGTA	CAAGAAGGTC	AGGTTTCATGG	CCTACACTGA	TGAAACCTTC
		AAGACCAGGG				
	4441	AGGCCATCCA	GCATGAGTCT	GGCATCCTGG	GCCCCCTGCT	GTATGGGGAG
60		GTGGGGGACA				

	4501	CCCTGCTGAT	CATCTTCAAG	AACCAGGCCA	GCAGGCCCTA	CAACATCTAC
		CCCCATGGCA				
	4561	TCACTGATGT	GAGGCCCCTG	TACAGCAGGA	GGCTGCCCAA	GGGGGTGAAG
		CACCTGAAGG				
5	4621	ACTTCCCCAT	CCTGCCTGGG	GAGATCTTCA	AGTACAAGTG	GACTGTGACT
		GTGGAGGATG				
	4681	GCCCCACCAA	GTCTGACCCC	AGGTGCCTGA	CCAGATACTA	CAGCAGCTTT
		GTGAACATGG				
10	4741	AGAGGGACCT	GGCCTCTGGC	CTGATTGGCC	CCCTGCTGAT	CTGCTACAAG
		GAGTCTGTGG				
	4801	ACCAGAGGGG	CAACCAGATC	ATGTCTGACA	AGAGGAATGT	GATCCTGTTC
		TCTGTGTTTG				
	4861	ATAGAACAG	GAGCTGGTAC	CTGACTGAGA	ACATCCAGAG	GTTCCCTGCCC
		AACCTGCTG				
15	4921	GGGTGCAGCT	GGAGGACCCT	GAGTTCAGG	CCAGCAACAT	CATGCACAGC
		ATCAATGGCT				
	4981	ATGTGTTTGA	CAGCCTGCAG	CTGTCTGTGT	GCCTGCATGA	GGTGGCCTAC
		TGGTACATCC				
20	5041	TGAGCATTGG	GGCCCAGACT	GACTTCCTGT	CTGTGTTCTT	CTCTGGCTAC
		ACCTTCAAGC				
	5101	ACAAGATGGT	GTATGAGGAC	ACCCTGACCC	TGTTCCCCTT	CTCTGGGGAG
		ACTGTGTTCA				
	5161	TGAGCATGGA	GAACCCTGGC	CTGTGGATTC	TGGGCTGCCA	CAACTCTGAC
		TTCAGGAACA				
25	5221	GGGGCATGAC	TGCCCTGCTG	AAAGTCTCCA	GCTGTGACAA	GAACACTGGG
		GACTACTATG				
	5281	AGGACAGCTA	TGAGGACATC	TCTGCCTACC	TGCTGAGCAA	GAACAATGCC
		ATTGAGCCCA				
30	5341	GGAGCTTCAG	CCAGAATGCC	ACTAATGTGT	CTAACAACAG	CAACACCAGC
		AATGACAGCA				
	5401	ATGTGTCTCC	CCCAGTGCTG	AAGAGGCACC	AGAGGGAGAT	CACCAGGACC
		ACCCTGCAGT				
	5461	CTGACCAGGA	GGAGATTGAC	TATGATGACA	CCATCTCTGT	GGAGATGAAG
		AAGGAGGACT				
35	5521	TTGACATCTA	CGACGAGGAC	GAGAACCAGA	GCCCCAGGAG	CTTCCAGAAG
		AAGACCAGGC				
	5581	ACTACTTCAT	TGCTGCTGTG	GAGAGGCTGT	GGGACTATGG	CATGAGCAGC
		AGCCCCCATG				
40	5641	TGCTGAGGAA	CAGGGCCCAG	TCTGGCTCTG	TGCCCCAGTT	CAAGAAGGTG
		GTGTTCCAGG				
	5701	AGTTCACTGA	TGGCAGCTTC	ACCCAGCCCC	TGTACAGAGG	GGAGCTGAAT
		GAGCACCTGG				
	5761	GCCTGCTGGG	CCCCTACATC	AGGGCTGAGG	TGGAGGACAA	CATCATGGTG
		ACCTTCAGGA				
45	5821	ACCAGGCCAG	CAGGCCCTAC	AGCTTCTACA	GCAGCCTGAT	CAGCTATGAG
		GAGGACCAGA				
	5881	GGCAGGGGGC	TGAGCCCAGG	AAGAACTTTG	TGAAGCCCAA	TGAAACCAAG
		ACCTACTTCT				
	5941	GGAAGGTGCA	GCACCACATG	GCCCCACCA	AGGATGAGTT	TGACTGCAAG
50		GCCTGGGCCT				
	6001	ACTTCTCTGA	TGTGGACCTG	GAGAAGGATG	TGCACTCTGG	CCTGATTGGC
		CCCCTGCTGG				
	6061	TGTGCCACAC	CAACACCCTG	AACCCTGCCC	ATGGCAGGCA	GGTGACTGTG
		CAGGAGTTTG				
55	6121	CCCTGTTCTT	CACCATCTTT	GATGAAACCA	AGAGCTGGTA	CTTCACTGAG
		AACATGGAGA				
	6181	GGAAGTGCAG	GGCCCCCTGC	AACATCCAGA	TGGAGGACCC	CACCTTCAAG
		GAGAACTACA				
60	6241	GGTTCCATGC	CATCAATGGC	TACATCATGG	ACACCCTGCC	TGGCCTGGTG
		ATGGCCCAGG				

	6301	ACCAGAGGAT	CAGGTGGTAC	CTGCTGAGCA	TGGGCAGCAA	TGAGAACATC
		CACAGCATCC				
	6361	ACTTCTCTGG	CCATGTGTTC	ACTGTGAGGA	AGAAGGAGGA	GTACAAGATG
		GCCCTGTACA				
5	6421	ACCTGTACCC	TGGGGTGTTC	GAGACTGTGG	AGATGCTGCC	CAGCAAGGCT
		GGCATCTGGA				
	6481	GGGTGGAGTG	CCTGATTGGG	GAGCACCTGC	ATGCTGGCAT	GAGCACCTG
		TTCCTGGTGT				
10	6541	ACAGCAACAA	GTGCCAGACC	CCCCTGGGCA	TGGCCTCTGG	CCACATCAGG
		GACTTCCAGA				
	6601	TCACTGCCTC	TGGCCAGTAT	GGCCAGTGGG	CCCCCAAGCT	GGCCAGGCTG
		CACTACTCTG				
	6661	GCAGCATCAA	TGCCTGGAGC	ACCAAGGAGC	CCTTCAGCTG	GATCAAGGTG
		GACCTGCTGG				
15	6721	CCCCATGAT	CATCCATGGC	ATCAAGACCC	AGGGGGCCAG	GCAGAAGTTC
		AGCAGCCTGT				
	6781	ACATCAGCCA	GTTTCATCATC	ATGTACAGCC	TGGATGGCAA	GAAGTGGCAG
		ACCTACAGGG				
20	6841	GCAACAGCAC	TGGCACCTG	ATGGTGTTC	TTGGCAATGT	GGACAGCTCT
		GGCATCAAGC				
	6901	ACAACATCTT	CAACCCCCC	ATCATTGCCA	GATACATCAG	GCTGCACCCC
		ACCCACTACA				
	6961	GCATCAGGAG	CACCCTGAGG	ATGGAGCTGA	TGGGCTGTGA	CCTGAACAGC
		TGCAGCATGC				
25	7021	CCCTGGGCAT	GGAGAGCAAG	GCCATCTCTG	ATGCCCAGAT	CACTGCCAGC
		AGCTACTTCA				
	7081	CCAACATGTT	TGCCACCTGG	AGCCCCAGCA	AGGCCAGGCT	GCACCTGCAG
		GGCAGGAGCA				
30	7141	ATGCCTGGAG	GCCCCAGGTC	AACAACCCCA	AGGAGTGGCT	GCAGGTGGAC
		TTCCAGAAGA				
	7201	CCATGAAGGT	GACTGGGGTG	ACCACCCAGG	GGGTGAAGAG	CCTGCTGACC
		AGCATGTATG				
	7261	TGAAGGAGTT	CCTGATCAGC	AGCAGCCAGG	ATGGCCACCA	GTGGACCCTG
		TTCTTCCAGA				
35	7321	ATGGCAAGGT	GAAGGTGTTC	CAGGGCAACC	AGGACAGCTT	CACCCCTGTG
		GTGAACAGCC				
	7381	TGGACCCCCC	CCTGCTGACC	AGATACCTGA	GGATTCACCC	CCAGAGCTGG
		GTGCACCAGA				
40	7441	TTGCCCTGAG	GATGGAGGTG	CTGGGCTGTG	AGGCCCAGGA	CCTGTACTGA
		GCGGCCGCGG				
	7501	GCCCAATCAA	CCTCTGGATT	ACAAAATTTG	TGAAAGATTG	ACTGGTATTC
		TTAACTATGT				
	7561	TGCTCCTTTT	ACGCTATGTG	GATACGCTGC	TTTAATGCCT	TTGTATCATG
		CTATTGCTTC				
45	7621	CCGTATGGCT	TTCATTTTCT	CCTCCTTGTA	TAAATCCTGG	TTGCTGTCTC
		TTTATGAGGA				
	7681	GTTGTGGCCC	GTTGTCAGGC	AACGTGGCGT	GGTGTGCACT	GTGTTTGCTG
		ACGCAACCCC				
50	7741	CACTGGTTGG	GGCATTGCCA	CCACCTGTCA	GCTCCTTTCC	GGGACTTTCC
		CTTTCCCCCT				
	7801	CCCTATTGCC	ACGGCGGAAC	TCATCGCCGC	CTGCCTTGCC	CGCTGCTGGA
		CAGGGGCTCG				
	7861	GCTGTTGGGC	ACTGACAATT	CCGTGGTGTT	GTCGGGGAAA	TCATCGTCCT
		TTCCTTGGCT				
55	7921	GCTCGCCTGT	GTTGCCACCT	GGATTCTGCG	CGGGACGTCC	TTCTGCTACG
		TCCCTTCGGC				
	7981	CCTCAATCCA	GCGGACCTTC	CTTCCCGCGG	CCTGCTGCCG	GCTCTGCGGC
		CTCTTCCGCG				
60	8041	TCTTCGCCTT	CGCCCTCAGA	CGAGTCGGAT	CTCCCTTTGG	GCCGCCTCCC
		CGCAAGCTTC				

5	8101	GCACTTTTTA	AAAGAAAAGG	GAGGACTGGA	TGGGATTTAT	TACTCCGATA	
		GGACGCTGGC					
	8161	TTGTAACTCA	GTCTCTTACT	AGGAGACCAG	CTTGAGCCTG	GGTGTTCGCT	
		GGTTAGCCTA					
10	8221	ACCTGGTTGG	CCACCAGGGG	TAAGGACTCC	TTGGCTTAGA	AAGCTAATAA	
		ACTTGCCTGC					
	8281	ATTAGAGCTC	TTACGCGTCC	CGGGCTCGAG	ATCCGCATCT	CAATTAGTCA	
		GCAACCATAG					
15	8341	TCCCGCCCCCT	AACTCCGCCC	ATCCCGCCCC	TAACTCCGCC	CAGTTCCGCC	
		CATTCTCCGC					
	8401	CCCATGGCTG	ACTAATTTTT	TTTATTTATG	CAGAGGCCGA	GGCCGCCTCG	
		GCCTCTGAGC					
20	8461	TATTCCAGAA	GTAGTGAGGA	GGCTTTTTTG	GAGGCCTAGG	CTTTTGCAAA	
		AAGCTAACTT					
	8521	GTTTATTGCA	GCTTATAATG	GTTACAAATA	AAGCAATAGC	ATCACAAATT	
		TCACAAATAA					
25	8581	AGCATTTTTT	TCACTGCATT	CTAGTTGTGG	TTTGTCCAAA	CTCATCAATG	
		TATCTTATCA					
	8641	TGTCTGTCCG	CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG	TCGTTCCGGCT	
		GCGGCGAGCG					
30	8701	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA	
		TAACGCAGGA					
	8761	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAAGGC	
		CGCGTTGCTG					
35	8821	GCGTTTTTCC	ATAGGCTCCG	CCCCCTGAC	GAGCATCACA	AAAATCGACG	
		CTCAAGTCAG					
	8881	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	
		AAGCTCCCTC					
40	8941	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	
		TCTCCCTTCG					
	9001	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	
		GTAGGTCGTT					
45	9061	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	
		CGCCTTATCC					
	9121	GGTAACTATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCGCCACT	
		GGCAGCAGCC					
50	9181	ACTGGTAACA	GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTG	CTACAGAGTT	
		CTTGAAGTGG					
	9241	TGGCCTAACT	ACGGCTACAC	TAGAAGAACA	GTATTTGGTA	TCTGCGCTCT	
		GCTGAAGCCA					
55	9301	GTTACCTTCG	GAAAAAGAGT	TGGTAGCTCT	TGATCCGGCA	AACAAACCAC	
		CGCTGGTAGC					
	9361	GGTGGTTTTT	TTGTTTGCAA	GCAGCAGATT	ACGCGCAGAA	AAAAAGGATC	
		TCAAGAAGAT					
60	9421	CCTTTGATCT	TTTCTACGGG	GTCTGACGCT	CAGTGGAACG	AAA	AACTCACG
		TTAAGGGATT					
	9481	TTGGTCATGA	GATTATCAAA	AAGGATCTTC	ACCTAGATCC	TTTTAAATTA	
		AAAATGAAGT					
60	9541	TTTAAATCAA	TCTAAAGTAT	ATATGAGTAA	ACTTGGTCTG	ACAGTTAGAA	
		AAACTCATCG					
	9601	AGCATCAAAT	GAAACTGCAA	TTTATTCATA	TCAGGATTAT	CAATACCATA	
		TTTTTGAAAA					
60	9661	AGCCGTTTCT	GTAATGAAGG	AGAAAACTCA	CCGAGGCAGT	TCCATAGGAT	
		GGCAAGATCC					
	9721	TGGTATCGGT	CTGCGATTCC	GACTCGTCCA	ACATCAATAC	AACCTATTAA	
		TTTCCCCTCG					
60	9781	TCAAAAATAA	GGTTATCAAG	TGAGAAATCA	CCATGAGTGA	CGACTGAATC	
		CGGTGAGAAT					
	9841	GGCAACAGCT	TATGCATTTT	TTTCCAGACT	TGTTCAACAG	GCCAGCCATT	
		ACGCTCGTCA					

	9901	TCAAAATCAC	TCGCATCAAC	CAAACCGTTA	TTCATTTCGTG	ATTGCGCCTG
		AGCGAGACGA				
	9961	AATACGCGAT	CGCTGTAAAA	AGGACAATTA	CAAACAGGAA	TCGAATGCAA
		CCGGCGCAGG				
5	10021	AACACTGCCA	GCGCATCAAC	AATATTTTCA	CCTGAATCAG	GATATTCTTC
		TAATACCTGG				
	10081	AATGCTGTTT	TTCCGGGGAT	CGCAGTGGTG	AGTAACCATG	CATCATCAGG
		AGTACGGATA				
10	10141	AAATGCTTGA	TGGTCGGAAG	AGGCATAAAT	TCCGTCAGCC	AGTTTAGTCT
		GACCATCTCA				
	10201	TCTGTAACAT	CATTGGCAAC	GCTACCTTTG	CCATGTTTCA	GAAACAACCTC
		TGGCGCATCG				
	10261	GGCTTCCCAT	ACAATCGATA	GATTGTCGCA	CCTGATTGCC	CGACATTATC
		GCGAGCCCAT				
15	10321	TTATACCCAT	ATAAATCAGC	ATCCATGTTG	GAATTTAATC	GCGGCCTAGA
		GCAAGACGTT				
	10381	TCCCGTTGAA	TATGGCTCAT	AACACCCCTT	GTATTACTGT	TTATGTAAGC
		AGACAGTTTT				
20	10441	ATTGTTTCATG	ATGATATATT	TTTATCTTGT	GCAATGTAAC	ATCAGAGATT
		TTGAGACACA				
	10501	ACAATTGGTC	GACGGATCC			
		SEQ ID NO:13				
	1	GGTACCTCAA	TATTGGCCAT	TAGCCATATT	ATTCATTGGT	TATATAGCAT
	61	TGGCTATTGG	CCATTGCATA	CGTTGTATCT	ATATCATAAT	ATGTACATTT
25	121	ATGTCCAATA	TGACCGCCAT	GTTGGCATTG	ATTATTGACT	AGTTATTAAT
		AGTAATCAAT				
	181	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCCGC	GTTACATAAC
		TTACGGTAAA				
30	241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA
		TGACGTATGT				
	301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT
		ATTTACGGTA				
	361	AACTGCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTCCGCCCC
		CTATTGACGT				
35	421	CAATGACGGT	AAATGGCCCG	CCTGGCATTG	TGCCCAGTAC	ATGACCTTAC
		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
40	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
		TCCACCCCAT				
	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
45	721	TATATAAGCA	GAGCTCGCTG	GCTTGTAACT	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTTCG	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
50	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCCCT
		CATTGACGCC				
	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
55	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTIONGAG
		AGGCCGTAGC				
60	1141	CGTAACTACT	CTGGGCAAGT	AGGGCAGGCG	GTGGGTACGC	AATGGGGGCG
		GCTACCTCAG				

	1201	CACTAAATAG	GAGACAATTA	GACCAATTTG	AGAAAATACG	ACTTCGCCCCG
		AACGGAAAGA				
	1261	AAAAGTACCA	AATTAAACAT	TTAATATGGG	CAGGCAAGGA	GATGGAGCGC
		TTCGGCCTCC				
5	1321	ATGAGAGGTT	GTTGGAGACA	GAGGAGGGGT	GTAAAAGAAT	CATAGAAGTC
		CTCTACCCCC				
	1381	TAGAACCAAC	AGGATCGGAG	GGCTTAAAAA	GTCTGTTCAA	TCTTGTGTGC
		GTGCTATATT				
10	1441	GCTTGCACAA	GGAACAGAAA	GTGAAAGACA	CAGAGGAAGC	AGTAGCAACA
		GTAAGACAAC				
	1501	ACTGCCATCT	AGTGGAAAAA	GAAAAAAGTG	CAACAGAGAC	ATCTAGTGGA
		CAAAAGAAAA				
	1561	ATGACAAGGG	AATAGCAGCG	CCACCTGGTG	GCAGTCAGAA	TTTTCCAGCG
		CAACAACAAG				
15	1621	GAAATGCCTG	GGTACATGTA	CCCTTGTGAC	CGCGCACCTT	AAATGCGTGG
		GTAAAAGCAG				
	1681	TAGAGGAGAA	AAAATTTGGA	GCAGAAATAG	TACCCATGTT	TCAAGCCCTA
		TCGAATTCCC				
20	1741	GTTTGTGCTA	GGGTTCTTAG	GCTTCTTGGG	GGCTGCTGGA	ACTGCAATGG
		GAGCAGCGGC				
	1801	GACAGCCCTG	ACGGTCCAGT	CTCAGCATTT	GCTTGCTGGG	ATACTGCAGC
		AGCAGAAGAA				
	1861	TCTGCTGGCG	GCTGTGGAGG	CTCAACAGCA	GATGTTGAAG	CTGACCATTT
		GGGGTGTTAA				
25	1921	AAACCTCAAT	GCCCGCGTCA	CAGCCCTTGA	GAAGTACCTA	GAGGATCAGG
		CACGACTAAA				
	1981	CTCCTGGGGG	TGCGCATGGA	AACAAGTATG	TCATACCACA	GTGGAGTGGC
		CCTGGACAAA				
30	2041	TCGGAATCCG	GATTGGCAAA	ATATGACTTG	GTTGGAGTGG	GAAAGACAAA
		TAGCTGATTT				
	2101	GGAAAGCAAC	ATTACGAGAC	AATTAGTGAA	GGCTAGAGAA	CAAGAGGAAA
		AGAATCTAGA				
	2161	TGCCTATCAG	AAGTTAACTA	GTTGGTCAGA	TTTCTGGTCT	TGGTTCGATT
		TCTCAAAATG				
35	2221	GCTTAACATT	TTAAAAATGG	GATTTTTAGT	AATAGTAGGA	ATAATAGGGT
		TAAGATTACT				
	2281	TTACACAGTA	TATGGATGTA	TAGTGAGGGT	TAGGCAGGGA	TATGTTCCCTC
		TATCTCCACA				
40	2341	GATCCATATC	CGCGGCAATT	TTAAAAGAAA	GGGAGGAATA	GGGGGACAGA
		CTTCAGCAGA				
	2401	GAGACTAATT	AATATAATAA	CAACACAATT	AGAAATACAA	CATTTACAAA
		CCAAAATTCA				
	2461	AAAAATTTTA	AATTTTAGAG	CCGCGGAGAT	CTCAATATTG	GCCATTAGCC
		ATATTATTCA				
45	2521	TTGGTTATAT	AGCATAAATC	AATATTGGCT	ATTGGCCATT	GCATACGTTG
		TATCTATATC				
	2581	ATAATATGTA	CATTTATATT	GGCTCATGTC	CAATATGACC	GCCATGTTGG
		CATTGATTAT				
50	2641	TGACTAGTTA	TTAATAGTAA	TCAATTACGG	GGTCATTAGT	TCATAGCCCA
		TATATGGAGT				
	2701	TCCGCGTTAC	ATAACTTACG	GTAAATGGCC	CGCCTGGCTG	ACCGCCCAAC
		GACCCCCGCC				
	2761	CATTGACGTC	AATAATGACG	TATGTTCCCA	TAGTAACGCC	AATAGGGACT
		TTCCATTGAC				
55	2821	GTCAATGGGT	GGAGTATTTA	CGGTAAACTG	CCCATTGGC	AGTACATCAA
		GTGTATCATA				
	2881	TGCCAAGTCC	GCCCCCTATT	GACGTCAATG	ACGGTAAATG	GCCCCGCTGG
		CATTATGCCC				
60	2941	AGTACATGAC	CTTACGGGAC	TTTCCTACTT	GGCAGTACAT	CTACGTATTA
		GTCATCGCTA				

	3001	TTACCATGGT	GATGCGGTTT	TGGCAGTACA	CCAATGGGCG	TGGATAGCGG
		TTTGACTCAC				
	3061	GGGGATTTCC	AAGTCTCCAC	CCCATTGACG	TCAATGGGAG	TTTGTTTTGG
		CACCAAAATC				
5	3121	AACGGGACTT	TCCAAAATGT	CGTAATAACC	CCGCCCCGTT	GACGCAAATG
		GGCGGTAGGC				
	3181	GTGTACGGTG	GGAGGTCTAT	ATAAGCAGAG	CTCGTTTAGT	GAACCGTCAG
		ATCACTAGAA				
10	3241	GCTTTATTGC	GGTAGTTTAT	CACAGTTAAA	TTGCTAACGC	AGTCAGTGCT
		TCTGACACAA				
	3301	CAGTCTCGAA	CTTAAGCTGC	AGAAGTTGGT	CGTGAGGCAC	TGGGCAGGCT
		AGCCACCAAT				
	3361	GCAGATTGAG	CTGAGCACCT	GCTTCTTCCT	GTGCCTGCTG	AGGTTCTGCT
		TCTCTGCCAC				
15	3421	CAGGAGATAC	TACCTGGGGG	CTGTGGAGCT	GAGCTGGGAC	TACATGCAGT
		CTGACCTGGG				
	3481	GGAGCTGCCT	GTGGATGCCA	GGTTCCCCCC	CAGAGTGCCC	AAGAGCTTCC
		CCTTCAACAC				
20	3541	CTCTGTGGTG	TACAAGAAGA	CCCTGTTTGT	GGAGTTCACT	GACCACCTGT
		TCAACATTGC				
	3601	CAAGCCCAGG	CCCCCTGGA	TGGGCCTGCT	GGGCCCCACC	ATCCAGGCTG
		AGGTGTATGA				
	3661	CACTGTGGTG	ATCACCTGA	AGAACATGGC	CAGCCACCCT	GTGAGCCTGC
		ATGCTGTGGG				
25	3721	GGTGAGCTAC	TGGAAGGCCT	CTGAGGGGGC	TGAGTATGAT	GACCAGACCA
		GCCAGAGGGA				
	3781	GAAGGAGGAT	GACAAGGTGT	TCCCTGGGGG	CAGCCACACC	TATGTGTGGC
		AGGTGCTGAA				
30	3841	GGAGAATGGC	CCCATGGCCT	CTGACCCCCT	GTGCCTGACC	TACAGCTACC
		TGAGCCATGT				
	3901	GGACCTGGTG	AAGGACCTGA	ACTCTGGCCT	GATTGGGGCC	CTGCTGGTGT
		GCAGGGAGGG				
	3961	CAGCCTGGCC	AAGGAGAAGA	CCCAGACCCT	GCACAAGTTC	ATCCTGCTGT
		TTGCTGTGTT				
35	4021	TGATGAGGGC	AAGAGCTGGC	ACTCTGAAAC	CAAGAACAGC	CTGATGCAGG
		ACAGGGATGC				
	4081	TGCCTCTGCC	AGGGCCTGGC	CCAAGATGCA	CACTGTGAAT	GGCTATGTGA
		ACAGGAGCCT				
40	4141	GCCTGGCCTG	ATTGGCTGCC	ACAGGAAGTC	TGTGTACTGG	CATGTGATTG
		GCATGGGCAC				
	4201	CACCCCTGAG	GTGCACAGCA	TCTTCCTGGA	GGGCCACACC	TTCCTGGTCA
		GGAACCACAG				
	4261	GCAGGCCAGC	CTGGAGATCA	GCCCCATCAC	CTTCCTGACT	GCCCAGACCC
		TGCTGATGGA				
45	4321	CCTGGGCCAG	TTCCTGCTGT	TCTGCCACAT	CAGCAGCCAC	CAGCATGATG
		GCATGGAGGC				
	4381	CTATGTGAAG	GTGGACAGCT	GCCCTGAGGA	GCCCCAGCTG	AGGATGAAGA
		ACAATGAGGA				
50	4441	GGCTGAGGAC	TATGATGATG	ACCTGACTGA	CTCTGAGATG	GATGTGGTGA
		GGTTTGTATGA				
	4501	TGACAACAGC	CCCAGCTTCA	TCCAGATCAG	GTCTGTGGCC	AAGAAGCACC
		CCAAGACCTG				
	4561	GGTGCACTAC	ATTGCTGCTG	AGGAGGAGGA	CTGGGACTAT	GCCCCCTGG
		TGCTGGCCCC				
55	4621	TGATGACAGG	AGCTACAAGA	GCCAGTACCT	GAACAATGGC	CCCCAGAGGA
		TTGGCAGGAA				
	4681	GTACAAGAAG	GTCAGGTTCA	TGGCCTACAC	TGATGAAACC	TTCAAGACCA
		GGGAGGCCAT				
60	4741	CCAGCATGAG	TCTGGCATCC	TGGGCCCCCT	GCTGTATGGG	GAGGTGGGGG
		ACACCCTGCT				

	4801	GATCATCTTC	AAGAACCAGG	CCAGCAGGCC	CTACAACATC	TACCCCCATG
		GCATCACTGA				
	4861	TGTGAGGCCC	CTGTACAGCA	GGAGGCTGCC	CAAGGGGGTG	AAGCACCTGA
		AGGACTTCCC				
5	4921	CATCCTGCCT	GGGAGATCT	TCAAGTACAA	GTGGACTGTG	ACTGTGGAGG
		ATGGCCCCAC				
	4981	CAAGTCTGAC	CCCAGGTGCC	TGACCAGATA	CTACAGCAGC	TTTGTGAACA
		TGGAGAGGGA				
	5041	CCTGGCCTCT	GGCCTGATTG	GCCCCCTGCT	GATCTGCTAC	AAGGAGTCTG
10		TGGACCAGAG				
	5101	GGGCAACCAG	ATCATGTCTG	ACAAGAGGAA	TGTGATCCTG	TTCTCTGTGT
		TTGATGAGAA				
	5161	CAGGAGCTGG	TACCTGACTG	AGAACATCCA	GAGGTTCTCTG	CCCAACCCTG
		CTGGGGTGCA				
15	5221	GCTGGAGGAC	CCTGAGTTCC	AGGCCAGCAA	CATCATGCAC	AGCATCAATG
		GCTATGTGTT				
	5281	TGACAGCCTG	CAGCTGTCTG	TGTGCCTGCA	TGAGGTGGCC	TACTGGTACA
		TCCTGAGCAT				
	5341	TGGGGCCACG	ACTGACTTCC	TGTCTGTGTT	CTTCTCTGGC	TACACCTTCA
20		AGCACAAGAT				
	5401	GGTGTATGAG	GACACCCTGA	CCCTGTTCCC	CTTCTCTGGG	GAGACTGTGT
		TCATGAGCAT				
	5461	GGAGAACCCT	GGCCTGTGGA	TTCTGGGCTG	CCACAACCTCT	GACTTCAGGA
		ACAGGGGGCAT				
25	5521	GACTGCCCTG	CTGAAAGTCT	CCAGCTGTGA	CAAGAACACT	GGGGACTACT
		ATGAGGACAG				
	5581	CTATGAGGAC	ATCTCTGCCT	ACCTGCTGAG	CAAGAACAAT	GCCATTGAGC
		CCAGGAGCTT				
	5641	CAGCCAGAAC	AGCAGGCACC	CCAGCACCAG	GCAGAAGCAG	TTCAATGCCA
30		CCACCATCCC				
	5701	TGAGAATGAC	ATAGAGAAGA	CAGACCCATG	GTTTGCCAC	CGGACCCCCA
		TGCCCAAGAT				
	5761	CCAGAATGTG	AGCAGCTCTG	ACCTGCTGAT	GCTGCTGAGG	CAGAGCCCCA
		CCCCCATGG				
35	5821	CCTGAGCCTG	TCTGACCTGC	AGGAGGCCAA	GTATGAAACC	TTCTCTGATG
		ACCCAGCCC				
	5881	TGGGGCCATT	GACAGCAACA	ACAGCCTGTC	TGAGATGACC	CACTTCAGGC
		CCCAGCTGCA				
	5941	CCACTCTGGG	GACATGGTGT	TCACCCCTGA	GTCTGGCCTG	CAGCTGAGGC
40		TGAATGAGAA				
	6001	GCTGGGCACC	ACTGCTGCCA	CTGAGCTGAA	GAAGCTGGAC	TTCAAAGTCT
		CCAGCACCAG				
	6061	CAACAACCTG	ATCAGCACCA	TCCCCTCTGA	CAACCTGGCT	GCTGGCACTG
		ACAACACCAG				
45	6121	CAGCCTGGGC	CCCCCCAGCA	TGCCTGTGCA	CTATGACAGC	CAGCTGGACA
		CCACCCTGTT				
	6181	TGGCAAGAAG	AGCAGCCCCC	TGACTGAGTC	TGGGGGCCCC	CTGAGCCTGT
		CTGAGGAGAA				
	6241	CAATGACAGC	AAGCTGCTGG	AGTCTGGCCT	GATGAACAGC	CAGGAGAGCA
50		GCTGGGGCAA				
	6301	GAATGTGAGC	AGCAGGGAGA	TCACCAGGAC	CACCCTGCAG	TCTGACCAGG
		AGGAGATTGA				
	6361	CTATGATGAC	ACCATCTCTG	TGGAGATGAA	GAAGGAGGAC	TTTGACATCT
		ACGACGAGGA				
55	6421	CGAGAACCAG	AGCCCCAGGA	GCTTCCAGAA	GAAGACCAGG	CACTACTTCA
		TTGCTGCTGT				
	6481	GGAGAGGCTG	TGGGACTATG	GCATGAGCAG	CAGCCCCCAT	GTGCTGAGGA
		ACAGGGCCCA				
	6541	GTCTGGCTCT	GTGCCCCAGT	TCAAGAAGGT	GGTGTTCCAG	GAGTTCACTG
60		ATGGCAGCTT				

	6601	CACCCAGCCC	CTGTACAGAG	GGGAGCTGAA	TGAGCACCTG	GGCCTGCTGG
		GCCCCTACAT				
	6661	CAGGGCTGAG	GTGGAGGACA	ACATCATGGT	GACCTTCAGG	AACCAGGCCA
		GCAGGCCCTA				
5	6721	CAGCTTCTAC	AGCAGCCTGA	TCAGCTATGA	GGAGGACCAG	AGGCAGGGGG
		CTGAGCCCAG				
	6781	GAAGAACTTT	GTGAAGCCCA	ATGAAACCAA	GACCTACTTC	TGGAAGGTGC
		AGCACCACAT				
	6841	GGCCCCCACC	AAGGATGAGT	TTGACTGCAA	GGCCTGGGCC	TACTTCTCTG
10		ATGTGGACCT				
	6901	GGAGAAGGAT	GTGCACTCTG	GCCTGATTGG	CCCCCTGCTG	GTGTGCCACA
		CCAACACCT				
	6961	GAACCCTGCC	CATGGCAGGC	AGGTGACTGT	GCAGGAGTTT	GCCCTGTTCT
		TCACCATCTT				
15	7021	TGATGAAACC	AAGAGCTGGT	ACTTCACTGA	GAACATGGAG	AGGAACTGCA
		GGGCCCCCTG				
	7081	CAACATCCAG	ATGGAGGACC	CCACCTTCAA	GGAGAACTAC	AGGTTCCATG
		CCATCAATGG				
	7141	CTACATCATG	GACACCCTGC	CTGGCCTGGT	GATGGCCCAG	GACCAGAGGA
20		TCAGGTGGTA				
	7201	CCTGCTGAGC	ATGGGCAGCA	ATGAGAACAT	CCACAGCATC	CACTTCTCTG
		GCCATGTGTT				
	7261	CACTGTGAGG	AAGAAGGAGG	AGTACAAGAT	GGCCCTGTAC	AACCTGTACC
		CTGGGGTGTT				
25	7321	TGAGACTGTG	GAGATGCTGC	CCAGCAAGGC	TGGCATCTGG	AGGGTGGAGT
		GCCTGATTGG				
	7381	GGAGCACCTG	CATGCTGGCA	TGAGCACCTT	GTTCTTGGTG	TACAGCAACA
		AGTGCCAGAC				
	7441	CCCCCTGGGC	ATGGCCTCTG	GCCACATCAG	GGACTTCCAG	ATCACTGCCT
30		CTGGCCAGTA				
	7501	TGGCCAGTGG	GCCCCCAAGC	TGGCCAGGCT	GCACTACTCT	GGCAGCATCA
		ATGCCTGGAG				
	7561	CACCAAGGAG	CCCTTCAGCT	GGATCAAGGT	GGACCTGCTG	GCCCCCATGA
		TCATCCATGG				
35	7621	CATCAAGACC	CAGGGGGCCA	GGCAGAAGTT	CAGCAGCCTG	TACATCAGCC
		AGTTCATCAT				
	7681	CATGTACAGC	CTGGATGGCA	AGAAGTGGCA	GACCTACAGG	GGCAACAGCA
		CTGGCACCTT				
	7741	GATGGTGTTT	TTTGGCAATG	TGGACAGCTC	TGGCATCAAG	CACAACATCT
40		TCAACCCCCC				
	7801	CATCATTGCC	AGATACATCA	GGCTGCACCC	CACCCACTAC	AGCATCAGGA
		GCACCCTGAG				
	7861	GATGGAGCTG	ATGGGCTGTG	ACCTGAACAG	CTGCAGCATG	CCCCTGGGCA
		TGGAGAGCAA				
45	7921	GGCCATCTCT	GATGCCCAGA	TCACTGCCAG	CAGCTACTTC	ACCAACATGT
		TTGCCACCTG				
	7981	GAGCCCCAGC	AAGGCCAGGC	TGCACCTGCA	GGGCAGGAGC	AATGCCTGGA
		GGCCCCAGGT				
	8041	CAACAACCCC	AAGGAGTGGC	TGCAGGTGGA	CTTCCAGAAG	ACCATGAAGG
50		TGACTGGGGT				
	8101	GACCACCCAG	GGGGTGAAGA	GCCTGCTGAC	CAGCATGTAT	GTGAAGGAGT
		TCCTGATCAG				
	8161	CAGCAGCCAG	GATGGCCACC	AGTGGACCCT	GTTCTTCCAG	AATGGCAAGG
		TGAAGGTGTT				
55	8221	CCAGGGCAAC	CAGGACAGCT	TCACCCCTGT	GGTGAACAGC	CTGGACCCCC
		CCCTGCTGAC				
	8281	CAGATACCTG	AGGATTCACC	CCCAGAGCTG	GGTGCACCAG	ATTGCCCTGA
		GGATGGAGGT				
	8341	GCTGGGCTGT	GAGGCCCAGG	ACCTGTACTG	AGCGGCCGCG	GGCCCAATCA
60		ACCTCTGGAT				

	8401	TACAAAATTT	GTGAAAGATT	GA CTGGTATT	CTTA ACTATG	TTGCTCCTTT
		TACGCTATGT				
	8461	GGATACGCTG	CTTTAATGCC	TTTGTATCAT	GCTATTGCTT	CCCGTATGGC
		TTTCATTTTC				
5	8521	TCCTCCTTGT	ATAAATCCTG	GTTGCTGTCT	CTTTATGAGG	AGTTGTGGCC
		CGTTGTCAGG				
	8581	CAACGTGGCG	TGGTGTGCAC	TGTGTTTGCT	GACGCAACCC	CCACTGGTTG
		GGGCATTGCC				
10	8641	ACCACCTGTC	AGCTCCTTTC	CGGGACTTTC	GCTTTCCCCC	TCCCTATTGC
		CACGGCGGAA				
	8701	CTCATCGCCG	CCTGCCTTGC	CCGCTGCTGG	ACAGGGGCTC	GGCTGTTGGG
		CACTGACAAT				
	8761	TCCGTGGTGT	TGTCGGGGAA	ATCATCGTCC	TTTCCTTGGC	TGCTCGCCTG
		TGTTGCCACC				
15	8821	TGATTCTGC	GCGGGACGTC	CTTCTGCTAC	GTCCCTTCGG	CCCTCAATCC
		AGCGGACCTT				
	8881	CCTTCCCGCG	GCCTGCTGCC	GGCTCTGCGG	CCTCTTCCGC	GTCTTCGCCT
		TCGCCCTCAG				
20	8941	ACGAGTCGGA	TCTCCCTTTG	GGCCGCCTCC	CCGCAAGCTT	CGCACTTTTT
		AAAAGAAAAG				
	9001	GGAGGACTGG	ATGGGATTTA	TTACTCCGAT	AGGACGCTGG	CTTGTA ACTC
		AGTCTCTTAC				
	9061	TAGGAGACCA	GCTTGAGCCT	GGGTGTTGCG	TGGTTAGCCT	AACCTGGTTG
		GCCACCAGGG				
25	9121	GTAAGGACTC	CTTGGCTTAG	AAAGCTAATA	AACTTGCCCTG	CATTAGAGCT
		CTTACGCGTC				
	9181	CCGGGCTCGA	GATCCGCATC	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC
		TAACTCCGCC				
30	9241	CATCCCGCCC	CTAACTCCGC	CCAGTTCCGC	CCATTCTCCG	CCCCATGGCT
		GACTAATTTT				
	9301	TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	GGCCTCTGAG	CTATTCCAGA
		AGTAGTGAGG				
	9361	AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTAACT	TGTTTATTGC
		AGCTTATAAT				
35	9421	GGTTACAAAT	AAAGCAATAG	CATCACAAAT	TTCACAAATA	AAGCATTTTT
		TTCACTGCAT				
	9481	TCTAGTTGTG	GTTTGTCCAA	ACTCATCAAT	GTATCTTATC	ATGTCTGTCC
		GCTTCCTCGC				
40	9541	TCACTGACTC	GCTGCGCTCG	GTCGTTCGGC	TGCGGCGAGC	GGTATCAGCT
		CACTCAAAGG				
	9601	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG	AAAGAACATG
		TGAGCAAAAAG				
	9661	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG	CCGCGTTGCT	GGCGTTTTTC
		CATAGGCTCC				
45	9721	GCCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAGTCA	GAGGTGGCGA
		AACCCGACAG				
	9781	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT
		CCTGTTCCGA				
50	9841	CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	TTCTCCCTTC	GGGAAGCGTG
		GCGCTTTCTC				
	9901	ATAGCTCACG	CTGTAGGTAT	CTCAGTTGCG	TGTAGGTCGT	TCGCTCCAAG
		CTGGGCTGTG				
	9961	TGCACGAACC	CCCCGTTTCC	CCCGACCGCT	GCGCCTTATC	CGGTA ACTAT
		CGTCTTGAGT				
55	10021	CCAACCCGGT	AAGACACGAC	TTATCGCCAC	TGGCAGCAGC	CACTGGTAAC
		AGGATTAGCA				
	10081	GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	TCTTGAAGTG	GTGGCCTAAC
		TACGGCTACA				
60	10141	CTAGAAGAAC	AGTATTTGGT	ATCTGCGCTC	TGCTGAAGCC	AGTTACCTTC
		GGAAAAAGAG				

	10201	TTGGTAGCTC	TTGATCCGGC	AAACAAACCA	CCGCTGGTAG	CGGTGGTTTT
		TTTGTTTGCA				
	10261	AGCAGCAGAT	TACGCGCAGA	AAAAAAGGAT	CTCAAGAAGA	TCCTTTGATC
		TTTTCTACGG				
5	10321	GGTCTGACGC	TCAGTGGAAC	GAAAACTCAC	GTTAAGGGAT	TTTGGTCATG
		AGATTATCAA				
	10381	AAAGGATCTT	CACCTAGATC	CTTTTAAATT	AAAAATGAAG	TTTTAAATCA
		ATCTAAAGTA				
	10441	TATATGAGTA	AACTTGGTCT	GACAGTTAGA	AAAACATCATC	GAGCATCAAA
10		TGAAACTGCA				
	10501	ATTTATTCAT	ATCAGGATTA	TCAATACCAT	ATTTTTGAAA	AAGCCGTTTC
		TGTAATGAAG				
	10561	GAGAAACTC	ACCGAGGCAG	TTCCATAGGA	TGGCAAGATC	CTGGTATCGG
		TCTGCGATTC				
15	10621	CGACTCGTCC	AACATCAATA	CAACCTATTA	ATTTCCCCTC	GTCAAAAATA
		AGGTTATCAA				
	10681	GTGAGAAATC	ACCATGAGTG	ACGACTGAAT	CCGGTGAGAA	TGGCAACAGC
		TTATGCATTT				
	10741	CTTTCCAGAC	TTGTTCAACA	GGCCAGCCAT	TACGCTCGTC	ATCAAAATCA
20		CTCGCATCAA				
	10801	CCAAACCGTT	ATTCATTCGT	GATTGCGCCT	GAGCGAGACG	AAATACGCGA
		TCGCTGTAA				
	10861	AAGGACAATT	ACAAACAGGA	ATCGAATGCA	ACCGGCGCAG	GAACACTGCC
		AGCGCATCAA				
25	10921	CAATATTTTC	ACCTGAATCA	GGATATTCTT	CTAATACCTG	GAATGCTGTT
		TTTCCGGGGA				
	10981	TCGCAGTGGT	GAGTAACCAT	GCATCATCAG	GAGTACGGAT	AAAATGCTTG
		ATGGTCGGAA				
	11041	GAGGCATAAA	TTCCGTCAGC	CAGTTTAGTC	TGACCATCTC	ATCTGTAACA
30		TCATTGGCAA				
	11101	CGCTACCTTT	GCCATGTTTC	AGAAACAAC	CTGGCGCATC	GGGCTTCCCA
		TACAATCGAT				
	11161	AGATTGTCGC	ACCTGATTGC	CCGACATTAT	CGCGAGCCCA	TTTATACCCA
		TATAAATCAG				
35	11221	CATCCATGTT	GGAATTTAAT	CGCGGCCTAG	AGCAAGACGT	TTCCCGTTGA
		ATATGGCTCA				
	11281	TAACACCCCT	TGTATTACTG	TTTATGTAAG	CAGACAGTTT	TATTGTTTCA
		GATGATATAT				
	11341	TTTTATCTTG	TGCAATGTAA	CATCAGAGAT	TTTGAGACAC	AACAATTGGT
40		CGACGGATCC				
		SEQ ID NO:14				
	1	GGTACCTCAA	TATTGGCCAT	TAGCCATATT	ATTCATTGGT	TATATAGCAT
	61	TGGCTATTGG	CCATTGCATA	CGTTGTATCT	ATATCATAAT	ATGTACATTT
	121	ATGTCCAATA	TGACCGCCAT	GTTGGCATTG	ATTATTGACT	AGTTATTAAT
45		AGTAATCAAT				
	181	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCCGC	GTTACATAAC
		TTACGGTAAA				
	241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	ACGTCAATAA
		TGACGTATGT				
50	301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT
		ATTTACGGTA				
	361	AACTGCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTCCGCCCC
		CTATTGACGT				
	421	CAATGACGGT	AAATGGCCCG	CCTGGCATTG	TGCCCAGTAC	ATGACCTTAC
55		GGGACTTTCC				
	481	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC
		GGTTTTGGCA				
	541	GTACACCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC
		TCCACCCCAT				

	601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA
		AATGTCGTAA				
	661	CAACTGCGAT	CGCCCGCCCC	GTTGACGCAA	ATGGGCGGTA	GGCGTGTACG
		GTGGGAGGTC				
5	721	TATATAAGCA	GAGCTCGCTG	GCTTGTAACT	CAGTCTCTTA	CTAGGAGACC
		AGCTTGAGCC				
	781	TGGGTGTTCTG	CTGGTTAGCC	TAACCTGGTT	GGCCACCAGG	GGTAAGGACT
		CCTTGGCTTA				
	841	GAAAGCTAAT	AAACTTGCCT	GCATTAGAGC	TTATCTGAGT	CAAGTGTCTT
10		CATTGACGCC				
	901	TCACTCTCTT	GAACGGGAAT	CTTCCTTACT	GGGTCTCTC	TCTGACCCAG
		GCGAGAGAAA				
	961	CTCCAGCAGT	GGCGCCCGAA	CAGGGACTTG	AGTGAGAGTG	TAGGCACGTA
		CAGCTGAGAA				
15	1021	GGCGTCGGAC	GCGAAGGAAG	CGCGGGGTGC	GACGCGACCA	AGAAGGAGAC
		TTGGTGAGTA				
	1081	GGCTTCTCGA	GTGCCGGGAA	AAAGCTCGAG	CCTAGTTAGA	GGACTAGGAG
		AGGCCGTAGC				
	1141	CGTAACTACT	CTTGGGCAAG	TAGGGCAGGC	GGTGGGTACG	CAATGGGGGC
20		GGCTACCTCA				
	1201	GCACTAAATA	GGAGACAATT	AGACCAATTT	GAGAAAATAC	GACTTCGCCC
		GAACGGAAAG				
	1261	AAAAAGTACC	AAATTAACA	TTAATATGG	GCAGGCAAGG	AGATGGAGCG
		CTTCGGCCTC				
25	1321	CATGAGAGGT	TGTTGGAGAC	AGAGGAGGGG	TGTAAAAGAA	TCATAGAAGT
		CCTCTACCCC				
	1381	CTAGAACCAA	CAGGATCGGA	GGGCTTAAAA	AGTCTGTTCA	ATCTTGTGTG
		CGTGCTATAT				
	1441	TGCTTGACACA	AGGAACAGAA	AGTGAAAGAC	ACAGAGGAAG	CAGTAGCAAC
30		AGTAAGACAA				
	1501	CACTGCCATC	TAGTGGAAAA	AGAAAAAAGT	GCAACAGAGA	CATCTAGTGG
		ACAAAAGAAA				
	1561	AATGACAAGG	GAATAGCAGC	GCCACCTGGT	GGCAGTCAGA	ATTTTCCAGC
		GCAACAACAA				
35	1621	GGAAATGCCT	GGGTACATGT	ACCCTTGTCA	CCGCGCACCT	TAAATGCGTG
		GGTAAAAGCA				
	1681	GTAGAGGAGA	AAAAATTTGG	AGCAGAAATA	GTACCCATGT	TTCAAGCCCT
		ATCGAATTCC				
	1741	CGTTTGTGCT	AGGGTTCTTA	GGCTTCTTGG	GGGCTGCTGG	AACTGCAATG
40		GGAGCAGCGG				
	1801	CGACAGCCCT	GACGGTCCAG	TCTCAGCATT	TGCTTGCTGG	GATACTGCAG
		CAGCAGAAGA				
	1861	ATCTGCTGGC	GGCTGTGGAG	GCTCAACAGC	AGATGTTGAA	GCTGACCATT
		TGGGGTGTTA				
45	1921	AAAACCTCAA	TGCCCGCGTC	ACAGCCCTTG	AGAAGTACCT	AGAGGATCAG
		GCACGACTAA				
	1981	ACTCCTGGGG	GTGCGCATGG	AAACAAGTAT	GTCATACCAC	AGTGGAGTGG
		CCCTGGACAA				
	2041	ATCGGACTCC	GGATTGGCAA	AATATGACTT	GGTTGGAGTG	GGAAAGACAA
50		ATAGCTGATT				
	2101	TGGAAAGCAA	CATTACGAGA	CAATTAGTGA	AGGCTAGAGA	ACAAGAGGAA
		AAGAATCTAG				
	2161	ATGCCTATCA	GAAGTTAACT	AGTTGGTCAG	ATTTCTGGTC	TTGGTTCGAT
		TTCTCAAAAT				
55	2221	GGCTTAACAT	TTTAAAAATG	GGATTTTTAG	TAATAGTAGG	AATAATAGGG
		TTAAGATTAC				
	2281	TTTACACAGT	ATATGGATGT	ATAGTGAGGG	TTAGGCAGGG	ATATGTTCTT
		CTATCTCCAC				
	2341	AGATCCATAT	CCGCGGCAAT	TTTAAAAGAA	AGGGAGGAAT	AGGGGGACAG
60		ACTTCAGCAG				

	2401	AGAGACTAAT	TAATATAATA	ACAACACAAT	TAGAAATACA	ACATTTACAA
		ACCAAAATTC				
	2461	AAAAAATTTT	AAATTTTAGA	GCCGCGGAGA	TCTGTTACAT	AACTTATGGT
		AAATGGCCTG				
5	2521	CCTGGCTGAC	TGCCCAATGA	CCCCTGCCCA	ATGATGTCAA	TAATGATGTA
		TGTTCCCATG				
	2581	TAATGCCAAT	AGGGACTTTC	CATTGATGTC	AATGGGTGGA	GTATTTATGG
		TAAGTGGCCA				
10	2641	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTATGCCC	CCTATTGATG
		TCAATGATGG				
	2701	TAAATGGCCT	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA	TGGGACTTTC
		CTACTTGGCA				
	2761	GTACATCTAT	GTATTAGTCA	TTGCTATTAC	CATGGGAATT	CACTAGTGGA
		GAAGAGCATG				
15	2821	CTTGAGGGCT	GAGTGCCCCT	CAGTGGGCAG	AGAGCACATG	GCCCACAGTC
		CCTGAGAAGT				
	2881	TGGGGGGAGG	GGTGGGCAAT	TGAACTGGTG	CCTAGAGAAG	GTGGGGGCTTG
		GGTAAACTGG				
20	2941	GAAAGTGATG	TGGTGTACTG	GCTCCACCTT	TTTCCCCAGG	GTGGGGGAGA
		ACCATATATA				
	3001	AGTGCAGTAG	TCTCTGTGAA	CATTCAAGCT	TCTGCCTTCT	CCCTCCTGTG
		AGTTTGCTAG				
	3061	CCACCAATGC	AGATTGAGCT	GAGCACCTGC	TTCTTCCTGT	GCCTGCTGAG
		GTTCTGCTTC				
25	3121	TCTGCCACCA	GGAGATACTA	CCTGGGGGCT	GTGGAGCTGA	GCTGGGACTA
		CATGCAGTCT				
	3181	GACCTGGGGG	AGCTGCCTGT	GGATGCCAGG	TTCCCCCCCA	GAGTGCCCAA
		GAGCTTCCCC				
30	3241	TTCAACACCT	CTGTGGTGTA	CAAGAAGACC	CTGTTTGTGG	AGTTCACTGA
		CCACCTGTTT				
	3301	AACATTGCCA	AGCCCAGGCC	CCCCTGGATG	GGCCTGCTGG	GCCCCACCAT
		CCAGGCTGAG				
	3361	GTGTATGACA	CTGTGGTGAT	CACCCTGAAG	AACATGGCCA	GCCACCCTGT
		GAGCCTGCAT				
35	3421	GCTGTGGGGG	TGAGCTACTG	GAAGGCCTCT	GAGGGGGCTG	AGTATGATGA
		CCAGACCAGC				
	3481	CAGAGGGAGA	AGGAGGATGA	CAAGGTGTTC	CCTGGGGGCA	GCCACACCTA
		TGTGTGGCAG				
40	3541	GTGCTGAAGG	AGAATGGCCC	CATGGCCTCT	GACCCCCTGT	GCCTGACCTA
		CAGCTACCTG				
	3601	AGCCATGTGG	ACCTGGTGAA	GGACCTGAAC	TCTGGCCTGA	TTGGGGGCCCT
		GCTGGTGTGC				
	3661	AGGGAGGGCA	GCCTGGCCAA	GGAGAAGACC	CAGACCCTGC	ACAAGTTCAT
		CCTGCTGTTT				
45	3721	GCTGTGTTTG	ATGAGGGCAA	GAGCTGGCAC	TCTGAAACCA	AGAACAGCCT
		GATGCAGGAC				
	3781	AGGGATGCTG	CCTCTGCCAG	GGCCTGGCCC	AAGATGCACA	CTGTGAATGG
		CTATGTGAAC				
50	3841	AGGAGCCTGC	CTGGCCTGAT	TGGCTGCCAC	AGGAAGTCTG	TGTACTGGCA
		TGTGATTGGC				
	3901	ATGGGCACCA	CCCCTGAGGT	GCACAGCATC	TTCTTGAGAG	GCCACACCTT
		CCTGGTCAGG				
	3961	AACCACAGGC	AGGCCAGCCT	GGAGATCAGC	CCCATCACCT	TCCTGACTGC
		CCAGACCCTG				
55	4021	CTGATGGACC	TGGGCCAGTT	CCTGCTGTTC	TGCCACATCA	GCAGCCACCA
		GCATGATGGC				
	4081	ATGGAGGCCT	ATGTGAAGGT	GGACAGCTGC	CCTGAGGAGC	CCCAGCTGAG
		GATGAAGAAC				
60	4141	AATGAGGAGG	CTGAGGACTA	TGATGATGAC	CTGACTGACT	CTGAGATGGA
		TGTGGTGAGG				

	4201	TTTGATGATG	ACAACAGCCC	CAGCTTCATC	CAGATCAGGT	CTGTGGCCAA
		GAAGCACCCC				
	4261	AAGACCTGGG	TGCACTACAT	TGCTGCTGAG	GAGGAGGACT	GGGACTATGC
		CCCCCTGGTG				
5	4321	CTGGCCCCTG	ATGACAGGAG	CTACAAGAGC	CAGTACCTGA	ACAATGGCCC
		CCAGAGGATT				
	4381	GGCAGGAAGT	ACAAGAAGGT	CAGGTTCATG	GCCTACACTG	ATGAAACCTT
		CAAGACCAGG				
	4441	GAGGCCATCC	AGCATGAGTC	TGGCATCCTG	GGCCCCCTGC	TGTATGGGGA
10		GGTGGGGGAC				
	4501	ACCCTGCTGA	TCATCTTCAA	GAACCAGGCC	AGCAGGCCCT	ACAACATCTA
		CCCCCATGGC				
	4561	ATCACTGATG	TGAGGCCCT	GTACAGCAGG	AGGCTGCCCA	AGGGGGTGAA
		GCACCTGAAG				
15	4621	GACTTCCCCA	TCCTGCCTGG	GGAGATCTTC	AAGTACAAGT	GGACTGTGAC
		TGTGGAGGAT				
	4681	GGCCCCACCA	AGTCTGACCC	CAGGTGCCTG	ACCAGATACT	ACAGCAGCTT
		TGTGAACATG				
	4741	GAGAGGGACC	TGGCCTCTGG	CCTGATTGGC	CCCCTGCTGA	TCTGCTACAA
20		GGAGTCTGTG				
	4801	GACCAGAGGG	GCAACCAGAT	CATGTCTGAC	AAGAGGAATG	TGATCCTGTT
		CTCTGTGTTT				
	4861	GATGAGAACA	GGAGCTGGTA	CCTGACTGAG	AACATCCAGA	GGTTCCTGCC
		CAACCCTGCT				
25	4921	GGGGTGCAGC	TGGAGGACCC	TGAGTTCCAG	GCCAGCAACA	TCATGCACAG
		CATCAATGGC				
	4981	TATGTGTTTG	ACAGCCTGCA	GCTGTCTGTG	TGCCTGCATG	AGGTGGCCTA
		CTGGTACATC				
	5041	CTGAGCATTG	GGGCCCAGAC	TGACTTCCTG	TCTGTGTTCT	TCTCTGGCTA
30		CACCTTCAAG				
	5101	CACAAGATGG	TGTATGAGGA	CACCCTGACC	CTGTTCCCCT	TCTCTGGGGA
		GACTGTGTTT				
	5161	ATGAGCATGG	AGAACCCTGG	CCTGTGGATT	CTGGGCTGCC	ACAACCTCTGA
		CTTCAGGAAC				
35	5221	AGGGGCATGA	CTGCCCTGCT	GAAAGTCTCC	AGCTGTGACA	AGAACACTGG
		GGACTACTAT				
	5281	GAGGACAGCT	ATGAGGACAT	CTCTGCCTAC	CTGCTGAGCA	AGAACAATGC
		CATTGAGCCC				
	5341	AGGAGCTTCA	GCCAGAACAG	CAGGCACCCC	AGCACCAGGC	AGAAGCAGTT
40		CAATGCCACC				
	5401	ACCATCCCTG	AGAATGACAT	AGAGAAGACA	GACCCATGGT	TTGCCACCG
		GACCCCATG				
	5461	CCCAAGATCC	AGAATGTGAG	CAGCTCTGAC	CTGCTGATGC	TGCTGAGGCA
		GAGCCCCACC				
45	5521	CCCCATGGCC	TGAGCCTGTC	TGACCTGCAG	GAGGCCAAGT	ATGAAACCTT
		CTCTGATGAC				
	5581	CCCAGCCCTG	GGGCCATTGA	CAGCAACAAC	AGCCTGTCTG	AGATGACCCA
		CTTCAGGCC				
	5641	CAGCTGCACC	ACTCTGGGGA	CATGGTGTTT	ACCCCTGAGT	CTGGCCTGCA
50		GCTGAGGCTG				
	5701	AATGAGAAGC	TGGGCACCAC	TGCTGCCACT	GAGCTGAAGA	AGCTGGACTT
		CAAAGTCTCC				
	5761	AGCACCAGCA	ACAACCTGAT	CAGCACCATC	CCCTCTGACA	ACCTGGCTGC
		TGGCACTGAC				
55	5821	AACACCAGCA	GCCTGGGCCC	CCCCAGCATG	CCTGTGCACT	ATGACAGCCA
		GCTGGACACC				
	5881	ACCCTGTTTG	GCAAGAAGAG	CAGCCCCCTG	ACTGAGTCTG	GGGGCCCCCT
		GAGCCTGTCT				
	5941	GAGGAGAACA	ATGACAGCAA	GCTGCTGGAG	TCTGGCCTGA	TGAACAGCCA
60		GGAGAGCAGC				

	6001	TGGGGCAAGA	ATGTGAGCAG	CAGGGAGATC	ACCAGGACCA	CCCTGCAGTC
		TGACCAGGAG				
	6061	GAGATTGACT	ATGATGACAC	CATCTCTGTG	GAGATGAAGA	AGGAGGACTT
		TGACATCTAC				
5	6121	GACGAGGACG	AGAACCAGAG	CCCCAGGAGC	TTCCAGAAGA	AGACCAGGCA
		CTACTTCATT				
	6181	GCTGCTGTGG	AGAGGCTGTG	GGACTATGGC	ATGAGCAGCA	GCCCCCATGT
		GCTGAGGAAC				
	6241	AGGGCCCAGT	CTGGCTCTGT	GCCCCAGTTC	AAGAAGGTGG	TGTTCCAGGA
10		GTTCACTGAT				
	6301	GGCAGCTTCA	CCCAGCCCCT	GTACAGAGGG	GAGCTGAATG	AGCACCTGGG
		CCTGCTGGGC				
	6361	CCCTACATCA	GGGCTGAGGT	GGAGGACAAC	ATCATGGTGA	CCTTCAGGAA
		CCAGGCCAGC				
15	6421	AGGCCCTACA	GCTTCTACAG	CAGCCTGATC	AGCTATGAGG	AGGACCAGAG
		GCAGGGGGCT				
	6481	GAGCCCAGGA	AGAACTTTGT	GAAGCCCAAT	GAAACCAAGA	CCTACTTCTG
		GAAGGTGCAG				
	6541	CACCACATGG	CCCCACCAA	GGATGAGTTT	GACTGCAAGG	CCTGGGCCTA
20		CTTCTCTGAT				
	6601	GTGGACCTGG	AGAAGGATGT	GCACTCTGGC	CTGATTGGCC	CCCTGCTGGT
		GTGCCACACC				
	6661	AACACCCTGA	ACCCTGCCCA	TGGCAGGCAG	GTGACTGTGC	AGGAGTTTGC
		CCTGTTCTTC				
25	6721	ACCATCTTTG	ATGAAACCAA	GAGCTGGTAC	TTCACTGAGA	ACATGGAGAG
		GAAGTGCAGG				
	6781	GCCCCCTGCA	ACATCCAGAT	GGAGGACCCC	ACCTTCAAGG	AGAACTACAG
		GTTCCATGCC				
	6841	ATCAATGGCT	ACATCATGGA	CACCCTGCCT	GGCCTGGTGA	TGGCCCAGGA
30		CCAGAGGATC				
	6901	AGGTGGTACC	TGCTGAGCAT	GGGCAGCAAT	GAGAACATCC	ACAGCATCCA
		CTTCTCTGGC				
	6961	CATGTGTTCA	CTGTGAGGAA	GAAGGAGGAG	TACAAGATGG	CCCTGTACAA
		CCTGTACCCT				
35	7021	GGGGTGTTTG	AGACTGTGGA	GATGCTGCCC	AGCAAGGCTG	GCATCTGGAG
		GGTGGAGTGC				
	7081	CTGATTGGGG	AGCACCTGCA	TGCTGGCATG	AGCACCCCTGT	TCCTGGTGTA
		CAGCAACAAG				
	7141	TGCCAGACCC	CCCTGGGCAT	GGCCTCTGGC	CACATCAGGG	ACTTCCAGAT
40		CACTGCCTCT				
	7201	GGCCAGTATG	GCCAGTGGGC	CCCCAAGCTG	GCCAGGCTGC	ACTACTCTGG
		CAGCATCAAT				
	7261	GCCTGGAGCA	CCAAGGAGCC	CTTCAGCTGG	ATCAAGGTGG	ACCTGCTGGC
		CCCCATGATC				
45	7321	ATCCATGGCA	TCAAGACCCA	GGGGGCCAGG	CAGAAGTTCA	GCAGCCTGTA
		CATCAGCCAG				
	7381	TTCATCATCA	TGTACAGCCT	GGATGGCAAG	AAGTGGCAGA	CCTACAGGGG
		CAACAGCACT				
	7441	GGCACCCCTGA	TGGTGTTCTT	TGGCAATGTG	GACAGCTCTG	GCATCAAGCA
50		CAACATCTTC				
	7501	AACCCCCCA	TCATTGCCAG	ATACATCAGG	CTGCACCCCA	CCCACTACAG
		CATCAGGAGC				
	7561	ACCCTGAGGA	TGGAGCTGAT	GGGCTGTGAC	CTGAACAGCT	GCAGCATGCC
		CCTGGGCATG				
55	7621	GAGAGCAAGG	CCATCTCTGA	TGCCCAGATC	ACTGCCAGCA	GCTACTTCAC
		CAACATGTTT				
	7681	GCCACCTGGA	GCCCCAGCAA	GGCCAGGCTG	CACCTGCAGG	GCAGGAGCAA
		TGCCTGGAGG				
	7741	CCCAGGTCA	ACAACCCCAA	GGAGTGGCTG	CAGGTGGACT	TCCAGAAGAC
60		CATGAAGGTG				

	7801	ACTGGGGTGA	CCACCCAGGG	GGTGAAGAGC	CTGCTGACCA	GCATGTATGT
		GAAGGAGTTC				
	7861	CTGATCAGCA	GCAGCCAGGA	TGGCCACCAG	TGGACCCTGT	TCTTCCAGAA
		TGGCAAGGTG				
5	7921	AAGGTGTTCC	AGGGCAACCA	GGACAGCTTC	ACCCCTGTGG	TGAACAGCCT
		GGACCCCCC				
	7981	CTGCTGACCA	GATACCTGAG	GATTCACCCC	CAGAGCTGGG	TGCACCAGAT
		TGCCCTGAGG				
	8041	ATGGAGGTGC	TGGGCTGTGA	GGCCCAGGAC	CTGTACTGAG	CGGCCGCGGG
10		CCCAATCAAC				
	8101	CTCTGGATTA	CAAAATTTGT	GAAAGATTGA	CTGGTATTCT	TAACATATGTT
		GCTCCTTTTA				
	8161	CGCTATGTGG	ATACGCTGCT	TTAATGCCTT	TGTATCATGC	TATTGCTTCC
		CGTATGGCTT				
15	8221	TCATTTTCTC	CTCCTTGTAT	AAATCCTGGT	TGCTGTCTCT	TTATGAGGAG
		TTGTGGCCCG				
	8281	TTGTCAGGCA	ACGTGGCGTG	GTGTGCACTG	TGTTTGCTGA	CGCAACCCCC
		ACTGGTTGGG				
	8341	GCATTGCCAC	CACCTGTCAG	CTCCTTTCCG	GGACTTTCGC	TTTCCCCCTC
20		CCTATTGCCA				
	8401	CGGCGGAACT	CATCGCCGCC	TGCCTTGCCC	GCTGCTGGAC	AGGGGCTCGG
		CTGTTGGGCA				
	8461	CTGACAATTC	CGTGGTGTG	TCGGGGAAAT	CATCGTCCTT	TCCTTGGCTG
		CTCGCCTGTG				
25	8521	TTGCCACCTG	GATTCTGCGC	GGGACGTCCT	TCTGCTACGT	CCCTTCGGCC
		CTCAATCCAG				
	8581	CGGACCTTCC	TTCCCGCGGC	CTGCTGCCGG	CTCTGCGGCC	TCTTCCGCGT
		CTTCGCCTTC				
	8641	GCCCTCAGAC	GAGTCGGATC	TCCCTTTGGG	CCGCCTCCCC	GCAAGCTTCG
30		CACTTTTTAA				
	8701	AAGAAAAGGG	AGGACTGGAT	GGGATTTATT	ACTCCGATAG	GACGCTGGCT
		TGTAACCTAG				
	8761	TCTCTTACTA	GGAGACCAGC	TTGAGCCTGG	GTGTTGCTG	GTTAGCCTAA
		CCTGGTTGGC				
35	8821	CACCAGGGGT	AAGGACTCCT	TGGCTTAGAA	AGCTAATAAA	CTTGCCTGCA
		TTAGAGCTCT				
	8881	TACGCGTCCC	GGGCTCGAGA	TCCGCATCTC	AATTAGTCAG	CAACCATAGT
		CCCGCCCCTA				
	8941	ACTCCGCCCA	TCCCGCCCCT	AACTCCGCCC	AGTTCCGCCC	ATTCTCCGCC
40		CCATGGCTGA				
	9001	CTAATTTTTT	TTATTTATGC	AGAGGCCGAG	GCCGCCTCGG	CCTCTGAGCT
		ATTCCAGAAG				
	9061	TAGTGAGGAG	GCTTTTTTGG	AGGCCTAGGC	TTTTGCAAAA	AGCTAACTTG
		TTTATTGCAG				
45	9121	CTTATAATGG	TTACAAATAA	AGCAATAGCA	TCACAAATTT	CACAAATAAA
		GCATTTTTTT				
	9181	CACTGCATTC	TAGTTGTGGT	TTGTCCAAAC	TCATCAATGT	ATCTTATCAT
		GTCTGTCCGC				
	9241	TTCCTCGCTC	ACTGACTCGC	TGCGCTCGGT	CGTTCGGCTG	CGGCGAGCGG
50		TATCAGCTCA				
	9301	CTCAAAGGCG	GTAATACGGT	TATCCACAGA	ATCAGGGGAT	AACGCAGGAA
		AGAACATGTG				
	9361	AGCAAAAGGC	CAGCAAAAGG	CCAGGAACCG	TAAAAAGGCC	GCGTTGCTGG
		CGTTTTTCCA				
55	9421	TAGGCTCCGC	CCCCCTGACG	AGCATCACAA	AAATCGACGC	TCAAGTCAGA
		GGTGGCGAAA				
	9481	CCCGACAGGA	CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG
		TGCGCTCTCC				
	9541	TGTTCCGACC	CTGCCGCTTA	CCGGATACCT	GTCCGCCTTT	CTCCCTTCGG
60		GAAGCGTGGC				

	9601	GCTTTCTCAT	AGCTCACGCT	GTAGGTATCT	CAGTTCGGTG	TAGGTCGTTC
		GCTCCAAGCT				
	9661	GGGCTGTGTG	CACGAACCCC	CCGTTACAGC	CGACCGCTGC	GCCTTATCCG
		GTAACATATCG				
5	9721	TCTTGAGTCC	AACCCGGTAA	GACACGACTT	ATCGCCACTG	GCAGCAGCCA
		CTGGTAACAG				
	9781	GATTAGCAGA	GCGAGGTATG	TAGGCGGTGC	TACAGAGTTC	TTGAAGTGGT
		GGCCTAACTA				
10	9841	CGGCTACACT	AGAAGAACAG	TATTTGGTAT	CTGCGCTCTG	CTGAAGCCAG
		TTACCTTCGG				
	9901	AAAAAGAGTT	GGTAGCTCTT	GATCCGGCAA	ACAAACCACC	GCTGGTAGCG
		GTGGTTTTTT				
	9961	TGTTTGCAAG	CAGCAGATTA	CGCGCAGAAA	AAAAGGATCT	CAAGAAGATC
		CTTTGATCTT				
15	10021	TTCTACGGGG	TCTGACGCTC	AGTGGAACGA	AAACTCACGT	TAAGGGATTT
		TGGTCATGAG				
	10081	ATTATCAAAA	AGGATCTTCA	CCTAGATCCT	TTTAAATTAA	AAATGAAGTT
		TTAAATCAAT				
	10141	CTAAAGTATA	TATGAGTAAA	CTTGGTCTGA	CAGTTAGAAA	AACTCATCGA
20		GCATCAAATG				
	10201	AAACTGCAAT	TTATTCATAT	CAGGATTATC	AATACCATAT	TTTTGAAAAA
		GCCGTTTCTG				
	10261	TAATGAAGGA	GAAAACTCAC	CGAGGCAGTT	CCATAGGATG	GCAAGATCCT
		GGTATCGGTC				
25	10321	TGCGATTCCG	ACTCGTCCAA	CATCAATACA	ACCTATTAAT	TTCCCCTCGT
		CAAAAATAAG				
	10381	GTTATCAAGT	GAGAAATCAC	CATGAGTGAC	GA CTGAATCC	GGTGAGAATG
		GCAACAGCTT				
	10441	ATGCATTTCT	TTCCAGACTT	GTTCAACAGG	CCAGCCATTA	CGCTCGTCAT
30		CAAAATCACT				
	10501	CGCATCAACC	AAACCGTTAT	TCATTCGTGA	TTGCGCCTGA	GCGAGACGAA
		ATACGCGATC				
	10561	GCTGTAAAAA	GGACAATTAC	AAACAGGAAT	CGAATGCAAC	CGGCGCAGGA
		ACACTGCCAG				
35	10621	CGCATCAACA	ATATTTTCAC	CTGAATCAGG	ATATTCTTCT	AATACCTGGA
		ATGCTGTTTT				
	10681	TCCGGGGATC	GCAGTGGTGA	GTAACCATGC	ATCATCAGGA	GTACGGATAA
		AATGCTTGAT				
	10741	GGTCGGAAGA	GGCATAAATT	CCGTCAGCCA	GTTTAGTCTG	ACCATCTCAT
40		CTGTAACATC				
	10801	ATTGGCAACG	CTACCTTTGC	CATGTTTCAG	AAACAACCTCT	GGCGCATCGG
		GCTTCCCAT				
	10861	CAATCGATAG	ATTGTCGCAC	CTGATTGCCC	GACATTATCG	CGAGCCCAT
		TATACCCATA				
45	10921	TAAATCAGCA	TCCATGTTGG	AATTTAATCG	CGGCCTAGAG	CAAGACGTTT
		CCCCTTGAAT				
	10981	ATGGCTCATA	ACACCCCTTG	TATTA CTGTT	TATGTAAGCA	GACAGTTTTA
		TTGTTTCATGA				
50	11041	TGATATATTT	TTATCTTGTG	CAATGTAACA	TCAGAGATTT	TGAGACACAA
		CAATTGGTCG				
	11101	ACGGATCC				
		SEQ ID NO: 15				
		ATGCCAGCTCTGTGTCCTGGGGCATTCTGCTGCTGGCTGGCCTGTGCTGTCTGGTGCCTG				
		TGTCCCTGGCTGAGGACCCTCAGGGGGATGCTGCCAGAAAAACAGACACCTCCCACCATGACC				
55		AGGACCACCCACCTTCAACAAGATCACCCCAACCTGGCAGAGTTTGCCTTCAGCCTGTACAG				
		ACAGCTGGCCACCCAGAGCAACAGCACCAACATCTTTTTTCAGCCCTGTGTCCATTGCCACAGCC				
		TTTGCCATGCTGAGCCTGGGCACCAAGGCTGACACCCATGATGAGATCCTGGAAGGCCTGAACT				
		TCAACCTGACAGAGATCCCTGAGGCCAGATCCATGAGGGCTTCCAGGAACTGCTGAGAACCCCT				
		GAACCAGCCAGACAGCCAGCTGCAGCTGACAAACAGGCAATGGGCTGTTCTGTCTGAGGGCCT				
60		GAAGCTGGTGGACAAGTTTCTGGAAGATGTGAAGAAGCTGTACCACTCTGAGGCCTTCACAGTG				

AACTTTGGGGACACAGAAGAGGCCAAGAAACAGATCAATGACTATGTGGAAAAGGGCACCCAGG
 GCAAGATTGTGGACCTTGTGAAAGAGCTGGACAGGGGACACTGTGTTTGCCCTTGTGAACTACAT
 CTTCTTCAAGGGCAAGTGGGAGAGGCCCTTTGAAGTGAAGGACACTGAGGAAGAGGACTTCCAT
 GTGGACCAAGTGACCACAGTGAAGGTGCCAATGATGAAGAGACTGGGGATGTTCAATATCCAGC
 5 ACTGCAAGAACTGAGCAGCTGGGTGCTGCTGATGAAGTACCTGGGCAATGCTACAGCCATATT
 CTTTCTGCCTGATGAGGGCAAGCTGCAGCACCTGGAAAATGAGCTGACCCATGACATCATCACC
 AAATTTCTGGAAAATGAGGACAGAAGATCTGCCAGCCTGCATCTGCCCAAGCTGAGCATCACAG
 GCACATATGACCTGAAGTCTGTGCTGGGACAGCTGGGAATCACCAAGGTGTTTCAGCAATGGGGC
 AGACCTGAGTGGAGTGACAGAGGAAGCCCCCTCTGAAGCTGTCCAAGGCTGTGCACAAGGCAGT
 10 GCTGACCATTGATGAGAAGGGCACAGAGGCTGCTGGGGCCATGTTTCTGGAAGCCATCCCCATG
 TCCATCCCCCAGAAGTGAAGTTCAACAAGCCCTTTGTGTTCTGATGATTGAGCAGAACACCAA
 GAGCCCCCTGTTTCATGGGCAAGGTTGTGAACCCACCCAGAAATGA
 SEQ ID NO: 16
 ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGCTGCTGAGGTTCTGCTTCTCTGCCAC
 15 CAGGAGATACTACCTGGGGGCTGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGGGA
 GCTGCCTGTGGATGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCTGTG
 GTGTACAAGAAGACCCTGTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAGCCCAGGC
 CCCCCTGGATGGGCCTGCTGGGCCCCACCATCCAGGCTGAGGTGTATGACACTGTGGTGATCA
 CCCTGAAGAACATGGCCAGCCACCCTGTGAGCCTGCATGCTGTGGGGGTGAGCTACTGGAAGG
 20 CCTCTGAGGGGGCTGAGTATGATGACCAGACCAGCCAGAGGGAGAAGGAGGATGACAAGGTGT
 TCCCTGGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATGGCCCCATGGCCTCTG
 ACCCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCTGAACCTCTGG
 CCTGATTGGGGCCCTGCTGGTGTGCAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCT
 GCACAAGTTTCATCCTGCTGTTTGTGTGTTTGTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGA
 25 ACAGCCTGATGCAGGACAGGGATGCTGCCTCTGCCAGGGCCTGGCCCAAGATGCACACTGTGA
 ATGGCTATGTGAACAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCA
 TGTGATTGGCATGGGCACCACCCCTGAGGTGCACAGCATCTTCTGGAGGGCCACACCTTCTG
 GTCAGGAACCACAGGCAGGCCAGCCTGGAGATCAGCCCCATCACCTTCTGACTGCCAGACC
 CTGCTGATGGACCTGGGCCAGTTCTGCTGTTCTGCCACATCAGCAGCCACCAGCATGATGGCA
 30 TGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGAAGAACAATG
 AGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTGAGGTTTGATGA
 TGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAGCACCCCAAGACCTGGGTG
 CACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCCCTGGTGTGGCCCCCTGATGAC
 AGGAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGAAG
 35 GTCAGGTTTCATGGCCTACACTGATGAAACCTTCAAGACAGGGAGGCCATCCAGCATGAGTCTG
 GCATCCTGGGCCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACC
 AGGCCAGCAGGCCCTACAACATCTACCCCATGGCATCACTGATGTGAGGCCCTGTACAGCAG
 GAGGCTGCCCAAGGGGGTGAAGCACCTGAAGGACTTCCCCTCTGCTGGGGAGATCTTCAA
 GTACAAGTGGACTGTGACTGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCCTGACCAG
 40 ATACTACAGCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCCCTGCTG
 ATCTGCTACAAGGAGTCTGTGGACCAGAGGGGCAACCAGATCATGTCTGACAAGAGGAATGTGA
 TCCTGTTCTCTGTGTTTGTGAGAACAGGAGCTGGTACCTGACTGAGAACATCCAGAGGTTCTG
 CCCAACCCCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACAGC
 ATCAATGGCTATGTGTTTGACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGGCCTACTGGT
 45 ACATCCTGAGCATTGGGGCCCAGACTGACTTCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAG
 CACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCCTTCTCTGGGGAGACTGTGTTTCATGA
 GCATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACCTCTGACTTCAGGAACAGGGGCAT
 GACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAACTCTGGGGACTACTATGAGGACAGCTAT
 GAGGACATCTCTGCCTACCTGCTGAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGA
 50 ACAGCAGGCACCCACAGCACCAGGCAGAAGCAGTTCAATGCCACCACCATCCCTGAGAATGACAT
 AGAGAAGACAGACCCATGGTTTGGCCACCGGACCCCATGCCCAAGATCCAGAATGTGAGCAGC
 TCTGACCTGCTGATGCTGCTGAGGCAGAGCCCCACCCCATGGCCTGAGCCTGTCTGACCTG
 CAGGAGGCCAAGTATGAAACCTTCTCTGATGACCCAGCCCTGGGGCCATTGACAGCAACAACA
 GCCTGTCTGAGATGACCCACTTCAGGCCCCAGCTGCACCACTCTGGGGACATGGTGTTACCCC
 55 TGAGTCTGGCCTGCAGCTGAGGCTGAATGAGAAGCTGGGCACCACTGCTGCCACTGAGCTGAA
 GAAGCTGGACTTCAAAGTCTCCAGCACCAGCAACAACCTGATCAGCACCATCCCCTCTGACAAC
 CTGGCTGCTGGCACTGACAACACCAGCAGCCTGGGCCCCCCCCAGCATGCCTGTGCACTATGAC
 AGCCAGCTGGACACCACCCTGTTTGGCAAGAAGAGCAGCCCCCTGACTGAGTCTGGGGGCCCC
 CTGAGCCTGTCTGAGGAGAACAATGACAGCAAGCTGCTGGAGTCTGGCCTGATGAACAGCCAG
 60 GAGAGCAGCTGGGGCAAGAATGTGAGCAGCAGGGAGATCACCAGGACCACCCTGCAGTCTGAC

CAGGAGGAGATTGACTATGATGACACCATCTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTA
 CGACGAGGACGAGAACCAGAGCCCCAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCT
 GCTGTGGAGAGGCTGTGGGACTATGGCATGAGCAGCAGCCCCATGTGCTGAGGAACAGGGCC
 CAGTCTGGCTCTGTGCCCCAGTTCAAGAAGGTGGTGTTCAGGAGTTCACTGATGGCAGCTTCA
 5 CCCAGCCCCTGTACAGAGGGGAGCTGAATGAGCACCTGGGCCTGCTGGGCCCCCTACATCAGGG
 CTGAGGTGGAGGACAACATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCCTACAGCTTCTA
 CAGCAGCCTGATCAGCTATGAGGAGGACCAGAGGCAGGGGGCTGAGCCCAGGAAGAACTTTGT
 GAAGCCCAATGAAACCAAGACCTACTTCTGGAAGGTGCAGCACCACATGGCCCCACCAAGGAT
 GAGTTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCTG
 10 GCCTGATTGGCCCCCTGCTGGTGTGCCACACCAACACCCTGAACCCTGCCCATGGCAGGCAGG
 TGAATGTGCAGGAGTTTGGCCTGTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTTCACT
 GAGAACATGGAGAGGAACCTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTTCAAG
 GAGAACTACAGTTCCATGCCATCAATGGCTACATCATGGACACCCTGCCTGGCCTGGTGTGATGG
 CCCAGGACCAGAGGATCAGGTGGTACCTGCTGAGCATGGGCAGCAATGAGAACATCCACAGCA
 15 TCCACTTCTCTGGCCATGTGTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAA
 CCTGTACCCTGGGGTGTGTTGAGACTGTGGAGATGCTGCCCAGCAAGGCTGGCATCTGGAGGGT
 GGAGTGCCTGATTGGGGAGCACCTGCATGCTGGCATGAGCACCTGTTCTGCTGTACAGCAAC
 AAGTGCCAGACCCCCCTGGGCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTG
 GCCAGTATGGCCAGTGGGCCCCCAAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAATGCCT
 20 GGAGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCATG
 GCATCAAGACCCAGGGGGGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCAT
 GTACAGCCTGGATGGCAAGAAGTGGCAGACCTACAGGGGCAACAGCACTGGCACCCCTGATGGT
 GTTCTTTGGCAATGTGGACAGCTCTGGCATCAAGCACAACATCTTCAACCCCCCATCATTGCCA
 GATACATCAGGCTGCACCCCACTACAGCATCAGGAGCACCCCTGAGGATGGAGCTGATGG
 25 GCTGTGACCTGAACAGCTGCAGCATGCCCCTGGGCATGGAGAGCAAGGCCATCTCTGATGCC
 AGATCACTGCCAGCAGCTACTTCACCAACATGTTTGCCACCTGGAGCCCCAGCAAGGCCAGGCT
 GCACCTGCAGGGCAGGAGCAATGCCTGGAGGGCCCCAGGTCAACAACCCCAAGGAGTGGCTGCA
 GGTGGACTTCCAGAAGACCATGAAGGTGACTGGGGTGACCACCCAGGGGGTGAAGAGCCTGCT
 GACCAGCATGTATGTGAAGGAGTTCCTGATCAGCAGCAGCCAGGATGGCCACCAGTGGACCCT
 30 GTTCTTCCAGAATGGCAAGGTGAAGGTGTTCCAGGGCAACCAGGACAGCTTACCCCTGTGGTG
 AACAGCCTGGACCCCCCCTGCTGACCAGATACCTGAGGATTACCCCCAGAGCTGGGTGCAC
 CAGATTGCCCTGAGGATGGAGGTGCTGGGCTGTGAGGCCCAGGACCTGTACTGA

SEQ ID NO: 17

CCGCGGAGATCTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATA
 35 TTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCC
 AATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
 AGTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGAC
 CGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG
 GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAG
 40 TGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT
 GCCCAGTACATGACCTTACGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT
 ACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAATCACGGGGAT
 TTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTT
 CCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
 45 GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTAT
 CACAGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGAAG
 TTGGTCGTGAGGCACTGGGCAGGCTAGC

SEQ ID NO: 18

TCGAGATGTGGTCTGAGTTAAAAATCAGGAGCAACGACGGAGGTGAAGGACCAGACGCCAA
 50 CGACCC

SEQ ID NO: 19

CCGGGGGTCTGTTGGCGTCTGGTCCTTACCTCCGTCGTTGCTCCTGATTTTAACTCAGACC
 ACATC

SEQ ID NO: 20

CCGGGGAAAGGGGGTGAACACATCCATATCCAGCCATCTCTACCTGTTTATGGACA
 55

SEQ ID NO: 21

ACCCTCTGTCCATAAACAGGTAGAGATGGCTGGATATGGATGTGTTGCACCCCTTTCC

SEQ ID NO: 22

GGGTTAGGTGGTTGCTGATTCTCTCATTACCCAGTGGG

60 SEQ ID NO: 23

GATCCCCACTGGGTGAATGAGAGAATCAGCAACCACCTA

SEQ ID NO: 24

GAGACTCGAGATGTGGTCTGAGTTAAAAATCAGG

SEQ ID NO: 25

5 AGAGGTAGACCAGTACGAGTCACGTTTGCCCCTATCACCATCCCTAACCCCTCTGTCATAAAC

SEQ ID NO: 26

TACGGGTCGAGACACAGGACCCCGTAAGACGACGACCGACCGGACACGACAGACCACGGA
CACAGGGACCGACTCCTGGGAGTCCCCCTACGACGGGTCTTTTGTCTGTGGAGGGTGGTACTG
GTCCTGGTGGGGTGGAAAGTTGTTCTAGTGGGGGTTGGACCGTCTCAAACGGAAGTCGGACATGT
10 CTGTCGACCGGGTGGTCTCGTTGTCGTGGTTGTAGAAAAAGTCGGGACACAGGTAACGGTGTCTG
GAAACGGTACGACTCGGACCCGTGGTTCGACTGTGGGTACTACTCTAGGACCTTCCGGACTTG
AAGTTGGACTGTCTCTAGGGACTCCGGGTCTAGGTACTCCCGAAGGTCTTGACGACTCTTGGG
ACTTGGTCTGGTCTGTCTGGTTCGACGTGACTGTTGTCCGTTACCCGACAAGGACAGACTCCCGGA
CTTCGACCACCTGTTCAAAGACCTTCTACACTTCTTCGACATGGTGAGACTCCGGAAGTGTCACT
15 TGAAACCCCTGTGTCTTCTCCGGTTCTTTGTCTAGTTACTGATACACCTTTTCCCGTGGGTCCCGT
TCTAACACCTGGAACACTTTCTCGACCTGTCCCTGTGACACAAACGGGAACACTTGATGTAGAAG
AAGTTCCCGTTCCACCTCTCCGGGAAACTTCACCTTCTGTGACTCCTTCTCCTGAAGGTACACCT
GGTTCACCTGGTGTCACTTCCACGGTTACTACTTCTCTGACCCCTACAAGTTATAGGTCTGTGACGT
TCTTTGACTCGTCGACCCACGACGACTACTTCATGGACCCGTTACGATGTCCGTATAAGAAAGAC
20 GGACTACTCCCGTTTCGACGTCTGGACCTTTTACTCGACTGGGTACTGTAGTAGTGGTTTAAAGA
CCTTTTACTCCTGTCTTCTAGACGGTTCGGACGTAGACGGGTTTCGACTCGTAGTGTCCGTGTATAC
TGGACTTCAGACACGACCCTGTCTGACCCCTAGTGGTTCACAAAGTCGTTACCCCGTCTGGACTC
ACCTCACTGTCTCCTTCGGGGAGACTTCGACAGGTTCCGACACGTGTTCCGTCACGACTGGTAA
CTACTCTTCCCGTGTCTCCGACGACCCCGGTACAAAGACCTTCGGTAGGGGTACAGGTAGGGG
25 GGTCTTCACTTCAAGTTGTTCCGGGAAACACAAGGACTACTAACTCGTCTTGTGGTTCTCGGGGGA
CAAGTACCCGTTCCAACACTTGGGGTGGGTCTTTACT

SEQ ID NO: 27

AEDPQGDAQAQKTDTSHHQDQDHPTFAEDPQGDAQAQKTDTSHHQDQDHPTFNKITPNLAEFAFLYR
QLAHQSNSTNIFFSPVSIATAFAMLSLGTADTHDEILEGLNFNLTIEPEAQIHEGFQELLRTLNPDSQ
30 LQLTTGNGLFLSEGLKLVDFLEDVKKLYHSEFTVNFQDTEEAQKQINDYVEKGTQGGKIVDLVKELD
RDTVFALVNIFFKQKWERPFVVDTEEDFHVDQVTTVKVPMKRLGMFNIQHCKKLSSWVLLMKY
LGNATAIFFLPDEGLQHLLENELTHDIITKFLNEDRRSASLHLPKLSITGTLDKSVLGQLGITKVFNSG
ADLSGVTEEAPLKLKSAVHKAVLTIDEKGTAAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLF
MGKVVNPTQK

SEQ ID NO: 28

TACGTCTAACTCGACTCGTGGACGAAGAAGGACACGGACGACTCCAAGACGAAGAGACGGT
GGTCTCTA
TGATGGACCCCCGACACCTCGACTCGACCCTGATGTACGTCAGACTGGACCCCCCTCGACGG
ACACCTACG
40 GTCCAAGGGGGGGTCTCACGGGTTCTCGAAGGGGAAGTTGTGGAGACACCACATGTTCTTC
TGGGACAAA
CACCTCAAGTGACTGGTGGACAAGTTGTAACGGTTCGGGTCCGGGGGGGACCTACCCGGACG
ACCCGGGGT
GGTAGGTCCGACTCCACATACTGTGACACCACTAGTGGGACTTCTTGTACCGGTCCGTGGGA
45 CACTCGGA
CGTACGACACCCCCACTCGATGACCTTCCGGAGACTCCCCCGACTCATACTACTGGTCTGGT
CGGTCTCC
CTCTTCTCCTACTGTTCCACAAGGGACCCCGTCCGGTGTGGATACACACCGTCCACGACTT
CCTCTTAC
50 CGGGGTACCGGAGACTGGGGGACACGGACTGGATGTCGATGGACTCGGTACACCTGGACC
ACTTCTCTGA
CTTGAGACCGGACTAACCCCGGGACGACCACACGTCCCTCCCGTCGGACCGGTTCTCTTCTC
TGGGTCTGG
GACGTGTTCAAGTAGGACGACAAACGACACAACTACTCCCGTTCTCGACCGTGAGACTTTG
55 GTTCTTGT
CGGACTACGTCCTGTCCCTACGACGGAGACGGTCCCGGACCGGGTCTACGTGTGACACTT
ACCGATACA
CTTGTCTCTCGGACGGACCGGACTAACCGACGGTGTCTTTCAGACACATGACCGTACACTAAC
CGTACCCG

TGGTGGGGACTCCACGTGTCGTAGAAAGGACCTCCCGGTGTGGAAGGACCAGTCCTTGGTGT
 CCGTCCGGT
 CGGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGGACGACTACCTGGACCCGG
 TCAAGGACGA
 5 CAAGACGGTGTAGTCGTCTGGTGGTCTACTACCGTACCTCCGGATACACTTCCACCTGTCGA
 CGGGACTC
 CTCGGGGTCTGACTCCTACTTCTTGTTACTCCTCCGACTCCTGATACTACTACTGGACTGACTG
 AGACTCT
 ACCTACACCACTCCAACTACTACTGTTGTCGGGGTCTGAAGTAGGTCTAGTCCAGACACCGG
 10 TTCTTCGT
 GGGGTTCTGGACCCACGTGATGTAACGACGACTCCTCCTCCTGACCCTGATACGGGGGGAC
 CACGACCGG
 GGACTACTGTCCTCGATGTTCTCGGTCTGACTTGTACCGGGGGTCTCCTAACCGTCCTT
 CATGTTCT
 15 TCCAGTCCAAGTACCGGATGTGACTACTTTGGAAGTTCTGGTCCCTCCGGTAGGTCTGACTC
 AGACCGTA
 GGACCCGGGGGACGACATACCCCTCCACCCCTGTGGGACGACTAGTAGAAGTTCTTGGTC
 CGGTCGTCC
 GGGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCTCCGACG
 20 GGTTCCCCC
 ACTTCGTGGACTTCCTGAAGGGGTAGGACGGACCCCTCTAGAAGTTCATGTTACCTGACAC
 TGACACCT
 CCTACCGGGGTGGTTCAGACTGGGGTCCACGGACTGGTCTATGATGTCGTCGAAACACTTGT
 ACCTCTCC
 25 CTGGACCGGAGACCGGACTAACCGGGGGACGACTAGACGATGTTCTCAGACACCTGGTCT
 CCCCCTTG
 TCTAGTACAGACTGTTCTCCTTACACTAGGACAAGAGACACAACTACTCTTGTCTCGACCA
 TGGACTG
 ACTCTTGTAGGTCTCCAAGGACGGGTTGGGACGACCCACGTCGACCTCCTGGGACTCAAG
 30 GTCCGGTCG
 TTGTAGTACGTGTCGTAGTTACCGATACACAACTGTCTGGACGTCGACAGACACACGGACGT
 ACTCCACC
 GGATGACCATGTAGGACTCGTAACCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACC
 GATGTGGAA
 35 GTTCGTGTTCTACCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACA
 AGTACTCG
 TACCTCTTGGGACCGGACACCTAAGACCCGACGGTGTGAGACTGAAGTCCTTGTCCCCGTA
 CTGACGGG
 ACGACTTTCAGAGGTGACACTGTTCTTGTGACCCCTGATGATACTCCTGTCGATACTCCTGT
 40 AGAGACG
 GATGGACGACTCGTTCTTGTACGGTAACTCGGGTCTCGAAGTCGGTCTTGTCTGTCGGTGG
 GGTCGTGG
 TCCGTCTTCGTCAAGTTACGGTGGTGGTAGGGACTCTTACTGTATCTCTTGTCTGGGTACC
 AAACGGG
 45 TGGCCTGGGGGTACGGGTTCTAGGTCTTACACTCGTCGAGACTGGACGACTACGACGACTC
 CGTCTCGGG
 GTGGGGGGTACCGGACTCGGACAGACTGGACGTCCTCCGGTTCATACTTTGGAAGAGACTA
 CTGGGGTCTG
 GGACCCCGGTAACCTGTCGTTGTTGTCTGGACAGACTCTACTGGGTGAAGTCCGGGGTCTGACG
 50 TGGTGAGAC
 CCCTGTACCACAAGTGGGGACTCAGACCGGACGTCGACTCCGACTTACTCTTCGACCCGTG
 GTGACGACG
 GTGACTCGACTTCTTCGACCTGAAGTTTCAGAGGTCGTGGTCGTTGTTGGACTAGTCGTGGT
 AGGGGAGA
 55 CTGTTGGACCGACGACCGTGACTGTTGTGGTCGTCTGGACCCGGGGGGTCTGACGGACAC
 GTGATACTGT
 CGGTCGACCTGTGGTGGGACAAACCGTTCTTCTCGTCGGGGGACTGACTCAGACCCCGGG
 GGACTCGGA
 CAGACTCCTCTTGTTACTGTCGTTCTGACGACCTCAGACCGGACTACTTGTCTGGTCCTCTCGT
 60 CGACCCCG

TTCTTACACTCGTCGTCCCTCTAGTGGTCCTGGTGGGACGTCAGACTGGTCCTCCTCTAACT
 GATACTAC
 TGTGGTAGAGACACCTCTACTTCTTCCTCCTGAAACTGTAGATGCTGCTCCTGCTCTTGGTCT
 CGGGGTC
 5 CTCGAAGGTCTTCTTCTGGTCCGTGATGAAGTAACGACGACACCTCTCCGACACCCTGATAC
 CGTACTCG
 TCGTCGGGGGTACACGACTCCTTGTCCCGGGTCAGACCGAGACACGGGGTCAAGTTCTTCC
 ACCACAAGG
 TCCTCAAGTGACTACCGTCGAAGTGGGTGCGGGACATGTCTCCCCTCGACTTACTCGTGGAC
 10 CCGGACGA
 CCCGGGGATGTAGTCCCGACTCCACCTCCTGTTGTAGTACCACTGGAAGTCCTTGGTCCGGT
 CGTCCGGG
 ATGTCGAAGATGTCGTGCGACTAGTCGATACTCCTCCTGGTCTCCGTCCCCCGACTCGGGTC
 CTTCTTGA
 15 AACACTTCGGGTACTTTGGTTCTGGATGAAGACCTTCCACGTCGTGGTGTACCGGGGGTGG
 TTCTACT
 CAACTGACGTTCCGGACCCGGATGAAGAGACTACACCTGGACCTCTTCTACACGTGAGAC
 CGGACTAA
 CCGGGGGACGACCACACGGTGTGGTTGTGGGACTTGGGACGGGTACCGTCCGTCCACTGA
 20 CACGTCCTCA
 AACGGGACAAGAAGTGGTAGAACTACTTTGGTTCTCGACCATGAAGTGACTCTTGTACCTCT
 CCTTGAC
 GTCCCGGGGGACGTTGTAGGTCTACCTCCTGGGGTGGGAAGTTCCTCTTGATGTCCAAGGTA
 CGGTAGTTA
 25 CCGATGTAGTACCTGTGGGACGGACCGGACCACTACCGGGTCCTGGTCTCCTAGTCCACCA
 TGGACGACT
 CGTACCCGTCGTTACTCTTGTAGGTGTCTAGGTGAAGAGACCGGTACACAAGTGACACTCC
 TTCTTCT
 CCTCATGTTCTACCGGGACATGTTGGACATGGGACCCACAACTCTGACACCTCTACGACG
 30 GGTCGTT
 CGACCGTAGACCTCCACCTCACGGACTAACCCTCGTGGACGTACGACCGTACTCGTGGG
 ACAAGGACC
 ACATGTCGTTGTTACGGTCTGGGGGGACCCGTACCGGAGACCGGTGTAGTCCCTGAAGGT
 CTAGTGACG
 35 GAGACCGGTCATACCGGTCACCCGGGGGTTTCGACCGGTCCGACGTGATGAGACCGTCGTA
 GTTACGGACC
 TCGTGGTTCCTCGGGAAGTCGACCTAGTTCCACCTGGACGACCGGGGGTACTAGTAGGTAC
 CGTAGTTCT
 GGGTCCCCCGGTCCGTCTTCAAGTCGTGCGGACATGTAGTCGGTCAAGTAGTAGTACATGTG
 40 GACCTACC
 GTTCTTACCGTCTGGATGTCCCGTTGTCTGACCGTGGGACTACCACAAGAAACCGTTAC
 ACCTGTG
 AGACCGTAGTTCGTGTTGTAGAAGTTGGGGGGTAGTAACGGTCTATGTAGTCCGACGTGG
 GGTGGGTGA
 45 TGTGCTAGTCCTCGTGGGACTCCTACCTCGACTACCCGACACTGGACTTGTGACGTCGTAC
 GGGGACCC
 GTACCTCTCGTTCCGGTAGAGACTACGGGTCTAGTGACGGTCGTGATGAAGTGGTTGTACA
 AACGGTGG
 ACCTCGGGGTCGTTCCGGTCCGACGTGGACGTCCCGTCCTCGTTACGGACCTCCGGGGTCC
 50 AGTTGTTGG
 GGTTCTCACCAGCTCCACCTGAAGGTCTTCTGGTACTTCCACTGACCCCACTGGTGGGTG
 CCCCCTT
 CTCGGACGACTGGTCGTACATACACTTCCTCAAGGACTAGTCGTGTCGGTCCTACCGGTGG
 TCACCTGG
 55 GACAAGAAGGTCTTACCGTTCCACTCCACAAGGTCCCGTTGGTCTGTGCAAGTGGGGACA
 CCACTTGT
 CGGACCTGGGGGGGGACGACTGGTCTATGGACTCCTAAGTGGGGGTCTCGACCCACGTGG
 TCTAACGGGA
 CTCCTACCTCCACGACCCGACACTCCGGGTCTGGACATGACT
 60 SEQ ID NO: 29

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRPVKSPFPNTSVVY
 KKTLFVEFTDHLFNIAPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
 DDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCR
 EGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIG
 5 CHRKSVYWHVIGMGTTPREVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQ
 HDGMEAYVKVDSCEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNPSFIQIRSVAKKHPKTW
 VHYIAAEEEDWDYAPLV LAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILG
 PLYGEVGDTLIIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDG
 PTKSDPRCLTRYSSFFVNMERDLASGLIGPLLCYKESVDQQRGNQIMSDKRNVLFSVFDENRSWYLT
 10 ENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYT
 FKHKMVYEDTLTLFPFSGETVFMSENPGLWLGCNNSDFRNRGMTALLKVSSCDKNTGDYYESDY
 EDISAYLLSKNNAIEPRFSQNSRHPSTRQKQFNATTIPENDIEKTDPFWAHRTMPMKIQNVSSDILLM
 LLRQSPTPHGLSLSDLQEAKEYETFSDDPSGPAIDSNNSLSEMTDFRPQLHHSMDVLTTPESGLQLRL
 NEKLGTTAATELKKLDFKVSSTSNNLISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDITTLFGKKSSP
 15 LTESGGPLSLSEENNSDKLLESGLMNSQESSWGKNVSSREITRTTLQSDQEEIDYDDTISVEMKKEDF
 DIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVQFKKVVFQEFTDGSF
 TQPLYRGELNEHLGLLPYIRAEVEDNIMVTFRNQASRPYSFYSSLSIYEEDQRQGAEPKRNFKPNE
 TKTYFWKVQHMAPTKDEFDCWAYFSDVDLEKDVHSGLIGPLLCHTNTLNPAHGRQVTVQEFA
 LFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQIRWYLL
 20 SMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTL
 FLVYSNKCQTPGLMASGHIRDFQITASGQYQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMII
 HGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRL
 HPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNA
 WRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLTSMYVKEFLISSSQDGHQWTLFFQNGKVQVF
 25 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY

SEQ ID NO: 30

ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCAC
 CAGGAGATACTACCTGGGGGCTGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGGGGA
 GCTGCCTGTGGATGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCCTTCAACACCTCTGTG
 30 GTGTACAAGAAGACCCTGTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAGCCCAGGC
 CCCCCTGGATGGGCTGCTGGGCCCCACCATCCAGGCTGAGGTGTATGACACTGTGGTGATCA
 CCCTGAAGAACATGGCCAGCCACCCTGTGAGCCTGCATGCTGTGGGGGTGAGCTACTGGAAGG
 CCTCTGAGGGGGCTGAGTATGATGACCAGACCAGCCAGAGGGAGAAGGAGGATGACAAGGTGT
 TCCCTGGGGGCAGCCACACCTATGTGTGGCAGGTGCTGAAGGAGAATGGCCCCATGGCCTCTG
 35 ACCCCCTGTGCCTGACCTACAGCTACCTGAGCCATGTGGACCTGGTGAAGGACCTGACCTCTGG
 CCTGATTGGGGCCCTGCTGGTGTGAGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCT
 GCACAAGTTCATCCTGCTGTTTGTGTGTTTGTATGAGGGCAAGAGCTGGCACTCTGAAACCAAGA
 ACAGCCTGATGCAGGACAGGGATGCTGCCTCTGCCAGGGCCTGGCCCAAGATGCACACTGTGA
 ATGGCTATGTGAACAGGAGCCTGCCTGGCCTGATTGGCTGCCACAGGAAGTCTGTGTACTGGCA
 40 TGTGATTGGCATGGGCACCAACCCTGAGGTGCACAGCATCTTCTGGAGGGCCACACCTTCTG
 GTCAGGAACCACAGGCAGGCCAGCCTGGAGATCAGCCCCATCACCTTCTGACTGCCAGACC
 CTGCTGATGGACCTGGGCCAGTTCTGCTGTTCTGCCACATCAGCAGCCACCAGCATGATGGCA
 TGGAGGCCTATGTGAAGGTGGACAGCTGCCCTGAGGAGCCCCAGCTGAGGATGAAGAACAATG
 AGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTGAGGTTTGTGA
 45 TGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAGCACCCCAAGACCTGGGTG
 CACTACATTGCTGCTGAGGAGGAGGACTGGGACTATGCCCCCCTGGTGTGGCCCCCTGATGAC
 AGGAGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGAGGATTGGCAGGAAGTACAAGAAG
 GTCAGGTTTCATGGCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATCCAGCATGAGTCTG
 GCATCCTGGGGCCCCCTGCTGTATGGGGAGGTGGGGGACACCCTGCTGATCATCTTCAAGAACC
 50 AGGCCAGCAGGCCCTACAACATCTACCCCATGGCATCACTGATGTGAGGCCCTGTACAGCAG
 GAGGCTGCCCAAGGGGGTGAAGCACCTGAAGGACTTCCCCTCTGCCTGGGGAGATCTTCAA
 GTACAAGTGGACTGTGACTGTGGAGGATGGCCCCACCAAGTCTGACCCCAGGTGCCTGACCAG
 ATACTACAGCAGCTTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCCCCTGCTG
 55 TCCTGTTCTCTGTGTTTGTATGAGAACAGGAGCTGGTACCTGACTGAGAACATCCAGAGGTTCTG
 CCCAACCCCTGCTGGGGTGCAGCTGGAGGACCCTGAGTTCCAGGCCAGCAACATCATGCACAGC
 ATCAATGGCTATGTGTTTGACAGCCTGCAGCTGTCTGTGTGCCTGCATGAGGTGGCCTACTGGT
 ACATCCTGAGCATTGGGGCCCAGACTGACTTCTGTCTGTGTTCTTCTCTGGCTACACCTTCAAG
 CACAAGATGGTGTATGAGGACACCCTGACCCTGTTCCCTTCTCTGGGGAGACTGTGTTTCATGA
 60 GCATGGAGAACCCTGGCCTGTGGATTCTGGGCTGCCACAACCTCTGACTTCAGGAACAGGGGCAT

GACTGCCCTGCTGAAAGTCTCCAGCTGTGACAAGAACAACACTGGGGACTACTATGAGGACAGCTAT
 GAGGACATCTCTGCCTACCTGCTGAGCAAGAACAATGCCATTGAGCCCAGGAGCTTCAGCCAGA
 ATGCCACTAATGTGTCTAACAACAGCAACACCAGCAATGACAGCAATGTGTCTCCCCCAGTGCTG
 AAGAGGCACCAGAGGGAGATCACCAGGACCACCCTGCAGTCTGACCAGGAGGAGATTGACTAT
 5 GATGACACCATCTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTACGACGAGGACGAGAACC
 AGAGCCCCAGGAGCTTCCAGAAGAAGACCAGGCACTACTTCATTGCTGCTGTGGAGAGGCTGTG
 GGACTATGGCATGAGCAGCAGCCCCCATGTGCTGAGGAACAGGGCCCAGTCTGGCTCTGTGCC
 CCAGTTCAAGAAGGTGGTGTTCAGGAGTTCACTGATGGCAGCTTCACCCAGCCCCCTGTACAGA
 GGGGAGCTGAATGAGCACCTGGGCCTGCTGGGCCCTACATCAGGGCTGAGGTGGAGGACAAC
 10 ATCATGGTGACCTTCAGGAACCAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCT
 ATGAGGAGGACCAGAGGCAGGGGGCTGAGCCCAGGAAGAAGTGTGAAGCCCAATGAAACCA
 AGACCTACTTCTGGAAGGTGCAGCACCATGAGCCCCACCAAGGATGAGTTTGACTGCAAGGC
 CTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCTGGCCTGATTGGCCCCCTG
 CTGGTGTGCCACACCAACACCCTGAACCCTGCCATGGCAGGCAGGTGACTGTGCAGGAGTTT
 15 GCCCTGTTCTTCACCATCTTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGAGGAA
 CTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTTCAAGGAGAACTACAGGTTCCAT
 GCCATCAATGGCTACATCATGGACACCCTGCCTGGCCTGGTATGGCCCAGGACCAGAGGATCA
 GGTGGTACCTGCTGAGCATGGGCAGCAATGAGAACATCCACAGCATCCACTTCTCTGGCCATGT
 GTTCACTGTGAGGAAGAAGGAGGAGTACAAGATGGCCCTGTACAACCTGTACCCTGGGGTGT
 20 GAGACTGTGGAGATGCTGCCAGCAAGGCTGGCATCTGGAGGGTGGAGTGCCTGATTGGGGAG
 CACCTGCATGCTGGCATGAGCACCTGTTCTGTGTACAGCAACAAGTGCCAGACCCCCCTGG
 GCATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGCCAGTATGGCCAGTGGG
 CCCCCAAGCTGGCCAGGCTGCACTACTCTGGCAGCATCAATGCCTGGAGCACCAAGGAGCCCT
 TCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAGGGGG
 25 CCAGGCAGAAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGCCTGGATGGCAA
 GAAGTGGCAGACCTACAGGGGCAACAGCACTGGCACCCCTGATGGTGTCTTTGGCAATGTGGAC
 AGCTCTGGCATCAAGCACAACATCTTCAACCCCCCATCATTGCCAGATACATCAGGCTGCACCC
 CACCCACTACAGCATCAGGAGCACCTGAGGATGGAGCTGATGGGCTGTGACCTGAACAGCTG
 CAGCATGCCCTGGGCATGGAGAGCAAGGCCATCTCTGATGCCCAGATCACTGCCAGCAGCTA
 30 CTTACCAACATGTTTGGCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAG
 CAATGCCTGGAGGCCCCAGGTCAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGAC
 CATGAAGGTGACTGGGGTGACCAACCAGGGGGTGAAGAGCCTGCTGACCAGCATGTATGTGAA
 GGAGTTCTGATCAGCAGCAGCCAGGATGGCCACCAGTGGACCCTGTTCTTCCAGAATGGCAAG
 GTGAAGGTGTTCCAGGGCAACCAGGACAGCTTACCCCTGTGGTGAACAGCCTGGACCCCCC
 35 CTGCTGACCAGATACCTGAGGATTCACCCCCAGAGCTGGGTGCACCAGATTGCCCTGAGGATGG
 AAGTGCTGGCTGTGAGGCCCAGGACCTGTACTGA

SEQ ID NO: 31

TACGTCTAACTCGACTCGTGACGAAGAAGGACACGGACGACTCCAAGACGAAGAGACGGT
 GGTCTCTATGATGGACCCCCGACACCTCGACTCGACCCTGATGTACGTCAGACTGGACCCCT
 40 CGACGGACACCTACGGTCCAAGGGGGGGTCTCACGGGTTCTCGAAGGGGAAGTTGTGGAGACA
 CCACATGTTCTTCTGGGACAAACACCTCAAGTGACTGGTGGACAAGTTGTAACGGTTCGGGTCC
 GGGGGGACCTACCCGGACGACCCGGGGTGGTAGGTCCGACTCCACATACTGTGACACCACTAG
 TGGGACTTCTTGTACCGGTGGTGGGACACTCGGACGTACGACACCCCCACTCGATGACCTTCC
 GGAGACTCCCCGACTCATACTACTGGTCTGGTCTCCCTCTTCTCTCTACTGTTCCACAAG
 45 GGACCCCCGTGGTGTGGATACACACCGTCCACGACTTCTTACCGGGGTACCGGAGACTG
 GGGGACACGGACTGGATGTGATGGACTCGGTACACCTGGACCACTTCTTGGACTTGAGACCG
 GACTAACCCCCGGGACGACCACACGTCCCTCCCGTGGGACCGGTTCTTCTTCTGGGTCTGGGAC
 GTGTTCAAGTAGGACGACAAACGACACAACTACTCCCGTTCTCGACCGTGAGACTTTGGTTCTT
 GTCGGACTACGTCTCTGCCCTACGACGGAGACGGTCCCGGACCGGGTTCTACGTGTGACACTTA
 50 CCGATACACTTGTCTCGGACGGACCGGACTAACCGACGGTGTCTTCAGACACATGACCGTAC
 ACTAACCGTACCCGTGGTGGGGACTCCACGTGTCGTAGAAGGACCTCCCGGTGTGGAAGGACC
 AGTCCTTGGTGTCCGTCCGGTCCGACCTCTAGTCGGGGTAGTGGAAGGACTGACGGGTCTGGG
 ACGACTACCTGGACCCGGTCAAGGACGACAAGACGGTGTAGTCGTGGTGGTCTACTACCGTA
 CCTCCGGATACACTTCCACCTGTGACGGGACTCCTCGGGGTGACTCCTACTTCTTGTACTCC
 55 TCCGACTCCTGATACTACTACTGGACTGACTGAGACTCTACCTACACCACTCCAACTACTACTG
 TTGTCCGGGTGCAAGTAGGTCTAGTCCAGACACCGGTTCTTCTGTGGGTCTGGACCCACGTGA
 TGTAACGACGACTCCTCCTCCTGACCCTGATACGGGGGACACGACCGGGGACTACTGTCTC
 GATGTTCTCGGTGATGGACTTGTACCGGGGTCTCCTAACCGTCTTCTGTTCTCCAGTCCA
 AGTACCGGATGTGACTACTTGGAAAGTTCTGGTCCCTCCGGTAGGTCGACTCAGACCGTAGGA
 60 CCCGGGGGACGACATAACCCCTCCACCCCTGTGGGACGACTAGTAGAAGTTCTTGGTCCGGTC

GTCCGGGATGTTGTAGATGGGGGTACCGTAGTGACTACACTCCGGGGACATGTCGTCCTCCGAC
 GGGTTCCCCCACTTCGTGGACTTCCTGAAGGGGTAGGACGGACCCCTCTAGAAGTTCATGTTCA
 CCTGACACTGACACCTCCTACCGGGGTGGTTTCAGACTGGGGTCCACGGACTGGTCTATGATGTC
 5 GTCGAAACACTTGTACCTCTCCCTGGACCGGAGACCGGACTAACCGGGGGACGACTAGACGAT
 GTTCCTCAGACACCTGGTCTCCCCGTTGGTCTAGTACAGACTGTTCTCCTTACACTAGGACAAGA
 GACACAACTACTCTTGTCTCGACCATGGACTGACTCTTGTAGGTCTCCAAGGACGGGTTGGG
 ACGACCCACGTCGACCTCCTGGGACTCAAGGTCCGGTCGTTGTAGTACGTGTCGTAGTTACCG
 ATACACAACTGTCCGACGTCGACAGACACACGGACGTAACCGGATGACCATGTAGGACT
 10 CGTAACCCCGGGTCTGACTGAAGGACAGACACAAGAAGAGACCGATGTGGAAGTTCGTGTTCTA
 CCACATACTCCTGTGGGACTGGGACAAGGGGAAGAGACCCCTCTGACACAAGTACTCGTACCTC
 TTGGGACCGGACACCTAAGACCCGACGGTGTTGAGACTGAAGTCCTTGTCCCCGTAAGTACGCGG
 ACGACTTTCAGAGGTCGACACTGTTCTTGTGACCCCTGATGATACTCCTGTGCGATACTCCTGTAG
 AGACGGATGCGACGACTCGTTCTTGTACGGTAACTCGGGTCCTCGAAGTCGGTCTTACGGTGAT
 15 TACACAGATTGTTGTCGTTGTGGTCGTTACTGTCGTTACACAGAGGGGGTACGACACTTCTCCGTG
 GTCTCCCTCTAGTGGTCCTGGTGGGACGTCAGACTGGTCCTCCTCTAACTGATACTACTGTGGTA
 GAGACACCTCTACTTCTTCTCCTGAACTGTAGATGCTGCTCCTGCTCTTGGTCTCGGGGTCTCT
 CGAAGGTCTTCTTCTGGTCCGTGATGAAGTAACGACGACACCTCTCCGACACCCCTGATACCGTA
 CTCGTCGTCGGGGGTACACGACTCCTTGTCCCGGGTCAGACCGAGACACGGGGTCAAGTTCTT
 CCACCACAAGGTCCTCAAGTGAAGTACCGTCGAAGTGGGTCGGGGACATGTCTCCCCTCGACTTA
 20 CTCGTGGACCCGGACGACCCGGGGATGTAGTCCCGACTCCACCTCCTGTTGTAGTACCACTGGA
 AGTCCTTGGTCCGGTCGTCCGGGATGTGCAAGATGTCGTCGGACTAGTCGATACTCCTCCTGGT
 CTCCGTCCCCCGACTCGGGTCCTTCTTGAACACTTCGGGTACTTTGGTTCTGGATGAAGACCT
 TCCACGTCGTGGTGTACCGGGGGTGGTTCCTACTCAAACGACGTTCCGGACCCGGATGAAGAG
 ACTACACCTGGACCTCTTCTACACGTGAGACCGGACTAACCGGGGGGACGACCACACGGTGTG
 25 GTTGTGGGACTTGGGACGGGTACCGTCCGTCCACTGACACGTCCTCAAACGGGACAAGAAGTG
 GTAGAACTACTTTGGTTCTCGACCATGAAGTGAAGTACTTTGTACCTCTCCTTGACGTCCCGGGGA
 CGTTGTAGGTCTACCTCCTGGGGTGGAAAGTTCCTCTTGATGTCCAAGGTACGGTAGTTACCGAT
 GTAGTACCTGTGGGACGGACCGGACCACTACCGGGTCCTGGTCTCCTAGTCCACCATGGACGA
 CTCGTACCCGTCGTTACTCTTGTAGGTGTCGTAGGTGAAGAGACCGGTACACAAGTGACACTCC
 30 TTCTTCCTCCTCATGTTCTACCGGGACATGTTGGACATGGGACCCACAACTCTGACACCTCTA
 CGACGGGTCGTTCCGACCGTAGACCTCCACCTCACGGACTAACCCCTCGTGGACGTACGACC
 GACTCGTGGGACAAGGACCACATGTCGTTGTTACGGTCTGGGGGGACCCGTACCGGAGACC
 GGTGTAGTCCCTGAAGGTCTAGTGACGGAGACCGGTCATACCGGTACCCGGGGGGTTCGACCG
 GTCCGACGTGATGAGACCGTCGTAGTTACGGACCTCGTGTTCTCGGGAAGTCGACCTAGTTC
 35 CACCTGGACGACCGGGGGTACTAGTAGGTACCGTAGTTCTGGGTCCCCCGGTCCGTCTTCAAGT
 CGTCGGACATGTAGTCGGTCAAGTAGTAGTACATGTCGGACCTACCGTTCTTCACCGTCTGGAT
 GTCCCCGTTGTCGTGACCGTGGGACTACCACAAGAAACCGTTACACCTGTCGAGACCGTAGTTC
 GTGTTGTAGAAGTTGGGGGGGTAGTAACGGTCTATGTAGTCCGACGTGGGGTGGGTGATGTCGT
 AGTCCTCGTGGGACTCCTACCTCGACTACCCGACACTGGACTTGTGACGTCGTACGGGGACCC
 40 GTACCTCTCGTTCGGGTAGAGACTACGGGTCTAGTGACGGTCTGTCGATGAAGTGGTTGTACAAA
 CGGTGGACCTCGGGGTCGTTCCGGTCCGACGTGGACGTCCCGTCCTCGTTACGGACCTCCGGG
 GTCCAGTTGTTGGGGTCTCACCACGTCCACCTGAAGGTCTTCTGGTACTTCCACTGACCCC
 ACTGGTGGGTCCCCCACTTCTCGGACGACTGGTCGTACATACTTCTCAAGGACTAGTCGTC
 GTCGGTCTACCGGTGGTCACTGGGACAAGAAGGTCTTACCGTTCACCTCCACAAGGTCCCG
 45 TTGGTCTGTGCAAGTGGGGACACCACTTGTGCGACCTGGGGGGGACGACTGGTCTATGGAC
 TCCTAAGTGGGGGTCTCGACCCACGTGGTCTAACGGGACTCCTACCTCCACGACCCGACACTCC
 GGGTCCTGGACATGACT

SEQ ID NO: 32

50 MQIELSTCFFLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPFPRVPKSFNFNTSVVY
 KKTFLF
 VEFTDHLFNIAPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDD
 QTSQR
 EKEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREGSLA
 KEKTQT
 55 LHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCHRKSVYW
 HVIGMG
 TTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLMLDLGQFLLFCHISSHQHDGMEAYVKVDS
 CPE
 EPQLRMKNNEEAEDYDDDLTDSEMDVVRFDNDDNSPSFIQIRSVAKKHPKTWVHYIAAEEEDWDY
 60 APLVLA

PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILGPLLYGEVGD TLLIIFKN
 QASR
 PYNIYPHGITDVRPLYSRRLPKGVKHLKDFILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFV
 NMER
 5 DLASGLIGPLLYCYKESVDQQRGNQIMSDKRNILFSVFDENRSWYL TENIQRFLPNPAGVQLEDPE
 FQAS
 NIMHSINGYVFDSLQLSVCLHEVAYWYILSIGAQTDFLSVFFSGYT FKHKVMYEDTLTLFPFSGETV
 FMS
 MENPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNAIEPRSF SQNA
 10 TNVSN
 NSNTSNDENVSPV LKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQK
 KTRHY
 FIAAVERLWDYGMSSSPHVLNRNRAQSGSV PQFKVVFQEFTDGSFTQPLYRGELNEHLGLLGPY
 IRAEVE
 15 DNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEP RKNFVKPNETKTYFWKVQH HMAPTKDEFDC
 KAWAYF
 SDVDLEKDVHSLIGLPLLVCHTNTLNPAHGRQVTVQEFALFFTIFDETKSWYFTENMERNCRAPC
 NIQME
 DPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMG SNENIHSIHFSGHVFTVRKKEEYKM
 20 ALYNL
 YPGVFETVEMLP SKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQY
 GQWAP
 KLARLHYSGSINAWSTKEPFSWIKVDLLAPMIIHG IKTQGARQKFSSLYISQFIIMYSLDGKKWQTY
 RGN
 25 STGTLMVFFGNVDSSGIKH NIFNPPIARYIRLHP THYSIRSTLRMELMGCDLNSCSMPLGMESKAI
 SDA
 QITASSYFTNM FATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTVTTQGVKSL
 LTSMYVK
 EFLISSSQDGHQWTLFFQNGKV KVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEV
 30 LGCEA
 QDLY
 SEQ ID NO: 33
 GGCGCCTCTAGAGTTATAACCGGTAATCGGTATAATAAGTAACCAATATATCGTATTTAGTTAT
 AACCGA
 35 TAACCGGTAACGTATGCAACATAGATATAGTATTATACATGTAAATATAACCGAGTACAGGTTA
 TACTGG
 CGGTACAACCGTA ACTAATAACTGATCAATAATTATCATTAGTTAATGCCCCAGTAATCAAGTA
 TCGGGT
 ATATACCTCAAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTTGCT
 40 GGGGGCGG
 GTAAGTGCAGTTATTACTGCATACAAGGGTATCATTGCGGTTATCCCTGAAAGGTAAGTGCAG
 TTACCCA
 CCTCATAAATGCCATTTGACGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCAGGCG
 GGGGATAA
 45 CTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACTGGAATGCCCTGAA
 AGGATGAA
 CCGTCATGTAGATGCATAATCAGTAGCGATAATGGTACCACTACGCCAAAACCGTCATGTGG
 TTACCCGC
 ACCTATCGCCAAACTGAGTGCCCCTAAAGGTT CAGAGGTGGGGTAACTGCAGTTACCCTCAA
 50 ACAAAC
 GTGGTTTTAGTTGCCCTGAAAGGTTTTACAGCATTATTGGGGCGGGGCAACTGCGTTTACCC
 GCCATCCG
 CACATGCCACCCTCCAGATATATTCGTCTCGAGCAAATCACTTGGCAGTCTAGTGATCTTCGA
 AATAACG
 55 CCATCAAATAGTGTCAATTTAACGATTGCGTCAGTCACGAAGACTGTGTTGTCAGAGCTTGAA
 TTCGACG
 TCTTCAACCAGCACTCCGTGACCCGTCCGATCG

60 ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<110> Imperial Innovations Limited

Imperial College of Science, Technology & Medicine
 The Chancellor, Masters and Scholars of The University of
 Oxford
 The University Court of The University of Edinburgh
 ДНКVEC Corporation

5

<120> Лентивірусні Вектори

<130> P42911WO

10

<150> GB1409089.8

<151> 2014-05-21

<160> 33

15

<170> PatentIn version 3.5

<210> 1

<211> 10535

20

<212> ДНК

<213> штучна послідовність

<220>

<223> pGM326 плазміда

25

<400> 1

ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat 60

tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc 120

30

atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat 180

tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 240

35

tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 300

tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 360

aactgcccac ttggcagtag atcaagtgtg tcatatgccg agtccgcccc ctattgacgt 420

40

caatgacggt aaatggcccg cctggcatta tgcccagtag atgaccttac gggactttcc 480

tacttggcag tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca 540

45

gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 600

tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 660

caactgcatg cggccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggagggtc 720

50

tatataagca gagctcgctg gcttgtaact cagtctctta ctaggagacc agcttgagcc 780

tgggtgttcg ctggttagcc taacctggtt ggccaccagg ggtaaggact ccttggttta 840

55

gaaagctaata aaacttgcct gcattagagc ttatctgagt caagtgtcct cattgacgcc 900

tacttctctt gaacgggaat cttccttact gggttctctc tctgaccagc gcgagagaaa 960

ctccagcagt ggcgcccga cagggacttg agtgagagtg taggcacgta cagctgagaa 1020

60

ggcgtcggac gcgaaggaag cgcgggggtgc gacgcgacca agaaggagac ttggtgagta 1080

ggcttctcga gtgccgggaa aaagctcgag cctagttaga ggactaggag aggccgtagc 1140

65

cgtaactact cttgggcaag tagggcaggc ggtgggtacg caatgggggc ggctacctca 1200

gcactaaata ggagacaatt agaccaatctt gagaaaatac gacttcgccc gaacggaaaag 1260

aaaaagtacc aaattaaaca tttaatatgg gcaggcaagg agatggagcg cttcggcctc 1320

70

	catgagaggt	tgttgagac	agaggagggg	tgtaaaagaa	tcatagaagt	cctctacccc	1380
	ctagaaccaa	caggatcgga	gggcttaaaa	agtctgttca	atcttggtg	cgtgctatat	1440
5	tgcttgacac	aggaacagaa	agtgaagac	acagaggaag	cagtagcaac	agtaagacaa	1500
	cactgccatc	tagtggaaaa	agaaaaaagt	gcaacagaga	catctagtgg	acaaaagaaa	1560
10	aatgacaagg	gaatagcagc	gccacctggt	ggcagtcaga	attttccagc	gcaacaacaa	1620
	ggaaatgcct	gggtacatgt	acccttgtca	ccgcgcacct	taaatgcgtg	ggtaaaagca	1680
	gtagaggaga	aaaaatttgg	agcagaaata	gtaccatttt	ttttgtttca	agccctatcg	1740
15	aattcccgtt	tgtgctaggg	ttcttaggct	tcttgggggc	tgctggaact	gcaatgggag	1800
	cagcggcgac	agccctgacg	gtccagtctc	agcatttgct	tgctgggata	ctgcagcagc	1860
20	agaagaatct	gctggcggct	gtggaggctc	aacagcagat	gttgaagctg	accatttggg	1920
	gtgttaaaaa	cctcaatgcc	cgcgtcacag	cccttgagaa	gtacctagag	gatcaggcac	1980
	gactaaactc	ctgggggtgc	gcatggaaac	aagtatgtca	taccacagtg	gagtggccct	2040
25	ggacaaatcg	gactccggat	tggcaaaata	tgacttggtt	ggagtgggaa	agacaaatag	2100
	ctgatttggg	aagcaacatt	acgagacaat	tagtgaaggc	tagagaacaa	gaggaaaaga	2160
30	atctagatgc	ctatcagaag	ttaactagtt	ggtcagattt	ctggtcttgg	ttcgatttct	2220
	caaaatggct	taacatttta	aaaatgggat	ttttagtaat	agtaggaata	atagggttaa	2280
	gattacttta	cacagtatat	ggatgtatag	tgagggttag	gcagggatat	gttcctctat	2340
35	ctccacagat	ccatatccgc	ggcaatttta	aaagaaaggg	aggaataggg	ggacagactt	2400
	cagcagagag	actaattaat	ataataacaa	cacaattaga	aataacaacat	ttacaaacca	2460
40	aaattcaaaa	aattttaaat	tttagagccg	cggagatctg	ttacataact	tatggtaaata	2520
	ggcctgcctg	gctgactgcc	caatgacccc	tgcccaatga	tgtcaataat	gatgtatggt	2580
	cccatgtaat	gccaataggg	actttccatt	gatgtcaatg	ggtggagtat	ttatggtaac	2640
45	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	atgcccccta	ttgatgtcaa	2700
	tgatggtaaa	tggcctgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	2760
50	ttggcagtac	atctatgtat	tagtcattgc	tattaccatg	ggaattcact	agtggagaag	2820
	agcatgcttg	agggctgagt	gcccctcagt	gggcagagag	cacatggccc	acagtccctg	2880
	agaagtggg	gggaggggtg	ggcaattgaa	ctggtgccta	gagaaggtgg	ggcttgggta	2940
55	aactgggaaa	gtgatgtggt	gtactggctc	cacctttttc	cccaggggtg	gggagaacca	3000
	tatataagtg	cagtagtctc	tgtgaacatt	caagcttctg	ccttctccct	cctgtgagtt	3060
60	tgctagccac	catgcagaga	agccctctgg	agaaggcctc	tgtggtgagc	aagctgttct	3120
	tcagctggac	caggcccatc	ctgaggaagg	gctacaggca	gagactggag	ctgtctgaca	3180
	tctaccagat	cccctctgtg	gactctgctg	acaacctgtc	tgagaagctg	gagagggagt	3240
65	gggatagaga	gctggccagc	aagaagaacc	ccaagctgat	caatgccctg	aggagatgct	3300
	tcttctggag	attcatgttc	tatggcatct	tcctgtacct	gggggaagtg	accaaggctg	3360
70	tgcagcctct	gctgctgggc	agaatcattg	ccagctatga	ccctgacaac	aaggaggaga	3420

	ggagcattgc	catctacctg	ggcattggcc	tgtgcctgct	gttcattgtg	aggaccctgc	3480
	tgctgcaccc	tgccatcttt	ggcctgcacc	acattggcat	gcagatgagg	attgccatgt	3540
5	tcagcctgat	ctacaagaaa	accctgaagc	tgtccagcag	agtgtctggac	aagatcagca	3600
	ttggccagct	ggtgagcctg	ctgagcaaca	acctgaacaa	gtttgatgag	ggcctggccc	3660
10	tggcccactt	tgtgtggatt	gcccctctgc	aggtggccct	gctgatgggc	ctgatttggg	3720
	agctgctgca	ggcctctgcc	ttttgtggcc	tgggcttcct	gattgtgctg	gcccctgtttc	3780
	aggctggcct	gggcaggatg	atgatgaagt	acagggacca	gagggcaggc	aagatcagtg	3840
15	agaggctggg	gatcacctct	gagatgattg	agaacatcca	gtctgtgaag	gcctactggt	3900
	gggaggaagc	tatggagaag	atgattgaaa	acctgaggca	gacagagctg	aagctgacca	3960
20	ggaaggctgc	ctatgtgaga	tacttcaaca	gctctgcctt	cttcttctct	ggcttctttg	4020
	tgggtgttcct	gtctgtgctg	ccctatgccc	tgatcaaggg	gatcatcctg	agaaagattt	4080
	tcaccaccat	cagcttctgc	attgtgtctg	ggatggctgt	gaccagacag	ttcccctggg	4140
25	ctgtgcagac	ctggtatgac	agcctggggg	ccatcaacaa	gatccaggac	ttcctgcaga	4200
	agcaggagta	caagaccctg	gagtacaacc	tgaccaccac	agaagtgggtg	atggagaatg	4260
	tgacagcctt	ctgggaggag	ggctttgggg	agctgtttga	gaaggccaag	cagaacaaca	4320
30	acaacagaaa	gaccagcaat	ggggatgact	ccctgttctt	ctccaacttc	tccctgctgg	4380
	gcacacctgt	gctgaaggac	atcaacttca	agattgagag	ggggcagctg	ctggctgtgg	4440
35	ctggatctac	aggggctggc	aagaccagcc	tgctgatgat	gatcatgggg	gagctggagc	4500
	cttctgaggg	caagatcaag	cactctggca	ggatcagctt	ttgcagccag	ttcagctgga	4560
	tcatgcctgg	caccatcaag	gagaacatca	tctttggagt	gagctatgat	gagtacagat	4620
40	acaggagtgt	gatcaaggcc	tgccagctgg	aggaggacat	cagcaagttt	gctgagaagg	4680
	acaacattgt	gctggggggag	ggaggcatta	cactgtcttg	gggccagaga	gccagaatca	4740
45	gcctggccag	ggctgtgtac	aaggatgctg	acctgtacct	gctggactcc	ccctttggct	4800
	acctggatgt	gctgacagag	aaggagattt	ttgagagctg	tgtgtgcaag	ctgatggcca	4860
	acaagaccag	aatcctgggtg	accagcaaga	tggagcacct	gaagaaggct	gacaagatcc	4920
50	tgatcctgca	tgagggcagc	agctacttct	atgggacctt	ctctgagctg	cagaacctgc	4980
	agcctgactt	cagctctaag	ctgatgggct	gtgacagctt	tgaccagtcc	tctgctgaga	5040
55	ggaggaacag	catcctgaca	gagaccctgc	acagattcag	cctggagggg	gatgcccctg	5100
	tgagctggac	agagaccaag	aagcagagct	tcaagcagac	aggggagttt	ggggagaaga	5160
	ggaagaactc	catcctgaac	cccatcaaca	gcacaggaa	gttcagcatt	gtgcagaaaa	5220
60	ccccctgca	gatgaatggc	attgaggaag	attctgatga	gcccctggag	aggagactga	5280
	gcctggtgcc	tgattctgag	caggagagag	ccatcctgcc	taggatctct	gtgatcagca	5340
65	caggccctac	actgcaggcc	agaaggaggc	agtctgtgct	gaacctgatg	accactctg	5400
	tgaaccaggg	ccagaacatc	cacaggaaaa	ccacagcctc	caccaggaaa	gtgagcctgg	5460
70	cccctcaggc	caatctgaca	gagctggaca	tctacagcag	gaggctgtct	caggagacag	5520

	gcctggagat	ttctgaggag	atcaatgagg	aggacctgaa	agagtgcctt	tttgatgaca	5580
	tggagagcat	ccctgctgtg	accacctgga	acacctacct	gagatacatc	acagtgcaca	5640
5	agagcctgat	ctttgtgctg	atctggtgcc	tggatgatctt	cctggctgaa	gtggctgcct	5700
	ctctggtggt	gctgtggctg	ctgggaaaca	ccccactgca	ggacaagggc	aacagcaccc	5760
	acagcaggaa	caacagctat	gctgtgatca	tcacctccac	ctccagctac	tatgtgttct	5820
10	acatctatgt	gggagtggct	gataccctgc	tggctatggg	cttctttaga	ggcctgcccc	5880
	tggatgcacac	actgatcaca	gtgagcaaga	tcctccacca	caagatgctg	cactctgtgc	5940
15	tgcaggctcc	tatgagcacc	ctgaataccc	tgaaggctgg	gggcatcctg	aacagattct	6000
	ccaaggatat	tgccatcctg	gatgacctgc	tgcctctcac	catctttgac	ttcatccagc	6060
	tgtgtgctgat	tgtgattggg	gccattgctg	tggatggcagt	gctgcagccc	tacatctttg	6120
20	tggccacagt	gcctgtgatt	gtggccttca	tcagtctgag	ggcctacttt	ctgcagacct	6180
	cccagcagct	gaagcagctg	gagtctgagg	gcagaagccc	catcttcacc	cacctggtga	6240
25	caagcctgaa	gggcctgtgg	accctgagag	cctttggcag	gcagccctac	tttgagaccc	6300
	tgttccacaa	ggccctgaac	ctgcacacag	ccaactgggt	cctctacctg	tccaccctga	6360
	gatgggttcca	gatgagaatt	gagatgatct	ttgtcatctt	cttcattgct	gtgaccttca	6420
30	tcagcattct	gaccacagga	gagggagagg	gcagagtggg	cattatcctg	accctggcca	6480
	tgaacatcat	gagcacactg	cagtgggcag	tgaacagcag	cattgatgtg	gacagcctga	6540
35	tgaggagtgt	gagcagagtg	ttcaagttca	ttgatatgcc	cacagagggc	aagcctacca	6600
	agagcaccaa	gccctacaag	aatggccagc	tgagcaaagt	gatgatcatt	gagaacagcc	6660
	atgtgaagaa	ggatgatatc	tggcccagtg	gaggccagat	gacagtgaag	gacctgacag	6720
40	ccaagtacac	agaggggggc	aatgctatcc	tggagaacat	ctccttcagc	atctccccctg	6780
	gccagagagt	gggactgctg	ggaagaacag	gctctggcaa	gtctaccctg	ctgtctgcct	6840
45	tcctgaggct	gctgaacaca	gagggagaga	tccagattga	tggagtgtcc	tgggacagca	6900
	tcacactgca	gcagtggagg	aaggcctttg	gtgtgatccc	ccagaaagtg	ttcatcttca	6960
	gtggcacctt	caggaagaac	ctggaccctt	atgagcagtg	gtctgaccag	gagatttgga	7020
50	aagtggctga	tgaagtgggc	ctgagaagtg	tgattgagca	gttccctggc	aagctggact	7080
	ttgtcctggt	ggatgggggc	tgtgtgctga	gccatggcca	caagcagctg	atgtgcctgg	7140
55	ccagatcagt	gctgagcaag	gccaaagatcc	tgctgctgga	tgagccttct	gcccacctgg	7200
	atcctgtgac	ctaccagatc	atcaggagga	ccctcaagca	ggcctttgct	gactgcacag	7260
	tcacctctgtg	tgagcacagg	attgaggcca	tgctggagtg	ccagcagttc	ctggtgattg	7320
60	aggagaacaa	agtgaggcag	tatgacagca	tccagaagct	gctgaatgag	aggagcctgt	7380
	tcaggcaggc	catcagcccc	tctgatagag	tgaagctgtt	ccccacagg	aacagctcca	7440
65	agtgaagag	caagccccag	attgctgccc	tgaaggagga	gacagaggag	gaagtgcagg	7500
	acaccaggct	gtgagggccc	aatcaacctc	tggattacaa	aatttgatga	agattgactg	7560
70	gtattcttaa	ctatgttgct	ccttttacgc	tatgtggata	cgctgcttta	atgcctttgt	7620

	atcatgctat	tgcttcccgt	atggctttca	ttttctcctc	cttgataaaa	tcctggttgc	7680
	tgtctcttta	tgaggagttg	tggtccgttg	tcaggcaacg	tggtgtggtg	tgactgtgt	7740
5	ttgctgacgc	aacccccact	ggttggggca	ttgccaccac	ctgtcagctc	ctttccggga	7800
	ctttcgcttt	ccccctccct	attgccacgg	cggaactcat	cgccgcctgc	cttgcccgtc	7860
	gctggacagg	ggctcggtg	ttgggactg	acaattccgt	ggtgttgtcg	gggaaatcat	7920
10	cgctctttcc	ttggctgtc	gcctgtgttg	ccacctggat	tctgcgcggg	acgtccttct	7980
	gctacgtccc	ttcgccctc	aatccagcgg	accttccttc	ccgcggcctg	ctgccggctc	8040
15	tgccgcctct	tccgcgtctt	cgcttccgcc	ctcagacgag	tcggatctcc	ctttgggccg	8100
	cctccccgca	agcttcgcac	tttttaaaag	aaaagggagg	actggatggg	atttattact	8160
	ccgataggac	gctggcttgt	aactcagtct	cttactagga	gaccagcttg	agcctgggtg	8220
20	ttcgctgggt	agcctaacct	ggttggccac	caggggtaag	gactccttgg	cttagaaagc	8280
	taataaactt	gcctgcatta	gagctcttac	gcgtcccggg	ctcgagatcc	gcatctcaat	8340
25	tagtcagcaa	ccatagtccc	gcccctaact	ccgcccattc	cgcccctaac	tccgcccagt	8400
	tccgcccatt	ctccgcccc	tggtgacta	atTTTTTTT	tttatgcaga	ggccgaggcc	8460
	gcctcggcct	ctgagctatt	ccagaagtag	tgaggaggct	TTTTTggagg	cctaggcttt	8520
30	tgcaaaaagc	taacttggtt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	8580
	caaatttcac	aaataaagca	TTTTTTTcac	tgatttctag	ttgtgggttg	tccaaactca	8640
35	tcaatgtatc	ttatcatgtc	tgtccgcttc	ctcgctcact	gactcgtctg	gctcggtcgt	8700
	tcggctgcgg	cgagcgggat	cagctcactc	aaaggcggtg	atacggttat	ccacagaatc	8760
	aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	caaaaggcca	ggaaccgtaa	8820
40	aaaggccgcg	ttgctggcgt	ttttccatag	gctccgcccc	cctgacgagc	atcacaaaaa	8880
	tcgacgtca	agtcagaggt	ggcgaaaccc	gacaggacta	taaagatacc	aggcgtttcc	8940
45	ccctggaagc	tccctcgtgc	gctctcctgt	tccgaccctg	ccgcttaccg	gatacctgtc	9000
	cgcttttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	tcacgctgta	ggtatctcag	9060
	ttcgggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	gaaccccccg	ttcagcccga	9120
50	ccgctgcgcc	ttatccggtg	actatcgtct	tgagtccaac	ccggtgaagc	acgacttatc	9180
	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	aggatgtgag	gcggtgctac	9240
55	agagtctctg	aagtgggtgg	ctaactacgg	ctacactaga	agaacagtat	ttggtatctg	9300
	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggt	agctcttgat	ccggcaaaca	9360
	aaccaccgct	ggtagcgggt	gtttttttgt	ttgcaagcag	cagattacgc	gcagaaaaaa	9420
60	aggatctcaa	gaagatcctt	tgatcttttc	tacgggggtct	gacgctcagt	ggaacgaaaa	9480
	ctcacgttaa	gggatttttg	tcatgagatt	atcaaaaagg	atcttcacct	agatcctttt	9540
65	aaattaaaaa	tgaagtttta	aatcaatcta	aagtatatat	gagtaaaact	ggtctgacag	9600
	ttagaaaaac	tcatcgagca	tcaaatgaaa	ctgcaattta	ttcatatcag	gattatcaat	9660
70	accatatttt	tgaaaaagcc	gtttctgtaa	tgaaggagaa	aactcaccga	ggcagttcca	9720

	taggatggca agatcctggt atcgggtctgc gattccgact cgtccaacat caatacaacc	9780
	tattaatttc ccctcgtcaa aaataagggt atcaagtgag aaatcaccat gagtgcgac	9840
5	tgaatccggt gagaatggca acagcttatg cttttctttc cagacttggt caacaggcca	9900
	gccattacgc tcgtcatcaa aatcactcgc atcaaccaa cgtttattca ttcgtgattg	9960
10	cgcctgagcg agacgaaata cgcgatcgt gttaaaagga caattacaaa caggaatcga	10020
	atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata	10080
	ttcttctaata acctggaatg ctgtttttcc ggggatcgc gtggtgagta accatgcatc	10140
15	atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt	10200
	tagtctgacc atctcatctg taacatcatt ggcaacgcta cttttgccat gtttcagaaa	10260
20	caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgc	10320
	attatcgca gccatttat acccatataa atcagcatcc atgttggaat ttaatcgcg	10380
	cctagagcaa gacgtttccc gttgaatatg gtcataaca ccccttgat tactgtttat	10440
25	gtaagcagac agttttattg ttcatgatga tatattttta tcttggtgcaa tgtaacatca	10500
	gagattttga gacacaaca ttggtcgac gatcc	10535
30	<210> 2	
	<211> 9885	
	<212> ДНК	
	<213> штучна послідовність	
35	<220>	
	<223> pGM297 плазміда	
	<400> 2	
40	gctcgagact agtgacttgg tgagtaggct tcgagcctag ttagaggact aggagaggcc	60
	gtagccgtaa ctactctggg caagtagggc aggcggtggg tacgcaatgg gggcggtac	120
	ctcagcacta aataggagac aattagacca atttgagaaa atacgacttc gcccgaacgg	180
45	aaagaaaaag taccaaatta aacatttaat atgggcaggc aaggagatgg agcgcttcgg	240
	cctccatgag aggttggtgg agacagagga ggggtgtaaa agaatcatag aagtcctcta	300
50	ccccctagaa ccaacaggat cggagggctt aaaaagtctg ttcaatcttg tgtgcgtact	360
	atattgcttg cacaaggaac agaaagtga agacacagag gaagcagtag caacagtaag	420
	acaacactgc catctagtgg aaaaagaaaa aagtgaaca gagacatcta gtggacaaaa	480
55	gaaaaatgac aagggaatag cagcgccacc tgggtggcagt cagaattttc cagcgcaaca	540
	acaaggaaat gcctgggtac atgtaccctt gtcaccgcgc accttaaag cgtgggtaaa	600
60	agcagtagag gagaaaaaat ttggagcaga aatagtaccc atgtttcaag ccctatcaga	660
	aggctgcaca ccctatgaca ttaatcagat gcttaatgtg ctaggagatc atcaaggggc	720
	attacaaata gtgaaagaga tcattaatga agaagcagcc cagtgggatg taacacaccc	780
65	actaccgcga ggaccctac cagcaggaca gtcagggac cctcgcggt cagatatagc	840
	agggaccacc agctcagtag aagaacagtt agaattggatc tatactgcta acccccgggt	900
70	agatgtaggt gccatctacc ggagatggat tattctagga cttcaaaagt gtgtcaaaat	960

	gtacaacca	gtatcagtc	tagacattag	gcagggacct	aaagagccct	tcaaggatta	1020
	tgtggacaga	ttttacaagg	caattagagc	agaacaagcc	tcaggggaag	tgaaacaatg	1080
5	gatgacagaa	tcattactca	ttcaaaatgc	taatccagat	tgtaagggtca	tcctgaaggg	1140
	cctaggaatg	caccccaccc	ttgaagaaat	gttaacggct	tgtcaggggg	taggaggccc	1200
	aagctacaaa	gcaaaagtaa	tggcagaaat	gatgcagacc	atgcaaaatc	aaaacatggt	1260
10	gcagcagggg	ggtccaaaaa	gacaaagacc	cccactaaga	tgttataatt	gtggaaaatt	1320
	tggccatatg	caaagacaat	gtccggaacc	aaggaaaaca	aaatgtctaa	agtgtggaaa	1380
15	attgggacac	ctagcaaaag	actgcagggg	acaggtgaat	tttttagggg	atggacgggtg	1440
	gatgggggca	aaaccgagaa	atgttccgcg	cgctactctt	ggagcgggaa	cgagtgcgcc	1500
	tcctccaccg	agcggcacca	ccccatacga	cccagcaaag	aagctcctgc	agcaatatgc	1560
20	agagaaaggg	aaacaactga	gggagcaaaa	gaggaatcca	ccggcaatga	atccggattg	1620
	gaccgagggg	tattctttga	actccctctt	tggagaagac	caataaagac	agtgtatata	1680
25	gaaggggtcc	ccattaaggc	actgctagac	acaggggcag	atgacaccat	aattaaagaa	1740
	aatgatttac	aattatcagg	tccatggaga	cccaaaatta	tagggggcat	aggaggaggc	1800
	cttaatgtaa	aagaatataa	cgacagggaa	gtaaaaatag	aagataaaat	tttgagagga	1860
30	acaatattgt	taggagcaac	tcccattaat	ataataggta	gaaatttgct	ggccccggca	1920
	ggtgcccggg	tagtaatggg	acaattatca	gaaaaaattc	ctgtcacacc	tgtcaaattg	1980
35	aaggaagggg	ctcggggacc	ctgtgtaaga	caatggcctc	tctctaaaga	gaagattgaa	2040
	gctttacagg	aaatatgttc	ccaattagag	caggaaggaa	aaatcagtag	agtaggagga	2100
	gaaaatgcat	acaatacccc	aatattttgc	ataaagaaga	aggacaaatc	ccagtggagg	2160
40	atgctagtag	acttttagaga	gttaaataag	gcaacccaag	atgtctttga	agtgcatta	2220
	gggatacccc	accagcagg	attaagaaag	atgagacaga	taacagtttt	agatgtagga	2280
45	gacgcctatt	attccatacc	attggatcca	aatttttagga	aatatactgc	ttttactatt	2340
	cccacagtga	ataatcaggg	accggggatt	aggtatcaat	tcaactgtct	cccgaaggg	2400
	tggaaaggat	ctcctacaat	cttccaaaat	acagcagcat	ccattttgga	ggagataaaa	2460
50	agaaacttgc	cagcactaac	cattgtacaa	tacatggatg	atztatgggt	aggttctcaa	2520
	gaaaatgaac	acacccatga	caaattagta	gaacagttaa	gaacaaaatt	acaagcctgg	2580
55	ggcttagaaa	cccagaaaaa	gaaggtgcaa	aaagaaccac	cttatgagtg	gatgggatac	2640
	aaactttggc	ctcacaaatg	ggaactaagc	agaatacaac	tggaggaaaa	agatgaatgg	2700
	actgtcaatg	acatccagaa	gttagttggg	aaactaaatt	gggcagcaca	attgtatcca	2760
60	ggtcttagga	ccaagaatat	atgcaagtta	attagaggaa	agaaaaatct	gttagagcta	2820
	gtgacttggg	cacctgaggc	agaagctgaa	tatgcagaaa	atgcagagat	tcttaaaaca	2880
65	gaacaggaag	gaacctatta	caaaccagga	atacctatta	gggcagcagt	acagaaattg	2940
	gaaggaggac	agtggagtta	ccaattcaaa	caagaaggac	aagtcttgaa	agtaggaaaa	3000
70	tacaccaagc	aaaagaacac	ccatacaaat	gaacttcgca	cattagctgg	tttagtgcag	3060

	aagatttgca	aagaagctct	agttatttgg	gggatattac	cagttctaga	actccccgata	3120
	gaaagagagg	tatgggaaca	atggtgggcg	gattactggc	aggtaagctg	gattccccgaa	3180
5	tgggattttg	tcagcacccc	acctttgctc	aaactatggt	acacattaac	aaaagaaccc	3240
	atacccaagg	aggacgttta	ctatgtagat	ggagcatgca	acagaaattc	aaaagaagga	3300
10	aaagcaggat	acatctcaca	atacgaaaa	cagagagtag	aaacattaga	aaacactacc	3360
	aatcagcaag	cagaattaac	agctataaaa	atggccttgg	aagacagtgg	gcctaattgtg	3420
	aacatagtaa	cagactctca	atatgcaatg	ggaattttga	cagcacaacc	cacacaaagt	3480
15	gattcaccat	tagtagagca	aattatagcc	ttaatgatac	aaaagcaaca	aatatatttg	3540
	cagtgggtac	cagcacataa	aggaatagga	ggaaatgagg	agatagataa	attagttagt	3600
20	aaaggcatta	gaagagtttt	attcttagaa	aaaatagaag	aagctcaaga	agagcatgaa	3660
	agatatcata	ataattggaa	aaacctagca	gatacatatg	ggcttccaca	aatagtagca	3720
	aaagagatag	tggccatgtg	tccaaaatgt	cagataaagg	gagaaccagt	gcatggacaa	3780
25	gtggatgcct	cacctggaac	atggcagatg	gattgtactc	atctagaagg	aaaagtagtc	3840
	atagttgcgg	tccatgtagc	cagtggattc	atagaagcag	aagtcatacc	tagggaaaca	3900
30	ggaaaagaaa	cggcaaagtt	tctattaaaa	atactgagta	gatggcctat	aacacagtta	3960
	cacacagaca	atgggcctaa	ctttacctcc	caagaagtgg	cagcaatatg	ttggtgggga	4020
	aaaattgaac	atacaacagg	tataccatat	aacccccaat	ctcaaggatc	aatagaaagc	4080
35	atgaacaaac	aattaaaaga	gataattggg	aaaataagag	atgattgcca	atatacagag	4140
	acagcagtac	tgatggcttg	ccatattcac	aattttaaaa	gaaagggagg	aataggggga	4200
40	cagacttcag	cagagagact	aattaatata	ataacaacac	aattagaaat	acaacattta	4260
	caaaccaaaa	ttcaaaaaat	tttaaatttt	agagtctact	acagagaagg	gagagaccct	4320
	gtgtggaaag	gaccagcaca	attaatctgg	aaaggggaag	gagcagtggg	cctcaaggac	4380
45	ggaagtgacc	taaaggttgt	accaagaagg	aaagctaaaa	ttattaagga	ttatgaaccc	4440
	aaacaaagag	tgggtaatga	gggtgacgtg	gaagggtacca	ggggatctga	taactaaatg	4500
50	gcagggaata	gtcagatatt	ggatgagaca	aagaaatttg	aaatggaact	attatatgca	4560
	tcagctggcg	gccgcgaatt	cactagtgat	tcccgtttgt	gctaggggtc	ttaggcttct	4620
	tgggggctgc	tggaactgca	atgggagcag	cggcgacagc	cctgacgggtc	cagtctcagc	4680
55	atttgcttgc	tgggatactg	cagcagcaga	agaatctgct	ggcggctgtg	gaggctcaac	4740
	agcagatgtt	gaagctgacc	atttgggggtg	ttaaaaacct	caatgcccgc	gtcacagccc	4800
60	ttgagaagta	cctagaggat	caggcacgac	taaactcctg	ggggtgcgca	tggaaacaag	4860
	tatgtcatac	cacagtggag	tggccctgga	caaatcggac	tccggattgg	caaaatatga	4920
	cttggttggg	gtgggaaaga	caaatagctg	atttggaag	caacattacg	agacaattag	4980
65	tgaaggctag	agaacaagag	gaaaagaatc	tagatgccta	tcagaagtta	actagttggt	5040
	cagatttctg	gtcttggttc	gatttctcaa	aatggcttaa	cattttaaaa	atgggatttt	5100
70	tagtaatagt	aggaataata	gggttaagat	tactttacac	agtatatgga	tgtatagtga	5160

	gggttaggca	gggatatggt	cctctatctc	cacagatcca	tatccaatcg	aattcccgcg	5220
	gccgcaattc	actcctcagg	tgcaggctgc	ctatcagaag	gtggtggctg	gtgtggccaa	5280
5	tgccctggct	cacaaatacc	actgagatct	ttttccctct	gccaaaaatt	atggggacat	5340
	catgaagccc	cttgagcatc	tgacttctgg	ctaataaagg	aaattttatt	tcattgcaat	5400
	agtgtgttgg	aatTTTTTgt	gtctctcact	cggaaggaca	tatgggaggg	caaatcattt	5460
10	aaaacatcag	aatgagtatt	tggttttagag	tttggcaaca	tatgcccata	tgctggctgc	5520
	catgaacaaa	ggttggctat	aaagaggtca	tcagtatatg	aaacagcccc	ctgctgtcca	5580
15	ttccttattc	catagaaaag	ccttgacttg	aggttagatt	ttttttatat	tttgTTTTgt	5640
	gttattTTTT	tctttaacat	ccctaaaatt	ttcctttacat	gtttttactag	ccagattttt	5700
	cctcctctcc	tgactactcc	cagtcatagc	tgtccctctt	ctcttatgga	gatccctcga	5760
20	cctgcagccc	aagcttggcg	taatcatggt	catagctgtt	tcctgtgtga	aattgttatc	5820
	cgctcacaat	tccacacaac	atacgagccg	gaagcataaa	gtgtaaagcc	tggggtgcct	5880
25	aatgagtgag	ctaactcaca	ttaattgcgt	tgcgctcact	gcccgccttc	cagtcgggaa	5940
	acctgtcgtg	ccagcggatc	cgcattctca	ttagtcagca	accatagtcc	cggccctaac	6000
	tccgccccat	ccgcccctaa	ctccgcccag	ttccgcccac	tctccgcccc	atggctgact	6060
30	aatTTTTTTT	atttatgcag	aggccgaggc	cgcctcggcc	tctgagctat	tccagaagta	6120
	gtgaggaggc	TTTTTTggag	gcctaggctt	ttgcaaaaag	ctaacttggt	tattgcagct	6180
35	tataatgggt	acaaataaag	caatagcatc	acaaatttca	caaataaagc	atttttttca	6240
	ctgcattcta	gttgtgggtt	gtccaaactc	atcaatgtat	cttatcatgt	ctgtccgctt	6300
	cctcgctcac	tgactcgctg	cgctcggctg	ttcggtcgcg	gcgagcggta	tcagctcact	6360
40	caaaggcggg	aatacggtta	tccacagaat	caggggataa	cgcaggaaag	aacatgtgag	6420
	caaaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	gttgctggcg	tttttccata	6480
45	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	aagtcagagg	tggcgaaacc	6540
	cgacaggact	ataaagatac	caggcgTTTT	cccctggaag	ctccctcgctg	cgctctcctg	6600
	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	cccttcggga	agcgtggcgc	6660
50	tttctcatag	ctcacgctgt	aggtatctca	gttcggtgta	ggtcgttcgc	tccaagctgg	6720
	gctgtgtgca	cgaaccccc	gttcagcccg	accgctgcgc	cttatccggt	aactatcgtc	6780
55	ttgagtccaa	cccggtaaga	cacgacttat	cgccactggc	agcagccact	ggtaacagga	6840
	ttagcagagc	gaggtatgta	ggcggtgcta	cagagttctt	gaagtgggtg	cctaactacg	6900
	gctacactag	aagaacagta	tttggtatct	gcgctctgct	gaagccagtt	accttcggaa	6960
60	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	tggtagcggg	ggTTTTTTtg	7020
	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	agaagatcct	ttgatctttt	7080
65	ctacggggct	tgacgctcag	tggaacgaaa	actcacgtta	agggattttg	gtcatgagat	7140
	tatcaaaaag	gatcttcacc	tagatccttt	taaattaaaa	atgaagtttt	aatcaatct	7200
70	aaagtatata	tgagtaaact	tggtctgaca	gttagaaaaa	ctcatcgagc	atcaaatgaa	7260

	actgcaat	ttcatatca	ggattatcaa	taccatattt	ttgaaaaagc	cgtttctgta	7320			
	atgaaggaga	aaactcaccg	aggcagttcc	ataggatggc	aagatcctgg	tatcgggtctg	7380			
5	cgattccgac	tcgtccaaca	tcaataacaac	ctattaat	cccctcgta	aaaataaggt	7440			
	tatcaagtga	gaaatcacca	tgagtgcga	ctgaatccgg	tgagaatggc	aacagcttat	7500			
10	gcatttcttt	ccagacttgt	tcaacaggcc	agccattacg	ctcgtcatca	aaatcactcg	7560			
	catcaaccaa	accgttattc	attcgtgatt	gcgcctgagc	gagacgaaat	acgcgatcgc	7620			
	tgttaaaagg	acaattacaa	acaggaatcg	aatgcaaccg	gcgcaggaac	actgccagcg	7680			
15	catcaacaat	at	tttcac	ct	gaatcaggat	attc	ttctaa	tacctggaat	gctgtttttc	7740
	cggggatcgc	agtggtgagt	aaccatgcat	catcaggagt	acggataaaa	tgcttgatgg	7800			
20	tcggaagagg	cataaattcc	gtcagccagt	ttagtctgac	catctcatct	gtaacatcat	7860			
	tggcaacgct	acctttgcca	tgtttcagaa	acaactctgg	cgcatcgggc	ttcccataca	7920			
	atcgatagat	tgtcgcacct	gattgcccga	cattatcgcg	agccatttta	tacccatata	7980			
25	aatcagcatc	catgttgga	tttaatcgcg	gcctagagca	agacgtttcc	cgttgaatat	8040			
	ggctcataac	accccttgta	ttactgttta	tgtaagcaga	cagttttatt	gttcatgatg	8100			
30	atatat	ttttgtgca	atgtaacatc	agagattttg	agacacaaca	attgtcgaca	8160			
	ttgattattg	actagttatt	aatagtaatc	aattacgggg	tcattagttc	atagcccata	8220			
	tatggagttc	cgcgttacat	aacttacggg	aaatggccc	cctggctgac	cgcccaacga	8280			
35	ccccgc	ccca	ttgacgtcaa	taatgacgta	tgttcccata	gtaacgcca	tagggacttt	8340		
	ccattgacgt	caatgggtg	agtattttacg	gtaaactgcc	cacttggcag	tacatcaagt	8400			
40	gtatcatatg	ccaagtacgc	cccctattga	cgtcaatgac	ggtaa	aatggc	ccgcctggca	8460		
	ttatgcccag	tacatgacct	tatgggactt	tcctacttgg	cagtacatct	acgtattagt	8520			
	catcgctatt	accatggctg	aggtgagccc	cacgttctgc	ttcactctcc	ccatctcccc	8580			
45	cccctccca	cccccaat	ttgtatttatt	tattttttaa	ttattttgtg	cagcgatggg	8640			
	ggcggggggg	gggggggggc	gcgcgccagg	cggggcgggg	cggggcgagg	ggcggggcg	8700			
50	ggcgaggcgg	agaggtgcgg	cggcagcca	tcagagcggc	gcgctccgaa	agtttccttt	8760			
	tatggcgagg	cggcggcggc	ggcggcccta	taaaaagcga	agcgcgcggc	gggcgggagt	8820			
	cgctgcgcgc	tgccttcgcc	ccgtgccccg	ctccgcgcgc	gcctcgcgc	gcccgc	8880			
55	gctctgactg	accgcgttac	tcccacaggt	gagcgggcgg	gacggccctt	ctcctccggg	8940			
	ctgtaattag	cgcttggttt	aatgacggct	tgtttctttt	ctgtggctgc	gtgaaagcct	9000			
60	tgaggggctc	cgggagggcc	ctttgtgcgg	ggggagcggc	tcgggggggtg	cgtgcgtgtg	9060			
	tgtgtgcgtg	gggagcgccg	cgtgcggctc	cgcgctgccc	ggcggtgtg	agcgctgcgg	9120			
	gcgcggcgcg	gggctttgtg	cgctccgcag	tgtgcgcgag	gggagcgcg	ccgggggcgg	9180			
65	tgccccgcgg	tgcggggggg	gctgcgaggg	gaacaaaggc	tgcgtgcggg	gtgtgtgcgt	9240			
	gggggggtga	gcaggggggtg	tgggcgcgtc	ggtcgggctg	caaccccc	tgcaccccc	9300			
70	tccccgagtt	gctgagcacg	gcccggcttc	gggtgcgggg	ctccgtacgg	ggcgtggcgc	9360			

	ggggctcgcc	gtgccgggcg	gggggtggcg	gcaggtgggg	gtgccgggcg	gggcggggcc	9420
	gcctcgggccc	ggggagggct	cgggggaggg	gcgcggcggc	ccccggagcg	ccggcgggctg	9480
5	tcgaggcgcg	gcgagccgca	gccattgcct	tttatggtaa	tcgtgcgaga	gggcgcaggg	9540
	acttcctttg	tcccaaactct	gtgcggagcc	gaaatctggg	aggcgccgcc	gcacccccctc	9600
	tagcgggcg	ggggcgaagc	ggtgcggcg	cggcaggaag	gaaatgggcg	gggagggcct	9660
10	tcgtgcgtcg	ccgcgccgcc	gtcccccttct	ccctctccag	cctcggggct	gtccgcgggg	9720
	ggacggctgc	cttcgggggg	gacggggcag	ggcgggggttc	ggcttctggc	gtgtgaccgg	9780
15	cggctctaga	gcctctgcta	accatgttca	tgccttcttc	tttttctac	agctcctggg	9840
	caacgtgctg	gttattgtgc	tgtctcatca	ttttggcaaa	gaatt		9885
20	<210> 3						
	<211> 3384						
	<212> ДНК						
	<213> штучна послідовність						
25	<220>						
	<223> pGM299 плазміда						
	<400> 3						
30	tcaatattgg	ccattagcca	tattattcat	tggttatata	gcataaatca	atattggcta	60
	ttggccattg	catacgttgt	atctatatca	taatatgtac	atttatattg	gtcatgttcc	120
	aatatgaccg	ccatgtttgg	attgattatt	gactagttat	taatagtaat	caattacggg	180
35	gtcattagtt	catagcccat	atatggagtt	ccgcgttaca	taacttacgg	taaatggccc	240
	gcctggctga	ccgcccacg	acccccgccc	attgacgtca	ataatgacgt	atgttcccat	300
	agtaacgcca	atagggactt	tccattgacg	tcaatgggtg	gagtatttac	ggtaaactgc	360
40	ccacttggca	gtacatcaag	tgtatcatat	gccaagtccg	ccccctattg	acgtcaatga	420
	cggtaaatgg	ccgcctggc	attatgcccc	gtacatgacc	ttacgggact	ttcctacttg	480
45	gcagtacatc	tacgtattag	tcatcgctat	taccatgggtg	atgcggtttt	ggcagtacac	540
	caatgggctg	ggatagcggg	ttgactcacg	gggattttcca	agtctccacc	ccattgacgt	600
	caatgggagt	ttgttttggc	acaaaaatca	acgggacttt	ccaaaatgtc	gtaataaccc	660
50	cgccccgttg	acgcaaattg	gcggtagggc	tgtacgggtg	gagggtctata	taagcagagc	720
	tcgttttagtg	aaccgtcaga	tcactagaag	ctttatttgcg	gtagttttatc	acagttaaatt	780
55	tgctaacgca	gtcagtgtct	ctgacacaac	agtctcgaac	ttaagctgca	gaagttgggtc	840
	gtgaggcact	gggcaggtaa	gtatcaagg	tacaagacag	gtttaaggag	accaatagaa	900
	actgggcttg	tcgagacaga	gaagactctt	gcgtttctga	taggcaccta	ttggtcttac	960
60	tgacatccac	tttgcccttc	tctccacagg	tgtccactcc	cagttcaatt	acagctctta	1020
	aggctagagt	acttaatacg	actcactata	ggctagcctc	gagaattcga	ttatgcccct	1080
65	aggaccagaa	gaaagaagat	tgcttcgctt	gattttggctc	ctttacagca	ccaatccata	1140
	tccaccaagt	ggggaaggga	cggccagaca	acgccgacga	gccaggagaa	ggtggagaca	1200
70	acagcaggat	caaattagag	tcttggtaga	aagactccaa	gagcagggtgt	atgcagttga	1260

	ccgcctggct	gacgaggctc	aacacttggc	tatacaacag	ttgcctgacc	ctcctcattc	1320
	agcttagaat	cactagtga	ttcacgcgtg	gtacctctag	agtcgacccg	ggcggccgct	1380
5	tcgagcagac	atgataagat	acattgatga	gtttggacaa	accacaacta	gaatgcagtg	1440
	aaaaaaatgc	tttatttgtg	aaatttgtga	tgctattgct	ttatttgtaa	ccattataag	1500
10	ctgcaataaa	caagttaaca	acaacaattg	cattcatttt	atgtttcagg	ttcaggggga	1560
	gatgtgggag	gtttttttaa	gcaagtaaaa	cctctacaaa	tgtggtaaaa	tcgataagga	1620
	tccgtcgacc	aattgtttgtg	tctcaaaatc	tctgatgtta	cattgcacaa	gataaaaata	1680
15	tatcatcatg	aacaataaaa	ctgtctgctt	acataaacag	taatacaagg	ggtgttatga	1740
	gccatattca	acgggaaacg	tcttgctcta	ggccgcgatt	aaattccaac	atggatgctg	1800
20	atttatatgg	gtataaatgg	gctcgcgata	atgtcgggca	atcagggtgcg	acaatctatc	1860
	gattgtatgg	gaagcccgat	gcgccagagt	tgtttctgaa	acatggcaaa	ggtagcgttg	1920
	ccaatgatgt	tacagatgag	atggtcagac	taaactggct	gacggaattt	atgcctcttc	1980
25	cgaccatcaa	gcattttatc	cgtactcctg	atgatgcatg	gttactcacc	actgcgatcc	2040
	ccggaaaaac	agcattccag	gtattagaag	aatatcctga	ttcagggtgaa	aatattgttg	2100
30	atgcgctggc	agtgttcctg	cgccggttgc	attcgattcc	tgtttgtaat	tgtcctttta	2160
	acagcgatcg	cgtatttcgt	ctcgcctcagg	cgcaatcacg	aatgaataac	ggtttggttg	2220
	atgcgagtga	ttttgatgac	gagcgtaatg	gctggcctgt	tgaacaagtc	tggaaagaaa	2280
35	tgcataagct	gttgccattc	tcaccggatt	cagtcgtcac	tcattggtgat	ttctcacttg	2340
	ataaccttat	ttttgacgag	gggaaattaa	taggttgat	tgatgttgga	cgagtcggaa	2400
40	tcgcagaccg	ataccaggat	cttgccatcc	tatggaactg	cctcgggtgag	ttttctcctt	2460
	cattacagaa	acggcttttt	caaaaatatg	gtattgataa	tcctgatatg	aataaattgc	2520
	agtttcattt	gatgctcgat	gagtttttct	aactgtcaga	ccaagtttac	tcatatatac	2580
45	tttagattga	tttaaaactt	cattttttaat	ttaaaaggat	ctagggtgaag	atcctttttg	2640
	ataatctcat	gaccaaatac	ccttaacgtg	agttttcgtt	ccactgagcg	tcagaccccg	2700
50	tagaaaagat	caaaggatct	tcttgagatc	ctttttttct	gcgcgtaatc	tgctgcttgc	2760
	aaacaaaaaa	accaccgcta	ccagcgggtg	tttgtttgcc	ggatcaagag	ctaccaactc	2820
	tttttccgaa	ggtaactggc	ttcagcagag	cgcagatacc	aaatactggt	cttctagtgt	2880
55	agccgtagtt	aggccaccac	ttcaagaact	ctgtagcacc	gcctacatac	ctcgctctgc	2940
	taatcctggt	accagtggct	gctgccagtg	gcgataagtc	gtgtcttacc	gggttggaact	3000
60	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	aacgggggggt	tcgtgcacac	3060
	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	cctacagcgt	gagctatgag	3120
	aaagcgccac	gcttcccga	gggagaaaagg	cggacaggta	tccggtaagc	ggcagggctcg	3180
65	gaacaggaga	gcgcacgagg	gagcttccag	ggggaaacgc	ctggatatctt	tatagtcctg	3240
	tcgggtttcg	ccacctctga	cttgagcgtc	gattttttgtg	atgctcgtca	ggggggcgga	3300
70	gcctatggaa	aaacgccagc	aacgcggcct	ttttacggtt	cctggccttt	tgctggcctt	3360

ttgctcacat ggctcgacag atct

3384

5 <210> 4
 <211> 6264
 <212> ДНК
 <213> штучна послідовність

10 <220>
 <223> pGM301 плазміда

<400> 4
 attgattatt gactagttat taatagtaat caattacggg gtcattagtt catagcccat 60
 15 atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgcccacg 120
 acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 180
 20 tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca gtacatcaag 240
 tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaattg cccgcctggc 300
 attatgccc a gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag 360
 25 tcatcgctat taccatggtc gaggtgagcc ccacgttctg cttcactctc cccatctccc 420
 cccctcccc accccaatt ttgtatttat ttatttttta attattttgt gcagcgatgg 480
 gggcgggggg gggggggggg cgcgcgccag gcggggaggg gcggggcaag gggcggggag 540
 30 gggcgaggcg gaaagggtgcg gcggcagcca atcaaagcgg cgcgctccga aagtttcctt 600
 ttatggcgag gcggcgggcg cggcggccct ataaaaagcg aagcgcgcg cgggcgggag 660
 35 tcgctgcgag ctgccttcgc cccgtgcccc gctccgccc cgcttcgag cgcccgcccc 720
 ggctctgact gaccgcgtta ctcccacagg tgagcggggg ggacggccct tctcctccgg 780
 gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc 840
 40 ttgaggggct ccgggagggc cttttgtgag ggggggagcg ctcggggggg gcgtgagtg 900
 gtgtgtgagc ggggagcgcc gcgtgaggct ccgcgctgcc cggcggctgt gagcgctgag 960
 45 ggcgcgggcg ggggctttgt gcgctccgca gtgtgagcga ggggagcgcg gccgggggag 1020
 gtgccccgag gtgagggggg ggctgagagg ggaacaaagg ctgagtgag ggtgtgtgag 1080
 tggggggggt agcagggggg gtgggagcgt cggctggggt gcaaccccc ctgcaccccc 1140
 50 ctccccgagt tgctgagcac ggcccggtt cgggtgaggg gctccgtacg gggcggtggc 1200
 cggggctcgc cgtgccggg ggggggtggc ggcaggtggg ggtgccggg gggcggggag 1260
 55 cgcttcggg cggggagggc tcgggggagg ggcgagggcg ccccgaggc gccggcgagg 1320
 gtcgagggcg ggcgagccgc agccattgcc ttttatggta atcgtgagag agggcgagc 1380
 gacttccttt gtcccaaata tgtgagggc cgaaatctgg gaggcggcg cgacccccct 1440
 60 ctagcgggag cggggagga cggtgagggc ccggcaggaa ggaaatggg ggggagggcc 1500
 ttcgtgagc gccgcggcgc cgtcccttc tccctctcca gcctcggggc tgtccgaggg 1560
 65 gggagggctg ccttcggggg ggacggggca gggcgggggt cggcttctg cgtgtgagcg 1620
 gcggctctag agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg 1680
 70 gcaacgtgct gggtattgtg ctgtctcatc attttgga agaattcgat tgccatggca 1740

	acatatatcc	agagagtaca	gtgcatctca	acatcactac	tggttggttct	caccacattg	1800
	gtctcgtgtc	agattcccag	ggataggctc	tctaacatag	gggtcatagt	cgatgaaggg	1860
5	aaatcactga	agatagctgg	atcccacgaa	tcgaggtaca	tagtactgag	tctagttccg	1920
	ggggtagact	ttgagaatgg	gtgcggaaca	gcccagggtta	tccagtacaa	gagcctactg	1980
	aacaggctgt	taatcccatt	gagggatgcc	ttagatcttc	aggaggctct	gataactgtc	2040
10	accaatgata	cgacacaaaa	tgccggtgct	ccccagtcga	gattcttcgg	tgctgtgatt	2100
	ggtactatcg	cacttgaggt	ggcgacatca	gcacaaatca	ccgcagggat	tgcactagcc	2160
15	gaagcgaggg	aggccaaaag	agacatagcg	ctcatcaaag	aatcgatgac	aaaaacacac	2220
	aagtctatag	aactgctgca	aaacgctgtg	ggggaacaaa	ttcttgctct	aaagacactc	2280
	caggatttcg	tgaatgatga	gatcaaacc	gcaataagcg	aattaggctg	tgagactgct	2340
20	gccttaagac	tggttataaa	attgacacag	cattactccg	agctgttaac	tgcgttcggc	2400
	tcgaatttcg	gaaccatcgg	agagaagagc	ctcacgctgc	aggcgctgtc	ttcactttac	2460
25	tctgctaaca	ttactgagat	tatgaccaca	atcaggacag	ggcagtctaa	catctatgat	2520
	gtcatttata	cagaacagat	caaaggaacg	gtgatagatg	tggatctaga	gagatacatg	2580
	gtcacctgt	ctgtgaagat	ccctattctt	tctgaagtcc	cagggtgtgct	catacacaa	2640
30	gcatcatcta	tttcttacia	catagacggg	gaggaatggt	atgtgactgt	ccccagccat	2700
	atactcagtc	gtgcttcttt	cttagggggt	gcagacataa	ccgatttgtg	tgagtccaga	2760
35	ttgacctata	tatgccccag	ggatcccgc	caactgatac	ctgacagcca	gcaaaagtgt	2820
	atcctggggg	acacaacaag	gtgtcctgtc	acaaaagttg	tggacagcct	tatccccaag	2880
	tttgcttttg	tgaatggggg	cgttgttgct	aactgcatag	catccacatg	tacctgcggg	2940
40	acaggccgaa	gaccaatcag	tcaggatcgc	tctaaagggtg	tagtattcct	aacctatgac	3000
	aactgtgggtc	ttataggtgt	caatggggta	gaattgtatg	ctaaccggag	agggcacgat	3060
45	gccacttggg	gggtccagaa	cttgacagtc	ggtcctgcaa	ttgctatcag	accggttgat	3120
	atttctctca	accttgctga	tgctacgaat	ttcttgcaag	actctaaggc	tgagcttgag	3180
	aaagcacgga	aaatcctctc	ggaggtaggt	agatggtaca	actcaagaga	gactgtgatt	3240
50	acgatcatag	tagttatggt	cgtaatatgt	gtggtcatta	tagtgatcat	catcgtgctt	3300
	tatagactca	gaaggtgaaa	tcactagtga	attcactcct	cagggtgcagg	ctgcctatca	3360
55	gaaggtgggtg	gctgggtgtg	ccaatgccct	ggctcacaaa	taccactgag	atctttttcc	3420
	ctctgccaaa	aattatgggg	acatcatgaa	gccccttgag	catctgactt	ctggctaata	3480
	aaggaaatth	atthttcattg	caatagtgtg	ttggaattht	ttgtgtctct	cactcggaag	3540
60	gacatatggg	agggcaaath	atthaaaaca	tcagaatgag	tatttggtth	agagthtggc	3600
	aacatatgcc	catatgctgg	ctgccatgaa	caaaggthtgg	ctataaagag	gtcatcagta	3660
65	tatgaaacag	ccccctgctg	tccattcctt	attccataga	aaagccttga	cttgaggtht	3720
	gattththth	atattthgtt	ttgtgtthatt	ththththth	acatccctaa	aattththctt	3780
70	acatgtthth	ctagccagat	thththctct	ctcttgacta	ctcccagth	tagctgtccc	3840

	tcttctctta	tggagatccc	tcgacctgca	gcccgaagctt	ggcgtaatca	tgggtcatagc	3900
	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	caacatacga	gccggaagca	3960
5	taaagtgtaa	agcctgggggt	gcctaattgag	tgagctaact	cacattaatt	gcgttgcgct	4020
	cactgcccgc	tttccagtcg	ggaaacctgt	cgtgccagcg	gatccgcatac	tcaattagtc	4080
10	agcaaccata	gtcccgcccc	taactccgcc	catcccgcgc	ctaactccgc	ccagttccgc	4140
	ccatttctccg	ccccatggct	gactaatttt	ttttatttat	gcagaggccg	aggccgcctc	4200
	ggcctctgag	ctattccaga	agtagtgagg	aggctttttt	ggaggcctag	gcttttgcaa	4260
15	aaagctaact	tgtttattgc	agcttataat	ggttacaaat	aaagcaatag	catcacaaat	4320
	ttcacaaata	aagcattttt	ttcactgcat	tctagtgtgtg	gtttgtccaa	actcatcaat	4380
20	gtatcttatc	atgtctgtcc	gcttcctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	4440
	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatagc	gttatccaca	gaatcagggg	4500
	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	4560
25	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	4620
	gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	ataccaggcg	tttccccctg	4680
30	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	4740
	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	ctgtaggtat	ctcagttcgg	4800
	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgacgaacc	ccccgttcag	cccgaccgct	4860
35	gcgccttatc	cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	4920
	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	4980
40	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	agtatttggt	atctgcgctc	5040
	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	5100
	ccgctggtag	cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	5160
45	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	gaaaactcac	5220
	gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	cacctagatc	cttttaaat	5280
50	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttaga	5340
	aaaactcatc	gagcatcaaa	tgaaactgca	atttattcat	atcaggatta	tcaataccat	5400
	atttttgaaa	aagccgtttc	tgtaatgaag	gagaaaaactc	accgaggcag	ttccatagga	5460
55	tggcaagatc	ctggtatcgg	tctgcgattc	cgactcgtcc	aacatcaata	caacctatta	5520
	atttcccctc	gtcaaaaata	aggttatcaa	gtgagaaatc	accatgagtg	acgactgaat	5580
60	ccggtgagaa	tggcaacagc	ttatgcattt	ctttccagac	ttgttcaaca	ggccagccat	5640
	tacgctcgtc	atcaaaatca	ctcgcaccaa	ccaaaccggt	attcattcgt	gattgcgcct	5700
	gagcgagacg	aaatacgcga	tcgctgttaa	aaggacaatt	acaaacagga	atcgaatgca	5760
65	accggcgagc	gaacactgcc	agcgcaccaa	caatatcttc	acctgaatca	ggatattctt	5820
	ctaatacctg	gaatgctgtt	tttccgggga	tcgcagtggt	gagtaaccat	gcatcatcag	5880
70	gagtacggat	aaaatgcttg	atggtcggaa	gaggcataaa	ttccgtcagc	cagtttagtc	5940

	tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc agaaacaact	6000
	ctggcgcgcatc gggcttccca tacaatcgat agattgtcgc acctgattgc cgcacattat	6060
5	cgcgagccca ttataacca tataaatcag catccatgtt ggaatttaaat cgcggcctag	6120
	agcaagacgt ttcccgttga atatggctca taacaccctt tgtattactg tttatgtaag	6180
10	cagacagttt tattgttcat gatgatatat ttttatcttg tgcaatgtaa catcagagat	6240
	tttgagacac aacaattggg cgac	6264
15	<210> 5 <211> 6522 <212> ДНК <213> штучна послідовність	
20	<220> <223> pGM303 плазміда	
	<400> 5 attgattatt gactagttat taatagtaat caattacggg gtcattagtt catagcccat	60
25	atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgcccacg	120
	acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt	180
30	tccattgacg tcaatgggtg gagtattttac ggtaaaactgc ccacttgga gtacatcaag	240
	tgtatcatat gccaagtacg cccctatttg acgtcaatga cggtaaattgg cccgcctggc	300
	attatgccc gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag	360
35	tcatcgctat taccatgggc gaggtgagcc ccacgttctg cttcactctc cccatctccc	420
	ccccctcccc acccccaatt ttgtattttat ttatTTTTTTa attattttgt gcagcgatgg	480
40	gggcgggggg gggggggggg cgcgcgccag gcggggcggg gcggggcgag gggcggggcg	540
	gggcgaggcg gaaagggtgc gcggcagcca atcaaagcgg cgcgctccga aagtttctt	600
	ttatggcgag gcggcgggcg cggcggccct ataaaaagcg aagcgcgcg cgggcgggag	660
45	tcgctgcgcg ctgccttcgc cccgtgcccc gctccgccc gcctcgcgc cggcgcccc	720
	ggctctgact gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctccgg	780
50	gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc	840
	ttgaggggct ccgggagggc cttttgtgcg gggggagcgg ctcggggggt gcgtgcgtgt	900
	gtgtgtgcgt ggggagcgcc gcgtgcggct ccgcgtgcc cggcggtgt gagcgtgcg	960
55	ggcgcggcgc ggggctttgt gcgctccgca gtgtgcgcga ggggagcgcg gccgggggcg	1020
	gtgccccgcg gtgcgggggg ggctgcgagg ggaacaaagg ctgcgtgcgg ggtgtgtgcg	1080
60	tggggggggt agcaggggggt gtgggcgcgt cggctgggct gcaaccccc ctgcaccccc	1140
	ctccccgagt tgctgagcac ggcccggctt cgggtgcggg gctccgtacg gggcgtggcg	1200
	cggggctcgc cgtgccgggc ggggggtggc ggcaggtggg ggtgccgggc ggggcggggc	1260
65	cgctcgggc cggggagggc tcgggggagg ggcgcggcgg cccccggag gccggcggt	1320
	gtcgaggcgc ggcgagccgc agccattgcc ttttatggta atcgtgcgag agggcgagg	1380
70	gacttccttt gtcccaaadc tgtgcggagc cgaaatcttg gaggcggcg cgcaccccc	1440

	ctagcggg	cggggcgaag	cggtgcggcg	ccggcaggaa	ggaaatgggc	ggggagggcc	1500
	ttcgtg	gccgcgccgc	cgtcccttc	tccctctcca	gcctcggggc	tgtccgcggg	1560
5	gggacggggc	agggcggggt	tcggcttctg	gcgtgtgacc	ggcggctcta	gagcctctgc	1620
	taacatg	catgccttct	tctttttcct	acagctcctg	ggcaacgtgc	tggttattgt	1680
10	gctgtctcat	cattttggca	aagaattcct	cgagcatgtg	gtctgagtta	aaaatcagga	1740
	gcaacgacgg	aggtgaagga	ccagaggacg	ccaacgaccc	ccgggggaaag	gggggtgcaac	1800
	acatccatat	ccagccatct	ctacctgttt	atggacagag	ggttagggat	ggtgataggg	1860
15	gcaaacgtga	ctcgtactgg	tctacttctc	ctagtggtag	caccacaaaa	ccagcatcag	1920
	gttgggagag	gtcaagtaaa	gccgacacat	ggttgtctgat	tctctcattc	accagtg	1980
20	ctttgtcaat	tgccacagt	atcatctgta	tcataatttc	tgctagacaa	gggtatagta	2040
	tgaaagagta	ctcaatgact	gtagaggcat	tgaacatgag	cagcaggggag	gtgaaagagt	2100
	cacttaccag	tctaataagg	caagagggtta	tagcaagggc	tgtcaacatt	cagagctctg	2160
25	tgcaaaccgg	aatcccagtc	ttgttgaaca	aaaacagcag	ggatgtcatc	cagatgattg	2220
	ataagtcgtg	cagcagacaa	gagctcactc	agcactgtga	gagtacgatc	gcagtcacc	2280
30	atgccgatgg	aattgcccc	cttgagccac	atagtttctg	gagatgccct	gtcggagaa	2340
	cgtatcttag	ctcagatcct	gaaatctcat	tgctgcctgg	tccgagcttg	ttatctgggt	2400
	ctacaacgat	ctctggatgt	gttaggtctc	cttactctc	aattggcgag	gcaatctatg	2460
35	cctattcatc	aatctcatt	acacaagggt	gtgctgacat	agggaatca	tatcagggtc	2520
	tgcaagctagg	gtacatatca	ctcaattcag	atatgtttcc	tgatcttaac	cccgtagtgt	2580
40	cccacactta	tgacatcaac	gacaatcgga	aatcatgctc	tgtggtggca	accgggacta	2640
	gggggttatca	gctttgctcc	atgccgactg	tagacgaaag	aaccgactac	tctagtgatg	2700
	gtattgagga	tctggtcctt	gatgtcctgg	atctcaaagg	gagaactaag	tctcaccggt	2760
45	atcgcaacag	cgaggtagat	cttgatcacc	cgttctctgc	actatacccc	agtgtaggca	2820
	acggcattgc	aacagaaggc	tcattgatat	ttcttgggta	tggtggacta	accaccctc	2880
50	tgcaaggtga	tacaaaatgt	aggacccaag	gatgccaa	gggtgtcgaa	gacacatgca	2940
	atgaggctct	gaaaattaca	tggttaggag	ggaaacaggt	ggtcagcgtg	atcatccagg	3000
	tcaatgacta	tctctcagag	aggccaaaga	taagagtcac	aaccattcca	atcactcaaa	3060
55	actatctcgg	ggcggaagg	agattattaa	aattgggtga	tcgggtgtac	atctatacaa	3120
	gatcatcagg	ctggcactct	caactgcaga	taggagtact	tgatgtcagc	caccctttga	3180
60	ctatcaactg	gacacctcat	gaagccttgt	ctagaccagg	aaataaagag	tgcaattgg	3240
	acaataagtg	tccgaaggaa	tgcatatcag	gcgtatacac	tgatgcttat	ccattgtccc	3300
	ctgatgcagc	taacgtcgct	accgtcacgc	tatatgccaa	tacatcgctg	gtcaacccaa	3360
65	caatcatgta	ttctaacact	actaacatta	taaatatgtt	aaggataaag	gatgttcaat	3420
	tagaggctgc	atataccacg	acatcgtgta	tcacgcattt	tggtaaaggc	tactgctttc	3480
70	acatcatcga	gatcaatcag	aagagcctga	ataccttaca	gccgatgctc	tttaagacta	3540

	gcatccctaa attatgcaag gccgagtctt aagcgggccgc gcatgcgaat tcactcctca	3600
	ggtgcaggct gcctatcaga aggtggtggc tgggtgtggcc aatgccctgg ctcaaaata	3660
5	ccactgagat ctttttccct ctgccaaaaa ttatggggac atcatgaagc cccttgagca	3720
	tctgacttct ggctaataaa ggaaatttat tttcattgca atagtgtgtt ggaatttttt	3780
10	gtgtctctca ctcggaagga catatgggag ggcaaatcat ttaaaacatc agaatgagta	3840
	tttggtttag agtttgga catatgccca tatgctggct gccatgaaca aagggttgct	3900
	ataaagaggt catcagtata tgaaacagcc ccctgctgtc tttccttat tccatagaaa	3960
15	agccttgact tgaggttaga ttttttttat attttgtttt gtgttatttt tttctttaac	4020
	atccctaaaa ttttccttac atgttttact agccagattt ttcctcctct cctgactact	4080
20	cccagtcata gctgtccctc ttctcttatg gagatccctc gacctgcagc ccaagcttgg	4140
	cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca	4200
	acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg agctaactca	4260
25	cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagcgga	4320
	tccgcatctc aattagtcag caaccatagt cccgccccta actccgcca tcccgccct	4380
	aactccgccc agttccgccc attctccgcc catggctga ctaatttttt ttatttatgc	4440
30	agaggccgag gccgcctcgg cctctgagct attccagaag tagtgaggag gcttttttgg	4500
	aggcctaggc ttttgcaaaa agctaacttg tttattgcag cttataatgg ttacaaataa	4560
35	agcaatagca tcacaaattt cacaaataaa gcattttttt cactgcattc tagttgtggt	4620
	ttgtccaaac tcatcaatgt atcttatcat gtctgtccgc ttcctcgctc actgactcgc	4680
40	tgcgctcggc cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc	4740
	tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg	4800
	ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg	4860
45	agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat	4920
	accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgtta	4980
50	ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct	5040
	gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg cacgaacccc	5100
	ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aaccggtaa	5160
55	gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg	5220
	taggcggtgc tacagagttc ttgaagtggg ggcctaacta cggctacact agaagaacag	5280
60	tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt	5340
	gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta	5400
	cgcgagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc	5460
65	agtggaacga aaactcacgt taagggattt tggctatgag attatcaaaa aggatcttca	5520
	cctagatcct tttaaattaa aaatgaagtt ttaaataaat ctaaagtata tatgagtaaa	5580
70	cttggctctga cagttagaaa aactcatcga gcatcaaatg aaactgcaat ttattcatat	5640

	caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga gaaaactcac	5700
	cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg actcgtccaa	5760
5	catcaatata acctattaat ttcccctcgt caaaaataag gttatcaagt gagaaatcac	5820
	catgagtgac gactgaatcc ggtgagaatg gcaacagctt atgcattttct ttccagactt	5880
10	gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcatacaacc aaaccgttat	5940
	tcattcgtga ttgcgctga gcgagacgaa atacgcgcatc gctgttaaaa ggacaattac	6000
	aaacaggaat cgaatgcaac cggcgcagga aactgccag cgcatacaaca atattttcac	6060
15	ctgaatcagg atattcttct aatacctgga atgctgtttt tccggggatc gcagtgggtga	6120
	gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcgggaaga ggcataaatt	6180
20	ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg ctacctttgc	6240
	catgtttcag aaacaactct ggcgcacatcg gcttcccata caatcgatag attgtcgcac	6300
	ctgattgccc gacattatcg cgagcccatt tatacccata taaatcagca tccatgttgg	6360
25	aatttaatcg cggcctagag caagacgttt cccgttgaat atggctcata acacccttg	6420
	tattactgtt tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg	6480
30	caatgtaaca tcagagattt tgagacacaa caattgggtcg ac	6522
	<210> 6	
	<211> 574	
	<212> ДНК	
35	<213> штучна послідовність	
	<220>	
	<223> hCEF промотор	
40	<400> 6	
	agatctgtta cataacttat ggtaaattggc ctgcctggct gactgcccac tgacccttgc	60
	ccaatgatgt caataatgat gtatgttccc atgtaatgcc aatagggact ttccattgat	120
45	gtcaatgggt ggagtattta tggttaactgc ccacttggca gtacatcaag tgtatcatat	180
	gccaaagtatg ccccctattg atgtcaatga tggtaaattgg cctgcctggc attatgcccac	240
50	gtacatgacc ttatgggact ttcctacttg gcagtacatc tatgtattag tcattgctat	300
	taccatggga attcactagt ggagaagagc atgcttgagg gctgagtgcc cctcagtggg	360
	cagagagcac atggcccaca gtccctgaga agttgggggg aggggtgggc aattgaactg	420
55	gtgcctagag aagggtggggc ttgggtaaac tgggaaagtg atgtggtgta ctggctccac	480
	ctttttcccc aggggtggggg agaaccatat ataagtgcag tagtctctgt gaacattcaa	540
60	gcttctgcct tctccctcct gtgagtttgc tagc	574
	<210> 7	
	<211> 4459	
	<212> ДНК	
65	<213> штучна послідовність	
	<220>	
	<223> кодон-оптимізований збіднений CpG трансген CFTR	
70	<400> 7	

	gctagccacc	atgcagagaa	gccctctgga	gaaggcctct	gtggtgagca	agctgttctt	60
	cagctggacc	aggcccatcc	tgaggaaggg	ctacaggcag	agactggagc	tgtctgacat	120
5	ctaccagatc	ccctctgttg	actctgctga	caacctgtct	gagaagctgg	agagggagtg	180
	ggatagagag	ctggccagca	agaagaaccc	caagctgac	aatgccctga	ggagatgctt	240
10	cttctggaga	ttcatgttct	atggcatctt	cctgtacctg	ggggaagtga	ccaaggctgt	300
	gcagcctctg	ctgctgggca	gaatcattgc	cagctatgac	cctgacaaca	aggaggagag	360
	gagcattgcc	atctacctgg	gcattggcct	gtgcctgctg	ttcattgtga	ggaccctgct	420
15	gctgcaccct	gccatctttg	gcctgcacca	cattggcatg	cagatgagga	ttgccatggt	480
	cagcctgatc	tacaagaaaa	ccctgaagct	gtccagcaga	gtgctggaca	agatcagcat	540
20	tggccagctg	gtgagcctgc	tgagcaacaa	cctgaacaag	tttgatgagg	gcctggccct	600
	ggcccacttt	gtgtggattg	cccctctgca	ggtggccctg	ctgatgggcc	tgatttgga	660
	gctgctgcag	gcctctgcct	tttgtggcct	gggcttctctg	attgtgctgg	ccctgtttca	720
25	ggctggcctg	ggcaggatga	tgatgaagta	cagggaccag	agggcaggca	agatcagtga	780
	gaggctggtg	atcacctctg	agatgattga	gaacatccag	tctgtgaagg	cctactgttg	840
30	ggaggaagct	atggagaaga	tgattgaaaa	cctgaggcag	acagagctga	agctgaccag	900
	gaaggctgcc	tatgtgagat	acttcaacag	ctctgccttc	ttcttctctg	gcttctttgt	960
	ggtgttctctg	tctgtgctgc	cctatgccct	gatcaagggg	atcatcctga	gaaagatttt	1020
35	caccaccatc	agcttctgca	ttgtgctgag	gatggctgtg	accagacagt	tcccctgggc	1080
	tgtgcagacc	tggtatgaca	gcctgggggc	catcaacaag	atccaggact	tcctgcagaa	1140
40	gcaggagtac	aagaccctgg	agtacaacct	gaccaccaca	gaagtgggtga	tggagaatgt	1200
	gacagccttc	tgggaggagg	gctttgggga	gctgtttgag	aaggccaagc	agaacaacaa	1260
	caacagaaag	accagcaatg	gggatgactc	cctgttcttc	tccaacttct	ccctgctggg	1320
45	cacacctgtg	ctgaaggaca	tcaacttcaa	gattgagagg	gggcagctgc	tggctgtggc	1380
	tggatctaca	ggggctggca	agaccagcct	gctgatgatg	atcatggggg	agctggagcc	1440
50	ttctgagggc	aagatcaagc	actctggcag	gatcagcttt	tgcagccagt	tcagctggat	1500
	catgcctggc	accatcaagg	agaacatcat	ctttggagtg	agctatgatg	agtacagata	1560
	caggagtgtg	atcaaggcct	gccagctgga	ggaggacatc	agcaagtttg	ctgagaagga	1620
55	caacattgtg	ctgggggagg	gaggcattac	actgtctggg	ggccagagag	ccagaatcag	1680
	cctggccagg	gctgtgtaca	aggatgctga	cctgtacctg	ctggactccc	cctttggcta	1740
60	cctggatgtg	ctgacagaga	aggagatttt	tgagagctgt	gtgtgcaagc	tgatggccaa	1800
	caagaccaga	atcctggtga	ccagcaagat	ggagcacctg	aagaaggctg	acaagatcct	1860
	gatcctgcat	gagggcagca	gctacttcta	tgggaccttc	tctgagctgc	agaacctgca	1920
65	gcctgacttc	agctctaagc	tgatgggctg	tgacagcttt	gaccagttct	ctgctgagag	1980
	gaggaacagc	atcctgacag	agaccctgca	cagattcagc	ctggagggag	atgcccctgt	2040
70	gagctggaca	gagaccaaga	agcagagctt	caagcagaca	ggggagtttg	gggagaagag	2100

	gaagaactcc	atcctgaacc	ccatcaacag	catcaggaag	ttcagcattg	tgcagaaaac	2160
	ccccctgcag	atgaatggca	ttgaggaaga	ttctgatgag	cccctggaga	ggagactgag	2220
5	cctggtgcct	gattctgagc	aggagagaggc	catcctgcct	aggatctctg	tgatcagcac	2280
	aggccctaca	ctgcaggcca	gaaggaggca	gtctgtgctg	aacctgatga	cccactctgt	2340
10	gaaccagggc	cagaacatcc	acaggaaaac	cacagcctcc	accaggaaaag	tgagcctggc	2400
	ccctcaggcc	aatctgacag	agctggacat	ctacagcagg	aggctgtctc	aggagacagg	2460
	cctggagatt	tctgaggaga	tcaatgagga	ggacctgaaa	gagtgccttct	ttgatgacat	2520
15	ggagagcatc	cctgctgtga	ccacctggaa	cacctacctg	agatacatca	cagtgcacaa	2580
	gagcctgatc	tttgtgctga	tctggtgcct	ggtgatcttc	ctggctgaag	tggctgcctc	2640
20	tctggtggtg	ctgtggctgc	tgggaaacac	cccactgcag	gacaagggca	acagcaccca	2700
	cagcaggaac	aacagctatg	ctgtgatcat	cacctccacc	tccagctact	atgtgttcta	2760
	catctatgtg	ggagtggctg	ataccctgct	ggctatgggc	ttcttttagag	gcctgcccct	2820
25	ggtgcacaca	ctgatcacag	tgagcaagat	cctccaccac	aagatgctgc	actctgtgct	2880
	gcaggctcct	atgagcacc	tgaataccct	gaaggctggg	ggcatcctga	acagattctc	2940
30	caaggatatt	gccatcctgg	atgacctgct	gcctctcacc	atctttgact	tcattccagct	3000
	gctgctgatt	gtgattgggg	ccattgctgt	ggtggcagtg	ctgcagccct	acatctttgt	3060
	ggccacagtg	cctgtgattg	tggccttcat	catgctgagg	gcctactttc	tgcagacctc	3120
35	ccagcagctg	aagcagctgg	agtctgaggg	cagaagcccc	atcttcaccc	acctggtgac	3180
	aagcctgaag	ggcctgtgga	ccctgagagc	ctttggcagg	cagccctact	ttgagaccct	3240
40	gttcacacaag	gccctgaacc	tgcacacagc	caactgggtc	ctctacctgt	ccaccctgag	3300
	atggttccag	atgagaattg	agatgatctt	tgtcatcttc	ttcattgctg	tgaccttcat	3360
	cagcattctg	accacaggag	aggagagagg	cagagtgggc	attatcctga	ccctggccat	3420
45	gaacatcatg	agcacactgc	agtgggcagt	gaacagcagc	attgatgtgg	acagcctgat	3480
	gaggagtgtg	agcagagtgt	tcaagtccat	tgatatgccc	acagagggca	agcctaccaa	3540
50	gagcaccaag	ccctacaaga	atggccagct	gagcaaagtg	atgatcattg	agaacagcca	3600
	tgtgaagaag	gatgatattc	ggcccagtg	aggccagatg	acagtgaagg	acctgacagc	3660
	caagtacaca	gaggggggca	atgctatcct	ggagaacatc	tccttcagca	tctcccctgg	3720
55	ccagagagtg	ggactgctgg	gaagaacagg	ctctggcaag	tctaccctgc	tgtctgcctt	3780
	cctgaggctg	ctgaacacag	aggagagat	ccagattgat	ggagtgtcct	gggacagcat	3840
60	cacactgcag	cagtggagga	aggccttttg	tgtgatcccc	cagaaagtgt	tcattcttcag	3900
	tggcaccttc	aggaagaacc	tggaccctta	tgagcagtg	tctgaccagg	agatttgga	3960
	agtggctgat	gaagtgggcc	tgagaagtgt	gattgagcag	ttccctggca	agctggactt	4020
65	tgtcctgggtg	gatgggggct	gtgtgctgag	ccatggccac	aagcagctga	tgtgcctggc	4080
	cagatcagtg	ctgagcaagg	ccaagatcct	gctgctggat	gagccttctg	cccacctgga	4140
70	tcctgtgacc	taccagatca	tcaggaggac	cctcaagcag	gcctttgctg	actgcacagt	4200

	catcctgtgt gagcacagga ttgaggccat gctggagtgc cagcagttcc tgggtgattga	4260
	ggagaacaaa gtgaggcagt atgacagcat ccagaagctg ctgaatgaga ggagcctggt	4320
5	caggcaggcc atcagcccct ctgatagagt gaagctgttc cccacagga acagctccaa	4380
	gtgcaagagc aagccccaga ttgctgcctt gaaggaggag acagaggagg aagtgcagga	4440
10	caccaggctg tgagggcc	4459
	<210> 8	
	<211> 600	
	<212> ДНК	
15	<213> вірус гепатита сурков	
	<400> 8	
	gggcccaatc aacctctgga ttacaaaatt tgtgaaagat tgactgggtat tcttaactat	60
20	gttgctcctt ttacgctatg tggatacgt gctttaatgc ctttgatatca tgctattgct	120
	tcccgtatgg ctttcatttt ctccctcctt tataaatcct gggtgctgtc tctttatgag	180
	gagttgtggc ccgttgtcag gcaacgtggc gtggtgtgca ctgtgtttgc tgacgcaacc	240
25	cccactgggt ggggcattgc caccacctgt cagctccttt ccgggacttt cgctttcccc	300
	ctccctattg ccacggcgga actcatcgcc gcctgccttg cccgctgctg gacaggggct	360
30	cggctgttgg gactgacaa ttccgtggtg ttgtcgggga aatcatcgtc ctttccttgg	420
	ctgctcgctt gtgttgccac ctggattctg cgcgggacgt ccttctgcta cgccccctcg	480
	gccctcaatc cagcggacct tccttcccgc ggcctgctgc cggctctgcg gcctcttccg	540
35	cgtcttcgcc ttcgccctca gacgagtcgg atctcccttt gggccgcctc cccgcaagct	600
	<210> 9	
40	<211> 7349	
	<212> ДНК	
	<213> штучна послідовність	
	<220>	
45	<223> pGM407 плазміда	
	<400> 9	
	ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat	60
50	tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc	120
	atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat	180
	tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa	240
55	tggcccgctt ggctgaccgc ccaacgacc cggccattg acgtcaataa tgacgtatgt	300
	tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	360
60	aactgcccac ttggcagtac atcaagtgt tcatatgcca agtccgcccc ctattgacgt	420
	caatgacggt aaatggcccc cctggcatta tgcccagtac atgaccttac gggactttcc	480
	tacttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca	540
65	gtacaccaat gggcgtggat agcggtttga ctacgggga tttccaagtc tccaccccat	600
	tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa	660
70	caactgcat cgcccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtc	720

	tatataagca	gagctcgctg	gcttgtaact	cagtctctta	ctaggagacc	agcttgagcc	780
5	tgggtgttcg	ctggttagcc	taacctgggt	ggccaccagg	ggtaaggact	ccttggtcta	840
	gaaagcta	aaacttgctt	gcattagagc	ttatctgagt	caagtgtcct	cattgacgcc	900
	tcactctctt	gaacgggaat	cttccttact	gggttctctc	tctgaccag	gcgagagaaa	960
10	ctccagcagt	ggcgcccgaa	cagggacttg	agtgagagtg	taggcacgta	cagctgagaa	1020
	ggcgtcggac	gcgaaggaag	cgcggggtgc	gacgcgacca	agaaggagac	ttggtgagta	1080
	ggcttctcga	gtgccgggaa	aaagctcgag	cctagttaga	ggactaggag	aggccgtagc	1140
15	cgtaactact	cttgggcaag	tagggcaggc	ggtgggtacg	caatgggggc	ggctacctca	1200
	gcactaaata	ggagacaatt	agaccaattt	gagaaaatac	gacttcgccc	gaacggaaaag	1260
20	aaaaagtacc	aatattaaca	tttaatatgg	gcaggcaagg	agatggagcg	cttcggcctc	1320
	catgagaggt	tgttgagagc	agaggagggg	tgtaaaagaa	tcatagaagt	cctctacccc	1380
	ctagaaccaa	caggatcgga	gggcttaaaa	agtctgttca	atcttgtgtg	cgtgctatat	1440
25	tgcttgacaa	aggaacagaa	agtgaaagac	acagaggaag	cagtagcaac	agtaagacaa	1500
	cactgccatc	tagtggaata	agaaaaaagt	gcaacagaga	catctagtgg	acaaaagaaa	1560
30	aatgacaagg	gaatagcagc	gccacctggt	ggcagtcaga	atcttccagc	gcaacaacaa	1620
	ggaaatgcct	gggtacatgt	acccttgtca	ccgcgcacct	taaatgcgtg	ggtaaaagca	1680
	gtagaggaga	aaaaatttgg	agcagaaata	gtaccatttt	ttttgtttca	agccctatcg	1740
35	aattcccgtt	tgtgctaggg	ttcttaggct	tcttgggggc	tgctggaact	gcaatgggag	1800
	cagcggcgac	agccctgacg	gtccagtctc	agcatttgct	tgctgggata	ctgcagcagc	1860
40	agaagaatct	gctggcggct	gtggaggctc	aacagcagat	gttgaagctg	accatttggg	1920
	gtgttaaaaa	cctcaatgcc	cgcgctcacag	cccttgagaa	gtacctagag	gatcaggcac	1980
	gactaaactc	ctgggggtgc	gcatggaaac	aagtatgtca	taccacagtg	gagtggccct	2040
45	ggacaaatcg	gactccggat	tggcaaaata	tgacttggtt	ggagtgggaa	agacaaatag	2100
	ctgatttgga	aagcaacatt	acgagacaat	tagtgaaggc	tagagaacaa	gaggaaaaga	2160
50	atctagatgc	ctatcagaag	ttaactagtt	ggtcagattt	ctggtccttg	ttcgatttct	2220
	caaaatggct	taacatttta	aaaatgggat	ttttagtaat	agtaggaata	ataggggtta	2280
	gattacttta	cacagtatat	ggatgtatag	tgagggttag	gcagggatat	gttcctctat	2340
55	ctccacagat	ccatatccgc	ggcaatttta	aaagaaaggg	aggaataggg	ggacagactt	2400
	cagcagagag	actaattaat	ataataacaa	cacaattaga	aataacaacat	ttacaaacca	2460
60	aaattcaaaa	aattttaaat	tttagagccg	cggagatctg	ttacataact	tatggtaaat	2520
	ggcctgcctg	gctgactgcc	caatgacccc	tgcccaatga	tgtcaataat	gatgtatggt	2580
	cccatgtaat	gccaataggg	actttccatt	gatgtcaatg	ggtggagtat	ttatggtaac	2640
65	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	atgcccccta	ttgatgtcaa	2700
	tgatggtaaa	tggcctgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	2760
70	ttggcagtac	atctatgtat	tagtcattgc	tattaccatg	ggaattcact	agtggagaag	2820

	agcatgcttg	agggctgagt	gcccctcagt	gggcagagag	cacatggccc	acagtccctg	2880
5	agaagtggg	gggaggggtg	ggcaattgaa	ctggtgccta	gagaagggtg	ggcttgggta	2940
	aactgggaaa	gtgatgtggt	gtactggctc	cacctttttc	cccaggggtg	gggagaacca	3000
	tatataagt	cagtagtctc	tgtgaacatt	caagcttctg	ccttctccct	cctgtgagtt	3060
10	tgctagccac	catgcccagc	tctgtgtcct	ggggcattct	gctgctggct	ggcctgtgct	3120
	gtctgggtgc	tgtgtccctg	gctgaggacc	ctcaggggga	tgctgcccag	aaaacagaca	3180
15	cctcccacca	tgaccaggac	cacccacct	tcaacaagat	cacccccaac	ctggcagagt	3240
	ttgccttcag	cctgtacaga	cagctggccc	accagagcaa	cagcaccaac	atctttttca	3300
	gccctgtgtc	cattgccaca	gcctttgcc	tgctgagcct	gggcaccaag	gctgacaccc	3360
20	atgatgagat	cctggaaggc	ctgaacttca	acctgacaga	gatccctgag	gccagatcc	3420
	atgagggctt	ccaggaactg	ctgagaaccc	tgaaccagcc	agacagccag	ctgcagctga	3480
25	caacaggcaa	tgggctgttc	ctgtctgagg	gcctgaagct	ggtggacaag	tttctggaag	3540
	atgtgaagaa	gctgtaccac	tctgaggcct	tcacagtga	ctttggggac	acagaagagg	3600
	ccaagaaaca	gatcaatgac	tatgtggaaa	agggcaccca	gggcaagatt	gtggaccttg	3660
30	tgaaagagct	ggacagggac	actgtgtttg	cccttgtgaa	ctacatcttc	ttcaagggca	3720
	agtgggagag	gccctttgaa	gtgaaggaca	ctgaggaaga	ggacttccat	gtggaccaag	3780
35	tgaccacagt	gaagggtgcca	atgatgaaga	gactggggat	gttcaatatc	cagcactgca	3840
	agaaactgag	cagctgggtg	ctgctgatga	agtacctggg	caatgctaca	gccatattct	3900
	ttctgcctga	tgagggcaag	ctgcagcacc	tggaaaatga	gctgacccat	gacatcatca	3960
40	ccaaatttct	ggaaaatgag	gacagaagat	ctgccagcct	gcatctgccc	aagctgagca	4020
	tcacaggcac	atatgacctg	aagtctgtgc	tgggacagct	gggaatcacc	aagggtgttca	4080
45	gcaatggggc	agacctgagt	ggagtgacag	aggaagcccc	tctgaagctg	tccaaggctg	4140
	tgacaaggc	agtgtgacc	attgatgaga	agggcacaga	ggctgctggg	gccatgtttc	4200
	tggaagccat	ccccatgtcc	atccccccag	aagtgaagtt	caacaagccc	tttgtgttcc	4260
50	tgatgattga	gcagaacacc	aagagcccc	tgttcatggg	caaggttgtg	aacccccacc	4320
	agaaatgagg	gcccaatcaa	cctctggatt	acaaaatttg	tgaaagattg	actggtattc	4380
55	ttaactatgt	tgctcctttt	acgctatgtg	gatacgctgc	tttaatgcct	ttgtatcatg	4440
	ctattgcttc	ccgtatggct	ttcattttct	cctccttgta	taaatcctgg	ttgtgtctc	4500
	tttatgagga	gttgtggccc	gttgtcaggc	aacgtggcgt	ggtgtgcact	gtgtttgctg	4560
60	acgcaacccc	cactgggttg	ggcattgcca	ccacctgtca	gctcctttcc	gggactttcg	4620
	ctttccccct	ccctattgcc	acggcggaac	tcacgcccgc	ctgccttgcc	cgctgctgga	4680
65	caggggctcg	gctgttgggc	actgacaatt	ccgtggtgtt	gtcggggaaa	tcacgtcct	4740
	ttccttggt	gctcgcctgt	gttgccacct	ggattctgcg	cgggacgtcc	ttctgctacg	4800
	tcccttcggc	cctcaatcca	gcggaccttc	cttcccgcgg	cctgctgccg	gctctgcggc	4860
70	ctcttccgcg	tcttcgcctt	cgccctcaga	cgagtcggat	ctcccttttg	gccgcctccc	4920

	cgcaagcttc	gcacttttta	aaagaaaagg	gaggactgga	tgggatttat	tactccgata	4980
5	ggacgctggc	ttgtaactca	gtctcttact	aggagaccag	cttgagcctg	ggtgttcgct	5040
	ggttagccta	acctggttgg	ccaccagggg	taaggactcc	ttggcttaga	aagctaataa	5100
	acttgccctgc	attagagctc	ttacgcgtcc	cgggctcgag	atccgcatct	caattagtca	5160
10	gcaaccatag	tcccgccctt	aactccgccc	atcccgcctc	taactccgcc	cagttccgcc	5220
	cattctccgc	cccatggctg	actaattttt	tttatttatg	cagaggccga	ggccgcctcg	5280
15	gcctctgagc	tattccagaa	gtagtgagga	ggcttttttg	gaggcctagg	cttttgcaaa	5340
	aagctaactt	gtttattgca	gcttataatg	gttacaaata	aagcaatagc	atcacaaatt	5400
	tcacaaataa	agcatttttt	tcactgcatt	ctagttgtgg	tttgtccaaa	ctcatcaatg	5460
20	tatcttatca	tgtctgtccg	cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	5520
	gcggcgagcg	gtatcagctc	actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	5580
25	taacgcagga	aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	5640
	cgcgttgctg	gcgtttttcc	ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	5700
	ctcaagtcag	aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	5760
30	aagctccctc	gtgcgctctc	ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	5820
	tctcccttcg	ggaagcgtgg	cgctttctca	tagctcacgc	tgtaggtatc	tcagttcggt	5880
35	gtaggtcggt	cgctccaagc	tgggctgtgt	gcacgaacct	cccgttcagc	ccgaccgctg	5940
	cgcttatcc	ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	6000
	ggcagcagcc	actggtaaca	ggattagcag	agcgaggtat	gtaggcgggtg	ctacagagtt	6060
40	cttgaagtgg	tggcctaact	acggctacac	tagaagaaca	gtatttggtg	tctgcgctct	6120
	gctgaagcca	gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	6180
45	cgctggtagc	ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	6240
	tcaagaagat	cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	6300
	ttaagggatt	ttggtcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	6360
50	aaaatgaagt	tttaaataca	tctaaagtat	atatgagtaa	acttgggtctg	acagttagaa	6420
	aaactcatcg	agcatcaaata	gaaactgcaa	tttattcata	tcaggattat	caataccata	6480
55	tttttgaaaa	agccgtttct	gtaatgaagg	agaaaactca	ccgaggcagt	tccataggat	6540
	ggcaagatcc	tggatatcgg	ctgcgattcc	gactcgtcca	acatcaatac	aacctattaa	6600
	tttccccctg	tcaaaaataa	ggttatcaag	tgagaaatca	ccatgagtga	cgactgaatc	6660
60	cggtgagaat	ggcaacagct	tatgcatttc	tttccagact	tggtcaacag	gccagccatt	6720
	acgctcgtca	tcaaaatcac	tcgcatcaac	caaaccgtta	ttcattcgtg	attgcgcctg	6780
65	agcgagacga	aatacgcgat	cgctgttaaa	aggacaatta	caaacaggaa	tcgaatgcaa	6840
	ccggcgcagg	aacactgcc	gcgcatcaac	aatattttca	cctgaatcag	gatattcttc	6900
	taataacctg	aatgctgttt	ttccggggat	cgcagtggtg	agtaaccatg	catcatcagg	6960
70	agtacggata	aaatgcttga	tggtcggaag	aggcataaat	tccgtcagcc	agtttagtct	7020

	gaccatctca tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc	7080
	tggcgcatcg ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc	7140
5	gcgagcccat ttatacccat ataatcagc atccatgttg gaatttaatc gcggcctaga	7200
	gcaagacgtt tcccgttgaa tatggctcat aacaccctt gtattactgt ttatgtaagc	7260
10	agacagtttt attgttcatg atgatatatt tttatcttgt gcaatgtaac atcagagatt	7320
	ttgagacaca acaattggtc gacggatcc	7349
15	<210> 10 <211> 6653 <212> ДНК <213> штучна послідовність	
20	<220> <223> pGM358 плазміда	
	<400> 10	
25	ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat	60
	tggctatttg ccattgcata cgttgtatct atatacataat atgtacattt atattggctc	120
	atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat	180
30	tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa	240
	tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt	300
	tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	360
35	aactgcccac ttggcagtac atcaagtgtg tcatatgcc aagtcgcccc ctattgacgt	420
	caatgacggg aaatggcccc cctggcatta tgcccagtac atgaccttac gggactttcc	480
40	tacttggcag tacatctacg tattagtcac cgctattacc atgggtgatgc ggttttggca	540
	gtacaccaat gggcgtggat agcggtttga ctacgaggga tttccaagtc tccaccccat	600
	tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa	660
45	caactgcatg cggccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtc	720
	tatataagca gagctcgctg gcttgtaact cagtctctta ctaggagacc agcttgagcc	780
50	tgggtgttcg ctggttagcc taacctggtt ggccaccagg ggtaaggact ccttggctta	840
	gaaagctaata aaacttgcct gcattagagc ttatctgagt caagtgtcct cattgacgcc	900
	tcactctctt gaacgggaat ctcccttact gggttctctc tctgaccag gcgagagaaa	960
55	ctccagcagt ggcgcccga cagggacttg agtgagagt taggcacgta cagctgagaa	1020
	ggcgtcggac gcgaaggga cgcgggggtgc gacgcgacca agaaggagac ttggtgagta	1080
60	ggcttctcga gtgccgggaa aaagctcgag cctagttaga ggactaggag aggccgtagc	1140
	cgtaactact cttgggcaag tagggcaggc ggtgggtacg caatgggggc ggctacctca	1200
	gcactaaata ggagacaatt agaccaattt gagaaaatac gacttcgccc gaacggaaaag	1260
65	aaaaagtacc aaattaaaca tttaatatgg gcaggcaagg agatggagcg cttcggcctc	1320
	catgagaggt tgttgagac agaggagggg tgtaaaagaa tcatagaagt cctctacccc	1380
70	ctagaaccaa caggatcgga gggcttaaaa agtctgttca atcttgtgtg cgtgctatat	1440

	tgcttgacaca	aggaacagaa	agtgaaagac	acagaggaag	cagtagcaac	agtaagacaa	1500
5	cactgccatc	tagtggaana	agaaaaaagt	gcaacagaga	catctagtgg	acaaaagaaa	1560
	aatgacaagg	gaatagcagc	gccacctggt	ggcagtcaga	atcttccagc	gcaacaacaa	1620
	ggaaatgcct	gggtacatgt	acccttgtca	ccgcgcacct	taaatgcgtg	ggtaaaagca	1680
10	gtagaggaga	aaaaatttgg	agcagaaata	gtacccattt	ttttgtttca	agccctatcg	1740
	aattcccgtt	tgtgctaggg	ttcttaggct	tcttgggggc	tgctggaact	gcaatgggag	1800
	cagcggcgac	agccctgacg	gtccagtctc	agcatttgct	tgctgggata	ctgcagcagc	1860
15	agaagaatct	gctggcggct	gtggaggctc	aacagcagat	gttgaagctg	accatttggg	1920
	gtgttaaaaa	cctcaatgcc	cgcgctcacag	cccttgagaa	gtacctagag	gatcaggcac	1980
20	gactaaactc	ctgggggtgc	gcatggaaac	aagtatgtca	taccacagtg	gagtggccct	2040
	ggacaaatcg	gactccggat	tggcaaaaata	tgacttggtt	ggagtgggaa	agacaaatag	2100
25	ctgatttggg	aagcaacatt	acgagacaat	tagtgaaggc	tagagaacaa	gaggaaaaga	2160
	atctagatgc	ctatcagaag	ttaactagtt	ggtcagattt	ctggtcttgg	ttcgatttct	2220
	caaaatggct	taacatttta	aaaatgggat	ttttagtaat	agtaggaata	ataggggtta	2280
30	gattacttta	cacagtatat	ggatgtatag	tgagggttag	gcagggatat	gttcctctat	2340
	ctccacagat	ccatatccgc	ggcaatttta	aaagaaaggg	aggaataggg	ggacagactt	2400
35	cagcagagag	actaattaat	ataataacaa	cacaattaga	aataacaacat	ttacaaacca	2460
	aaattcaaaa	aattttaaat	tttagagccg	cggagatctg	ttacataact	tatggtaaata	2520
	ggcctgcctg	gctgactgcc	caatgacccc	tgcccaatga	tgtcaataat	gatgtatggt	2580
40	cccatgtaat	gccaataggg	actttccatt	gatgtcaatg	ggtggagtat	ttatggtaac	2640
	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	atgcccccta	ttgatgtcaa	2700
45	tgatggtaaa	tggcctgcct	ggcattatgc	ccagtacatg	accttatggg	actttcctac	2760
	ttggcagtac	atctatgtat	tagtcattgc	tattaccatg	ggaattcact	agtggagaag	2820
	agcatgcttg	agggctgagt	gcccctcagt	gggcagagag	cacatggccc	acagtccctg	2880
50	agaagtggg	gggaggggtg	ggcaattgaa	ctggtgccta	gagaagggtg	ggcttgggta	2940
	aactgggaaa	gtgatgtggt	gtactggctc	cacctttttc	cccaggggtg	gggagaacca	3000
55	tatataagt	cagtagtctc	tgtgaacatt	caagcttctg	ccttctccct	cctgtgagtt	3060
	tgctagccac	catgggagtg	aagggtgctgt	ttgccctgat	ctgcattgct	gtggctgagg	3120
	ccaagcccac	agagaacaat	gaggacttca	acattgtggc	tgtggccagc	aactttgcc	3180
60	ccacagacct	ggatgctgac	aggggcaagc	tgcctggcaa	gaagctgccc	ctggaagtcc	3240
	tgaagagat	ggaagccaat	gccaggaagg	ctggctgcac	aagaggctgt	ctgatctgcc	3300
65	tgagccacat	caagtgcacc	cccaagatga	agaagttcat	ccctggcagg	tgccacacct	3360
	atgaagggga	caaagagtct	gcccaggggg	gaattggaga	ggccattgtg	gacatccctg	3420
	agatccctgg	cttcaaggac	ctggaaccca	tggaaacagtt	cattgcccag	gtggacctgt	3480
70	gtgtggactg	cactacaggc	tgtctcaagg	gcctggccaa	tgtgcagtg	tctgacctgc	3540

	tgaagaagtg	gctgccccag	agatgtgcca	cctttgccag	caagatccag	ggccaggtgg	3600
5	acaagatcaa	gggagctggg	ggagattgat	gagggcccaa	tcaacctctg	gattacaaaa	3660
	tttgtgaaag	attgactggg	attcttaact	atgttgctcc	ttttacgcta	tgtggatacg	3720
	ctgctttaat	gcctttgtat	catgctattg	cttcccgtat	ggctttcatt	ttctcctcct	3780
10	tgtataaatc	ctggttgctg	tctctttatg	aggagttgtg	gcccgttgtc	aggcaacgtg	3840
	gcgtgggtgtg	cactgtgttt	gctgacgcaa	ccccactgg	ttggggcatt	gccaccacct	3900
15	gtcagctcct	ttccgggact	ttcgctttcc	ccctccctat	tgccacggcg	gaactcatcg	3960
	ccgcctgcct	tgcccgtgc	tggacagggg	ctcggtgtt	gggactgac	aattccgtgg	4020
	tgttgctcggg	gaaatcatcg	tcctttcctt	ggctgctcgc	ctgtgttgcc	acctggattc	4080
20	tgcgcgggac	gtccttctgc	tacgtccctt	cggccctcaa	tccagcggac	cttccttccc	4140
	gcggcctgct	gccggctctg	cggcctcttc	cgctcttcg	ccttcgccct	cagacgagtc	4200
25	ggatctccct	ttggggccgc	tccccgcaag	cttcgcactt	tttaaaagaa	aaggaggagc	4260
	tggatgggat	ttattactcc	gataggacgc	tggcttgtaa	ctcagtctct	tactaggaga	4320
	ccagcttgag	cctgggtgtt	cgctgggttag	cctaacctgg	ttggccacca	ggggtaagga	4380
30	ctccttggt	tagaaagcta	ataaacttgc	ctgcattaga	gctcttacgc	gtcccgggct	4440
	cgagatccgc	atctcaatta	gtcagcaacc	atagtccgc	ccctaactcc	gcccattccc	4500
35	cccctaactc	cgcccagttc	cgcccattct	ccgccccatg	gctgactaat	tttttttatt	4560
	tatgcagagg	ccgaggccgc	ctcggcctct	gagctattcc	agaagtagtg	aggaggcttt	4620
	tttggaggcc	taggcttttg	caaaaagcta	acttgtttat	tgcagcttat	aatggttaca	4680
40	aataaagcaa	tagcatcaca	aatttcacaa	ataaagcatt	tttttactg	cattctagtt	4740
	gtggtttgtc	caaactcatc	aatgtatctt	atcatgtctg	tccgcttcct	cgctcactga	4800
45	ctcgtgcgc	tcggtcgttc	ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	4860
	acggttatcc	acagaatcag	gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	4920
	aaaggccagg	aaccgtaaaa	aggccgcgtt	gctggcggtt	ttccataggc	tccgcccccc	4980
50	tgacgagcat	cacaaaaatc	gacgctcaag	tcagaggtgg	cgaaaccgga	caggactata	5040
	aagataccag	gcgtttcccc	ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	5100
55	gcttaccgga	tacctgtccg	cctttctccc	ttcggaagc	gtggcgcttt	ctcatagctc	5160
	acgctgtagg	tatctcagtt	cggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	5220
	acccccggtt	cagcccgacc	gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	5280
60	ggtaagacac	gacttatcgc	cactggcagc	agccactggg	aacaggatta	gcagagcgag	5340
	gtatgtaggc	ggtgctacag	agttcttgaa	gtggtggcct	aactacggct	acactagaag	5400
65	aacagtattt	ggtatctgcg	ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	5460
	ctcttgatcc	ggcaaaaaa	ccaccgctgg	tagcggtggt	ttttttgttt	gcaagcagca	5520
	gattacgcgc	agaaaaaaag	gatctcaaga	agatcctttg	atcttttcta	cggggtctga	5580
70	cgctcagtgg	aacgaaaact	cacgttaagg	gattttgggtc	atgagattat	caaaaaggat	5640

	cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa gtatatatga	5700
5	gtaaacttgg tctgacagtt agaaaaactc atcgagcatc aaatgaaact gcaattttatt	5760
	catatcagga ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa	5820
	ctcaccgagg cagttccata ggatggcaag atcctgggtat cggctctgca ttccgactcg	5880
10	tccaacatca atacaacctt ttaattttccc ctcgtcaaaa ataagggttat caagtgagaa	5940
	atcaccatga gtgacgactg aatccggtga gaatggcaac agcttatgca tttctttcca	6000
15	gacttgttca acaggccagc cattacgctc gtcatacaaaa tcaactcgcat caaccaaac	6060
	gttattcatt cgtgattgca cctgagcgag acgaaatacg cgatcgctgt taaaaggaca	6120
	attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt	6180
20	ttcacctgaa tcaggatatt cttctaatac ctggaatgct gtttttccgg ggatcgagct	6240
	ggtagagtaac catgcatcat caggagtacg gataaaatgc ttgatgggtcg gaagaggcat	6300
25	aaattccgctc agccagttta gtctgacctc ctcactctgta acatcattgg caacgctacc	6360
	tttgccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt	6420
	cgcacctgat tgcccagacat tatcgcgagc ccattttatac ccataataat cagcatccat	6480
30	gttggaattt aatcgcgggc tagagcaaga cgtttcccgt tgaatatggc tcataacacc	6540
	ccttgattta ctgtttatgt aagcagacag ttttattgtt catgatgata ttttttatc	6600
35	ttgtgcaatg taacatcaga gattttgaga cacaacaatt ggtcgacgga tcc	6653
	<210> 11	
	<211> 10812	
	<212> ДНК	
40	<213> штучна послідовність	
	<220>	
	<223> pGM411 плазміда	
45	<400> 11	
	ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat	60
	tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc	120
50	atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat	180
	tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa	240
	tggccgcct ggctgaccgc ccaacgacct ccgcccattg acgtcaataa tgacgtatgt	300
55	tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	360
	aactgcccac ttggcagtac atcaagtgtg tcatatgcc aagtcgcccc ctattgacgt	420
60	caatgacggt aaatggcccc cctggcatta tgcccagtac atgaccttac gggactttcc	480
	tacttggcag tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca	540
	gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat	600
65	tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa	660
	caactgcat cgccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtc	720
70	tatataagca gagctcgctg gcttgtaact cagtctctta ctaggagacc agcttgagcc	780

	tgggtgttcg	ctggttagcc	taacctggtt	ggccaccagg	ggtaaggact	ccttggctta	840
5	gaaagcta	aaacttgcct	gcattagagc	ttatctgagt	caagtgtcct	cattgacgcc	900
	tcactctctt	gaacgggaat	cttccttact	gggttctctc	tctgaccag	gcgagagaaa	960
	ctccagcagt	ggcgcccgaa	cagggacttg	agtgagagt	taggcacgta	cagctgagaa	1020
10	ggcgtcggac	gcgaaggaag	cgcggggtgc	gacgcgacca	agaaggagac	ttggtgagta	1080
	ggcttctcga	gtgccgggaa	aaagctcgag	cctagttaga	ggactaggag	aggccgtagc	1140
15	cgtaactact	ctgggcaagt	agggcaggcg	gtgggtacgc	aatgggggag	gctacctcag	1200
	cactaaatag	gagacaatta	gaccaatttg	agaaaatacg	acttcgcccc	aacggaaaaga	1260
	aaaagtacca	aattaaacat	ttaatattgg	caggcaagga	gatggagcgc	ttcggcctcc	1320
20	atgagagggt	gttgagagaca	gaggaggggt	gtaaaagaat	catagaagtc	ctctaccccc	1380
	tagaaccaac	aggatcggag	ggcttaaaaa	gtctgttcaa	tcttgtgtgc	gtgctatatt	1440
25	gcttgacaaa	ggaacagaaa	gtgaaagaca	cagaggaagc	agtagcaaca	gtaagacaac	1500
	actgccatct	agtggaaaaa	gaaaaaagt	caacagagac	atctagtggg	caaaagaaaa	1560
	atgacaaggg	aatagcagcg	ccacctggtg	gcagtcagaa	ttttccagcg	caacaacaag	1620
30	gaaatgcctg	ggtacatgta	cccttgtcac	cgcgcacctt	aaatgcgtgg	gtaaaagcag	1680
	tagaggagaa	aaaatttgga	gcagaaatag	taccatgtt	tcaagcccta	tcgaattccc	1740
35	gtttgtgcta	gggttcttag	gcttcttggg	ggctgctgga	actgcaatgg	gagcagcggc	1800
	gacagccctg	acggtccagt	ctcagcattt	gcttgcctgg	atactgcagc	agcagaagaa	1860
	tctgctggcg	gctgtggagg	ctcaacagca	gatgttgaag	ctgaccattt	ggggtgttaa	1920
40	aaacctcaat	gcccgcgtca	cagcccttga	gaagtaccta	gaggatcagg	cacgactaaa	1980
	ctcctggggg	tgcgcatgga	aacaagtatg	tcataccaca	gtggagtggc	cctggacaaa	2040
45	tcggactccg	gattggcaaa	atatgacttg	gttggagtgg	gaaagacaaa	tagctgattt	2100
	ggaaagcaac	attacgagac	aattagttaa	ggctagagaa	caagaggaaa	agaatctaga	2160
	tgcctatcag	aagttaacta	gttgggtcaga	tttctgggtc	tggttcgatt	tctcaaaatg	2220
50	gcttaacatt	ttaaaaatgg	gatttttagt	aatagtagga	ataatagggt	taagattact	2280
	ttacacagta	tatggatgta	tagtgagggt	taggcaggga	tatgttcctc	tatctccaca	2340
55	gatccatatt	cgcggaatt	ttaaaagaaa	gggaggaata	gggggacaga	cttcagcaga	2400
	gagactaatt	aataataata	caacacaatt	agaaatacaa	catttacaaa	ccaaaattca	2460
	aaaaatttta	aatttttagag	ccgcggagat	ctcaatattg	gccattagcc	atattattca	2520
60	ttggttatat	agcataaatc	aatattggct	attggccatt	gcatacgttg	tatctatatc	2580
	ataatatgta	catttatatt	ggctcatgtc	caatatgacc	gccatgttgg	cattgattat	2640
65	tgactagtta	ttaatagtaa	tcaattacgg	ggtcattagt	tcatagcccc	tatatggagt	2700
	tccgcgttac	ataacttacg	gtaaatggcc	cgcttggtcg	accgcccac	gacccccgcc	2760
	cattgacgtc	aataatgacg	tatgttccca	tagtaacgcc	aatagggact	ttccattgac	2820
70	gtcaatgggt	ggagtattta	cggtaaactg	cccacttggc	agtacatcaa	gtgtatcata	2880

	tgccaagtcc	gccccctatt	gacgtcaatg	acggtaaagt	gcccgcctgg	cattatgccc	2940
5	agtacatgac	cttacgggac	tttctacttt	ggcagtacat	ctacgtatta	gtcatcgcta	3000
	ttaccatggt	gatgcggttt	tggcagtaca	ccaatgggcg	tggatagcgg	tttgactcac	3060
	ggggatttcc	aagtctccac	cccattgacg	tcaatgggag	tttgtttttg	caccaaatac	3120
10	aacgggactt	tccaaaatgt	cgtaataacc	ccgccccgtt	gacgcaaagt	ggcggtaggc	3180
	gtgtacgggt	ggaggtctat	ataagcagag	ctcgtttagt	gaaccgtcag	atcactagaa	3240
15	gctttattgc	ggtagtttat	cacagttaaa	ttgctaacgc	agtcagtgtc	tctgacacaa	3300
	cagtctcgaa	cttaagctgc	agaagttggt	cgtgaggcac	tgggcaggct	agccaccaat	3360
	gcagattgag	ctgagcacct	gcttcttctt	gtgcctgtct	aggttctgtc	tctctgccac	3420
20	caggagatac	tacctggggg	ctgtggagct	gagctgggac	tacatgcagt	ctgacctggg	3480
	ggagctgcct	gtggatgcca	ggttcccccc	cagagtgtcc	aagagcttcc	ccttcaacac	3540
25	ctctgtggtg	tacaagaaga	ccctgtttgt	ggagttcact	gaccacctgt	tcaacattgc	3600
	caagcccagg	ccccctgga	tgggcctgct	gggccccacc	atccaggctg	aggtgtatga	3660
	cactgtggtg	atcacctga	agaacatggc	cagccaccct	gtgagcctgc	atgctgtggg	3720
30	ggtgagctac	tgggaaggcct	ctgagggggc	tgagtatgat	gaccagacca	gccagaggga	3780
	gaaggaggat	gacaaggtgt	tccctggggg	cagccacacc	tatgtgtggc	aggtgtgaa	3840
35	ggagaatggc	cccatggcct	ctgacccctt	gtgcctgacc	tacagctacc	tgagccatgt	3900
	ggacctgggt	aaggacctga	actctggcct	gattggggcc	ctgctgggtg	gcagggaggg	3960
	cagcctggcc	aaggagaaga	cccagaccct	gcacaagttc	atcctgctgt	ttgctgtgtt	4020
40	tgatgagggc	aagagctggc	actctgaaac	caagaacagc	ctgatgcagg	acagggatgc	4080
	tgcctctgcc	agggcctggc	ccaagatgca	cactgtgaat	ggctatgtga	acaggagcct	4140
45	gcctggcctg	attggctgcc	acaggaagtc	tgtgtactgg	catgtgattg	gcatgggcac	4200
	cacccttgag	gtgcacagca	tcttcttgga	gggccacacc	ttcctgggtc	ggaaccacag	4260
	gcaggccagc	ctggagatca	gccccatcac	cttcttgact	gcccagaccc	tgctgatgga	4320
50	cctggggccag	ttcctgctgt	tctgccacat	cagcagccac	cagcatgatg	gcatggaggc	4380
	ctatgtgaag	gtggacagct	gccctgagga	gccccagctg	aggatgaaga	acaatgagga	4440
55	ggctgaggac	tatgatgatg	acctgactga	ctctgagatg	gatgtgggtg	ggtttgatga	4500
	tgacaacagc	cccagcttca	tccagatcag	gtctgtggcc	aagaagcacc	ccaagacctg	4560
	ggtgcactac	attgctgctg	aggaggagga	ctgggactat	gccccctgg	tgctggcccc	4620
60	tgatgacagg	agctacaaga	gccagtacct	gaacaatggc	ccccagagga	ttggcaggaa	4680
	gtacaagaag	gtcaggttca	tggcctacac	tgatgaaacc	ttcaagacca	gggaggccat	4740
65	ccagcatgag	tctggcatcc	tgggccccct	gctgtatggg	gaggtggggg	acaccctgct	4800
	gatcatcttc	aagaaccagg	ccagcaggcc	ctacaacatc	tacccccatg	gcatcactga	4860
	tgtgaggccc	ctgtacagca	ggaggctgcc	caagggggtg	aagcacctga	aggacttccc	4920
70	catcctgcct	ggggagatct	tcaagtacaa	gtggactgtg	actgtggagg	atggccccac	4980

	caagtctgac	cccaggtgcc	tgaccagata	ctacagcagc	tttgtgaaca	tggagagggga	5040
5	cctggcctct	ggcctgattg	gccccctgct	gatctgctac	aaggagtctg	tggaccagag	5100
	gggcaaccag	atcatgtctg	acaagaggaa	tgtgatcctg	ttctctgtgt	ttgatgagaa	5160
	caggagctgg	tacctgactg	agaacatcca	gaggttcctg	cccaaccctg	ctggggtgca	5220
10	gctggaggac	cctgagttcc	aggccagcaa	catcatgcac	agcatcaatg	gctatgtggt	5280
	tgacagcctg	cagctgtctg	tgtgcctgca	tgagggtggc	tactggtaca	tcctgagcat	5340
	tggggccag	actgacttcc	tgtctgtggt	cttctctggc	tacaccttca	agcacaagat	5400
15	ggtgtatgag	gacaccctga	ccctgttccc	cttctctggg	gagactgtgt	tcattgagcat	5460
	ggagaaccct	ggcctgtgga	ttctgggctg	ccacaactct	gacttcagga	acaggggcat	5520
20	gactgccctg	ctgaaagtct	ccagctgtga	caagaacact	ggggactact	atgaggacag	5580
	ctatgaggac	atctctgcct	acctgctgag	caagaacaat	gccattgagc	ccaggagctt	5640
	cagccagaat	gccactaatg	tgtctaacia	cagcaacacc	agcaatgaca	gcaatgtgtc	5700
25	tccccagtg	ctgaagaggc	accagagggg	gatcaccagg	accaccctgc	agtctgacca	5760
	ggaggagatt	gactatgatg	acaccatctc	tgtggagatg	aagaaggagg	actttgacat	5820
30	ctacgacgag	gacgagaacc	agagccccag	gagcttccag	aagaagacca	ggcactactt	5880
	cattgctgct	gtggagaggc	tgtgggacta	tggcatgagc	agcagcccc	atgtgctgag	5940
	gaacagggcc	cagtctggct	ctgtgcccc	gttcaagaag	gtggtgttcc	aggagttcac	6000
35	tgatggcagc	ttcaccagc	ccctgtacag	aggggagctg	aatgagcacc	tgggcctgct	6060
	ggggccctac	atcagggtg	aggtggagga	caacatcatg	gtgaccttca	ggaaccaggc	6120
40	cagcaggccc	tacagcttct	acagcagcct	gatcagctat	gaggaggacc	agaggcaggg	6180
	ggctgagccc	aggaagaact	ttgtgaagcc	caatgaaacc	aagacctact	tctggaaggt	6240
	gcagcaccac	atggccccca	ccaaggatga	gtttgactgc	aaggcctggg	cctacttctc	6300
45	tgatgtggac	ctggagaagg	atgtgcactc	tggcctgatt	ggccccctgc	tgggtgtgcca	6360
	caccaacacc	ctgaaccctg	cccatggcag	gcagggtgact	gtgcaggagt	ttgccctgtt	6420
50	cttcaccatc	tttgatgaaa	ccaagagctg	gtactttact	gagaacatgg	agaggaactg	6480
	cagggccccc	tgcaacatcc	agatggagga	ccccaccttc	aaggagaact	acaggttcca	6540
	tgccatcaat	ggctacatca	tggacaccct	gcctggcctg	gtgatggccc	aggaccagag	6600
55	gatcaggtgg	tacctgctga	gcatgggcag	caatgagaac	atccacagca	tccacttctc	6660
	tggccatgtg	ttcactgtga	ggaagaagga	ggagtacaag	atggccctgt	acaacctgta	6720
60	ccctggggtg	tttgagactg	tggagatgct	gccagcaag	gctggcatct	ggaggggtgga	6780
	gtgcctgatt	ggggagcacc	tgcatgctgg	catgagcacc	ctgttcctgg	tgtacagcaa	6840
	caagtgccag	acccccctgg	gcatggcctc	tggccacatc	agggacttcc	agatcactgc	6900
65	ctctggccag	tatggccagt	gggcccccaa	gctggccagg	ctgcactact	ctggcagcat	6960
	caatgcctgg	agcaccaagg	agcccttcag	ctggatcaag	gtggacctgc	tggcccccat	7020
70	gatcatccat	ggcatcaaga	cccagggggc	caggcagaag	ttcagcagcc	tgtacatcag	7080

	ccagttcatc	atcatgtaca	gcctggatgg	caagaagtgg	cagacctaca	ggggcaacag	7140
5	cactggcacc	ctgatggtgt	tctttggcaa	tgtggacagc	tctggcatca	agcacaacat	7200
	cttcaacccc	cccatcattg	ccagatacat	caggctgcac	cccaccact	acagcatcag	7260
	gagcaccctg	aggatggagc	tgatgggctg	tgacctgaac	agctgcagca	tgcccctggg	7320
10	catggagagc	aaggccatct	ctgatgccca	gatcactgcc	agcagctact	tcaccaacat	7380
	gtttgccacc	tggagcccca	gcaaggccag	gctgcacctg	cagggcagga	gcaatgcctg	7440
	gaggccccag	gtcaacaacc	ccaaggagtg	gctgcagggtg	gacttccaga	agaccatgaa	7500
15	ggtgactggg	gtgaccaccc	aggggggtgaa	gagcctgctg	accagcatgt	atgtgaagga	7560
	gttcctgatc	agcagcagcc	aggatggcca	ccagtggacc	ctgttcttcc	agaatggcaa	7620
20	ggtgaagggtg	ttccagggca	accaggacag	cttcaccctt	gtggtgaaca	gcctggaccc	7680
	ccccctgctg	accagatacc	tgaggattca	ccccagagc	tgggtgcacc	agattgccct	7740
	gaggatggag	gtgctgggct	gtgaggccca	ggacctgtac	tgagcggccg	cgggccaat	7800
25	caacctctgg	attacaaaat	ttgtgaaaga	ttgactggta	ttcttaacta	tgttgctcct	7860
	tttacgctat	gtggatacgc	tgctttaatg	cctttgtatc	atgctattgc	ttcccgtatg	7920
30	gctttcattt	tctcctcctt	gtataaatcc	tggttgctgt	ctctttatga	ggagttgtgg	7980
	cccgtttgtca	ggcaacgtgg	cgtggtgtgc	actgtgtttg	ctgacgcaac	ccccactggt	8040
	tggggcattg	ccaccacctg	tcagctcctt	tccgggactt	tcgctttccc	cctccctatt	8100
35	gccacggcgg	aactcatcgc	cgctgcctt	gcccgcgtgct	ggacaggggc	tcggctgttg	8160
	ggcactgaca	attccgtggg	gttgtcgggg	aaatcatcgt	cctttccttg	gctgctcgcc	8220
40	tgtgttgcca	cctggattct	gcgcgggacg	tccttctgct	acgtcccttc	ggccctcaat	8280
	ccagcggacc	ttccttcccc	cggcctgctg	cgggctctgc	ggcctcttcc	gcgtcttcgc	8340
	cttcgcccctc	agacgagtcg	gatctccctt	tgggcccgcct	ccccgcaagc	ttcgactttt	8400
45	ttaaaagaaa	agggaggact	ggatgggatt	tattactccg	ataggacgct	ggcttgtaac	8460
	tcagtctctt	actaggagac	cagcttgagc	ctgggtgttc	gctgggttagc	ctaacctggt	8520
50	tggccaccag	gggtaaggac	tccttggctt	agaaagctaa	taaacttgcc	tgcattagag	8580
	ctcttacgcg	tcccgggctc	gagatccgca	tctcaattag	tcagcaacca	tagtcccgcc	8640
	cctaactccg	cccatcccgc	ccctaactcc	gcccagttcc	gcccattctc	cgccccatgg	8700
55	ctgactaatt	ttttttatth	atgcagaggc	cgaggccgcc	tcggcctctg	agctattcca	8760
	gaagtagtga	ggaggctttt	ttggaggcct	aggcttttgc	aaaaagctaa	cttgtttatt	8820
60	gcagcttata	atggttacaa	ataaagcaat	agcatcacia	atttcacaaa	taaagcattt	8880
	ttttcactgc	attctagttg	tggtttgtcc	aaactcatca	atgtatctta	tcatgtctgt	8940
	ccgcttcctc	gctcactgac	tcgctgcgct	cggctcgttcg	gctgcggcga	gcggtatcag	9000
65	ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	9060
	tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	9120
70	tccataggct	ccgccccctt	gacgagcatc	acaaaaatcg	acgctcaagt	cagagggtggc	9180

	gaaacccgac	aggactataa	agataaccagg	cgtttccccc	tggaagctcc	ctcgtgcgct	9240
5	ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	9300
	tggcgctttc	tcatagctca	cgctgtaggt	atctcagttc	ggtgtaggtc	gttcgctcca	9360
	agctgggctg	tgtgcacgaa	cccccgttc	agcccgaccg	ctgcgcctta	tccggtaact	9420
10	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	9480
	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	9540
	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	9600
15	tcggaaaaag	agttggtagc	tcttgatccg	gcaaacaac	caccgctggt	agcgggtggtt	9660
	tttttgtttg	caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gatcctttga	9720
20	tcttttctac	ggggtctgac	gctcagtggg	acgaaaactc	acgttaaggg	attttgggtca	9780
	tgagattatc	aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	agttttaaat	9840
	caatctaaag	tatatatgag	taaacttggt	ctgacagtta	gaaaaactca	tcgagcatca	9900
25	aatgaaactg	caattttattc	atatcaggat	tatcaatacc	atatttttga	aaaagccggt	9960
	tctgtaatga	aggagaaaaac	tcaccgaggc	agttccatag	gatggcaaga	tcctggtatc	10020
30	ggtctgcat	tccgactcgt	ccaacatcaa	tacaacctat	taatttcccc	tcgtcaaaaa	10080
	taaggttatc	aagtgagaaa	tcaccatgag	tgacgactga	atccggtgag	aatggcaaca	10140
	gcttatgcat	ttctttccag	acttggtcaa	caggccagcc	attacgctcg	tcatcaaaat	10200
35	cactcgcac	aaccaaaccg	ttattcattc	gtgattgcgc	ctgagcgaga	cgaaatacgc	10260
	gatcgctggt	aaaaggacaa	ttacaaacag	gaatcgaatg	caaccggcgc	aggaacactg	10320
40	ccagcgcac	aacaatattt	tcacctgaat	caggatattc	ttctaatacc	tggaatgctg	10380
	tttttccggg	gatcgcagtg	gtgagtaacc	atgcatcatc	aggagtacgg	ataaaatgct	10440
	tgatggctcg	aagaggcata	aattccgtca	gccagtttag	tctgaccatc	tcatctgtaa	10500
45	catcattggc	aacgctacct	ttgccatggt	tcagaaacaa	ctctggcgca	tcgggcttcc	10560
	catacaatcg	atagattgtc	gcacctgatt	gcccgcacatt	atcgcgagcc	catttatacc	10620
50	catataaatc	agcatccatg	ttggaattta	atcgcggcct	agagcaagac	gtttcccggt	10680
	gaatatggct	cataacaccc	cttgatttac	tgtttatgta	agcagacagt	tttattgttc	10740
	atgatgatat	atttttatct	tgtgcaatgt	aacatcagag	attttgagac	acaacaattg	10800
55	gtcgacggat	cc					10812
60	<210>	12					
	<211>	10519					
	<212>	ДНК					
	<213>	штучна послідовність					
65	<220>						
	<223>	pGM413 плазміда					
	<400>	12					
	ggtacctcaa	tattggccat	tagccatatt	attcattggg	tatatagcat	aatcaatat	60
70	tggtatttgg	ccattgcata	cgttgtatct	atatcataat	atgtacattt	atattggctc	120

	atgtccaata	tgaccgccat	gttggcattg	attattgact	agttattaat	agtaatcaat	180
5	tacgggggtca	ttagttcata	gcccataatat	ggagttccgc	gttacataac	ttacggtaaa	240
	tggcccgct	ggctgaccgc	ccaacgaccc	ccgcccattg	acgtcaataa	tgacgtatgt	300
	tcccatagta	acgccaatag	ggactttcca	ttgacgtcaa	tgggtggagt	atttacggta	360
10	aactgcccac	ttggcagtac	atcaagtgtg	tcatatgcc	agtccgcccc	ctattgacgt	420
	caatgacggt	aaatggcccc	cctggcatta	tgcccagtac	atgaccttac	gggactttcc	480
15	tacttggcag	tacatctacg	tattagtc	cgctattacc	atggtgatgc	ggttttggca	540
	gtacaccaat	gggctgggat	agcggtttga	ctcacgggga	ttccaagtc	tccaccccat	600
	tgacgtcaat	gggagtttgt	tttggcacca	aaatcaacgg	gactttccaa	aatgtcgtaa	660
20	caactgcat	cggccgcccc	gttgacgcaa	atgggcggta	ggcgtgtacg	gtgggaggtc	720
	tatataagca	gagctcgctg	gcttgtaact	cagtctctta	ctaggagacc	agcttgagcc	780
25	tgggtgttcg	ctggtagcc	taacctgggt	ggccaccagg	ggtaaggact	ccttggtcta	840
	gaaagcta	aaacttgcct	gcattagagc	ttatctgagt	caagtgtcct	cattgacgcc	900
	tcactctctt	gaacgggaat	cttccttact	gggttctctc	tctgaccag	gcgagagaaa	960
30	ctccagcagt	ggcgcccgaa	cagggacttg	agttagagtg	taggcacgta	cagctgagaa	1020
	ggcgtcggac	gcgaaggaag	cgcggggtgc	gacgcgacca	agaaggagac	ttggtgagta	1080
35	ggcttctcga	gtgccgggaa	aaagctcgag	cctagttaga	ggactaggag	agggcgtagc	1140
	cgtaactact	ctgggcaagt	agggcaggcg	gtgggtacgc	aatgggggcg	gctacctcag	1200
	cactaaatag	gagacaatta	gaccaatttg	agaaaatacg	acttcgcccc	aacggaaaga	1260
40	aaaagtacca	aattaaacat	ttaatatggg	caggcaagga	gatggagcgc	ttcggcctcc	1320
	atgagagggt	gttgagagaca	gaggaggggt	gtaaaagaat	catagaagtc	ctctaccccc	1380
45	tagaaccaac	aggatcggag	ggcttaaaaa	gtctgttcaa	tcttgtgtgc	gtgctatatt	1440
	gcttgacaaa	ggaacagaaa	gtgaaagaca	cagaggaagc	agtagcaaca	gtaagacaac	1500
	actgccatct	agtggaaaaa	gaaaaaagtg	caacagagac	atctagtggg	caaaagaaaa	1560
50	atgacaaggg	aatagcagcg	ccacctgggtg	gcagtcagaa	ttttccagcg	caacaacaag	1620
	gaaatgcctg	ggtacatgta	cccttgtcac	cgcgcacctt	aaatgcgtgg	gtaaaagcag	1680
55	tagaggagaa	aaaatttgga	gcagaaatag	tacccatgtt	tcaagcccta	tcgaattccc	1740
	gtttgtgcta	gggttcttag	gcttcttggg	ggctgtgga	actgcaatgg	gagcagcggc	1800
	gacagccctg	acggtccagt	ctcagcattt	gcttgtctgg	atactgcagc	agcagaagaa	1860
60	tctgtctggcg	gctgtggagg	ctcaacagca	gatgttgaag	ctgaccattt	ggggtgttaa	1920
	aaacctcaat	gcccgcgtca	cagcccttga	gaagtaccta	gaggatcagg	cacgactaaa	1980
65	ctcctggggg	tgcgcatgga	aacaagtatg	tcataaccaca	gtggagtggc	cctggacaaa	2040
	tcggactccg	gattggcaaa	atatgacttg	gttgagtggtg	gaaagacaaa	tagctgattt	2100
	ggaaagcaac	attacgagac	aattagtgaa	ggctagagaa	caagaggaaa	agaatctaga	2160
70	tgccatcag	aagttaacta	gttgggtcaga	tttctgggtct	tggttcgatt	tctcaaaatg	2220

	gcttaacatt	ttaaaaatgg	gatttttagt	aatagtagga	ataatagggg	taagattact	2280
	ttacacagta	tatggatgta	tagtgagggg	taggcaggga	tatgttcctc	tatctccaca	2340
5	gatccatatt	cgcggaatt	ttaaaagaaa	gggaggaata	gggggacaga	cttcagcaga	2400
	gagactaatt	aatataataa	caacacaatt	agaaatacaa	catttacaaa	ccaaaattca	2460
10	aaaaatttta	aatttttagag	ccgcggagat	ctgttacata	acttatggta	aatggcctgc	2520
	ctggctgact	gcccattgac	ccctgcccaa	tgatgtcaat	aatgatgtat	gttcccatgt	2580
	aatgccaata	gggactttcc	attgatgtca	atgggtggag	tatttatggg	aactgcccac	2640
15	ttggcagtag	atcaagtgtg	tcatatgcca	agtatgcccc	ctattgatgt	caatgatggg	2700
	aaatggcctg	cctggcatta	tgcccagtag	atgaccttat	gggactttcc	tacttggcag	2760
20	tacatctatg	tattagtcat	tgctattacc	atgggaattc	actagtggag	aagagcatgc	2820
	ttgagggctg	agtggccctc	agtgggcaga	gagcacatgg	cccacagtcc	ctgagaagtt	2880
	ggggggaggg	gtgggcaatt	gaactggtgc	ctagagaagg	tggggcttgg	gtaaactggg	2940
25	aaagtgatgt	ggtgtactgg	ctccaccttt	ttccccaggg	tgggggagaa	ccatatataa	3000
	gtgcagtagt	ctctgtgaac	attcaagctt	ctgccttctc	cctcctgtga	gtttgctagc	3060
30	caccaatgca	gattgagctg	agcacctgct	tcttcctgtg	cctgctgagg	ttctgcttct	3120
	ctgccaccag	gagatactac	ctgggggctg	tggagctgag	ctgggactac	atgcagtctg	3180
	acctggggga	gctgcctgtg	gatgccagggt	tcccccccag	agtggccaag	agcttccccct	3240
35	tcaacacctc	tgtggtgtac	aagaagacct	tgtttgtgga	gttcaactgac	cacctgttca	3300
	acattgccaa	gcccaggccc	ccctggatgg	gcctgctggg	ccccaccatc	caggctgagg	3360
40	tgtatgacac	tgtggtgatc	accctgaaga	acatggccag	ccaccctgtg	agcctgcatg	3420
	ctgtgggggt	gagctactgg	aaggcctctg	agggggctga	gtatgatgac	cagaccagcc	3480
	agagggagaa	ggaggatgac	aagggtgttc	ctgggggcag	ccacacctat	gtgtggcagg	3540
45	tgctgaagga	gaatggcccc	atggcctctg	acccctgtg	cctgacctac	agctacctga	3600
	gccatgtgga	cctggtgaag	gacctgaact	ctggcctgat	tggggccctg	ctggtgtgca	3660
50	gggagggcag	cctggccaag	gagaagacct	agaccctgca	caagttcatc	ctgctgtttg	3720
	ctgtgtttga	tgagggaag	agctggcact	ctgaaaccaa	gaacagcctg	atgcaggaca	3780
	gggatgctgc	ctctgccagg	gcctggccca	agatgcacac	tgtgaatggc	tatgtgaaca	3840
55	ggagcctgcc	tggcctgatt	ggctgccaca	ggaagtctgt	gtactggcat	gtgattggca	3900
	tgggcaccac	ccctgagggt	cacagcatct	tcctggaggg	ccacaccttc	ctggtcagga	3960
60	accacaggca	ggccagcctg	gagatcagcc	ccatcacctt	cctgactgcc	cagaccctgc	4020
	tgatggacct	gggccagttc	ctgctgttct	gccacatcag	cagccaccag	catgatggca	4080
	tggaggccta	tgtgaagggt	gacagctgcc	ctgaggagcc	ccagctgagg	atgaagaaca	4140
65	atgaggaggc	tgaggactat	gatgatgacc	tgactgactc	tgagatggat	gtggtgaggt	4200
	ttgatgatga	caacagcccc	agcttcatcc	agatcagggtc	tgtggccaag	aagcacccca	4260
70	agacctgggt	gcactacatt	gctgctgagg	aggaggactg	ggactatgcc	cccctgggtgc	4320

	tggcccctga	tgacaggagc	tacaagagcc	agtacctgaa	caatggcccc	cagaggattg	4380
5	gcaggaagta	caagaaggtc	aggttcatgg	cctacactga	tgaaaccttc	aagaccaggg	4440
	aggccatcca	gcatgagtct	ggcatcctgg	gccccctgct	gtatggggag	gtgggggaca	4500
	ccctgctgat	catcttcaag	aaccaggcca	gcaggcccta	caacatctac	ccccatggca	4560
10	tactgatgt	gaggcccctg	tacagcagga	ggctgcccac	gggggtgaag	cacctgaagg	4620
	acttccccat	cctgcctggg	gagatcttca	agtacaagtg	gactgtgact	gtggaggatg	4680
15	gccccaccaa	gtctgacccc	aggtgcctga	ccagatacta	cagcagcttt	gtgaacatgg	4740
	agagggacct	ggcctctggc	ctgattggcc	ccctgctgat	ctgctacaag	gagtctgtgg	4800
	accagagggg	caaccagatc	atgtctgaca	agaggaatgt	gatcctgttc	tctgtgtttg	4860
20	atgagaacag	gagctggtac	ctgactgaga	acatccagag	gttcctgccc	aaccctgctg	4920
	gggtgcagct	ggaggaccct	gagttccagg	ccagcaacat	catgcacagc	atcaatggct	4980
25	atgtgtttga	cagcctgcag	ctgtctgtgt	gcctgcatga	ggtggcctac	tggtacatcc	5040
	tgagcattgg	ggcccagact	gacttcctgt	ctgtgttctt	ctctggctac	accttcaagc	5100
	acaagatggg	gtatgaggac	accctgaccc	tgttccccct	ctctggggag	actgtgttca	5160
30	tgagcatgga	gaaccctggc	ctgtggattc	tgggctgcca	caactctgac	ttcaggaaca	5220
	ggggcatgac	tgccctgctg	aaagtctcca	gctgtgacaa	gaacactggg	gactactatg	5280
35	aggacagcta	tgaggacatc	tctgcctacc	tgctgagcaa	gaacaatgcc	attgagccca	5340
	ggagcttcag	ccagaatgcc	actaatgtgt	ctaacaacag	caacaccagc	aatgacagca	5400
	atgtgtctcc	cccagtgctg	aagaggcacc	agagggagat	caccaggacc	accctgcagt	5460
40	ctgaccagga	ggagattgac	tatgatgaca	ccatctctgt	ggagatgaag	aaggaggact	5520
	ttgacatcta	cgacaggagc	gagaaccaga	gccccaggag	cttcagaaag	aagaccaggc	5580
45	actacttcat	tgctgctgtg	gagaggctgt	gggactatgg	catgagcagc	agcccccatg	5640
	tgctgaggaa	cagggcccag	tctggctctg	tgccccagtt	caagaagggtg	gtgttccagg	5700
	agttactga	tggcagcttc	accagcccc	tgtacagagg	ggagctgaat	gagcacctgg	5760
50	gcctgctggg	cccctacatc	agggtgagg	tggaggacaa	catcatgggtg	accttcagga	5820
	accaggccag	caggccctac	agcttctaca	gcagcctgat	cagctatgag	gaggaccaga	5880
55	ggcagggggc	tgagcccagg	agaacttttg	tgaagcccaa	tgaaaccaag	acctacttct	5940
	ggaagggtgca	gcaccacatg	gccccacca	aggatgagtt	tgactgcaag	gcctgggcct	6000
	acttctctga	tgtggacctg	gagaaggatg	tgcactctgg	cctgattggc	cccctgctgg	6060
60	tgtgccacac	caacaccctg	aaccctgccc	atggcaggca	ggtgactgtg	caggagtttg	6120
	ccctgttctt	caccatcttt	gatgaaacca	agagctggta	cttactgag	aacatggaga	6180
65	ggaactgcag	ggccccctgc	aacatccaga	tggaggaccc	caccttcaag	gagaactaca	6240
	ggttccatgc	catcaatggc	tacatcatgg	acaccctgcc	tggcctgggtg	atggcccagg	6300
	accagaggat	caggtggtac	ctgctgagca	tgggcagcaa	tgagaacatc	cacagcatcc	6360
70	acttctctgg	ccatgtgttc	actgtgagga	agaaggagga	gtacaagatg	gccctgtaca	6420

	acctgtaccc	tggggtgttt	gagactgtgg	agatgctgcc	cagcaaggct	ggcatctgga	6480
5	gggtggagt	cctgattggg	gagcacctgc	atgctggcat	gagcaccctg	ttcctggtgt	6540
	acagcaacaa	gtgccagacc	cccctgggca	tggcctcttg	ccacatcagg	gacttccaga	6600
	tactgcctc	tggccagtat	ggccagtggg	cccccaagct	ggccaggctg	cactactctg	6660
10	gcagcatcaa	tgcctggagc	accaaggagc	ccttcagctg	gatcaagggt	gacctgctgg	6720
	cccccatgat	catccatggc	atcaagaccc	agggggccag	gcagaagttc	agcagcctgt	6780
	acatcagcca	gttcatcatc	atgtacagcc	tggatggcaa	gaagtggcag	acctacaggg	6840
15	gcaacagcac	tggcaccctg	atggtgttct	ttggcaatgt	ggacagctct	ggcatcaagc	6900
	acaacatctt	caaccccccc	atcattgcc	gatacatcag	gctgcacccc	accactaca	6960
20	gcatcaggag	caccctgagg	atggagctga	tgggctgtga	cctgaacagc	tgcagcatgc	7020
	ccctgggcat	ggagagcaag	gccatctctg	atgccagat	cactgccagc	agctacttca	7080
	ccaacatgtt	tgccacctgg	agccccagca	aggccaggct	gcacctgcag	ggcaggagca	7140
25	atgcctggag	gccccaggct	aacaacccca	aggagtggct	gcagggtggac	ttccagaaga	7200
	ccatgaagg	gactgggggtg	accacccagg	gggtgaagag	cctgctgacc	agcatgtatg	7260
30	tgaaggagtt	cctgatcagc	agcagccagg	atggccacca	gtggaccctg	ttcttccaga	7320
	atggcaagg	gaagggtgtt	cagggcaacc	aggacagctt	caccctgtg	gtgaacagcc	7380
	tggaccccc	cctgctgacc	agatacctga	ggattcaccc	ccagagctgg	gtgcaccaga	7440
35	ttgccctgag	gatggagggtg	ctgggctgtg	aggcccagga	cctgtactga	gcggccgcgg	7500
	gccaatcaa	cctctggatt	acaaaatttg	tgaagattg	actggtattc	ttaactatgt	7560
40	tgctcctttt	acgctatgtg	gatacgtgc	tttaatgcct	ttgtatcatg	ctattgcttc	7620
	ccgataggct	ttcatTTTTt	cctccttgta	taaatcctgg	ttgctgtctc	tttatgagga	7680
	gttgtggccc	gttgtcaggc	aacgtggcgt	ggtgtgcact	gtgtttgctg	acgcaacccc	7740
45	cactggttgg	ggcattgcc	ccacctgtca	gctcctttcc	gggactttcg	ctttccccct	7800
	ccctattgcc	acggcggaac	tcacgcgcgc	ctgccttgcc	cgctgctgga	caggggctcg	7860
50	gctgttgggc	actgacaatt	ccgtggtgtt	gtcggggaaa	tcacgtcct	ttccttggct	7920
	gctgcctgt	gttgccacct	ggattctgcg	cgggacgtcc	ttctgctacg	tcccttcggc	7980
	cctcaatcca	gcggaacctt	cttcccgcgg	cctgctgccg	gctctgcggc	ctcttcgcgg	8040
55	tcttcgcctt	cgccctcaga	cgagtcggat	ctccctttgg	gccgcctccc	cgcaagcttc	8100
	gcacttttta	aaagaaaagg	gaggactgga	tgggatttat	tactccgata	ggacgctggc	8160
60	ttgtaactca	gtctcttact	aggagaccag	cttgagcctg	ggtgttcgct	ggttagccta	8220
	acctggttgg	ccaccagggg	taaggactcc	ttggcttaga	aagctaataa	acttgccctgc	8280
	attagagctc	ttacgcgtcc	cgggctcgag	atccgcatct	caattagtca	gcaaccatag	8340
65	tcccgcccct	aactccgccc	atcccgcgcc	taactccgcc	cagttccgcc	cattctccgc	8400
	cccatggctg	actaattttt	tttatttatg	cagaggccga	ggccgcctcg	gcctctgagc	8460
70	tattccagaa	gtagtgagga	ggcttttttg	gaggcctagg	cttttgcaaa	aagctaactt	8520

	gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa	8580
5	agcatttttt tctactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca	8640
	tgtctgtccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	8700
	gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga	8760
10	aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	8820
	gcgttttttc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag	8880
15	agggtggcgaa acccgacagg actataaaga taccaggcgt tttccccctgg aagctccctc	8940
	gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg	9000
	ggaagcgtgg cgcttttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt	9060
20	cgctccaagc tgggctgtgt gcacgaacct cccgttcagc ccgaccgctg cgccttatcc	9120
	ggtaactatc gtcttgagtc caaccgggta agacacgact tatcgccact ggcagcagcc	9180
25	actggtaaca ggattagcag agcgaggtat gtaggcgggtg ctacagagtt cttgaagtgg	9240
	tggcctaact acggctacac tagaagaaca gtatttggtg tctgcgctct gctgaagcca	9300
	gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc	9360
30	ggtgggtttt ttgtttgcaa gcagcagatt acgcgagaa aaaaaggatc tcaagaagat	9420
	cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggatt	9480
35	ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt	9540
	tttaaataca tctaaagtat atatgagtaa acttggtctg acagttagaa aaactcatcg	9600
	agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa	9660
40	agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc	9720
	tggtatcggg ctgcgattcc gactcgcca acatcaatac aacctattaa tttcccctcg	9780
45	tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat	9840
	ggcaacagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca	9900
	tcaaaatcac tcgcatcaac caaacggtta ttcattcgtg attgcgctg agcgagacga	9960
50	aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgagc	10020
	aacactgcca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg	10080
55	aatgctgttt ttccggggat cgcagtgggtg agtaaccatg catcatcagg agtacggata	10140
	aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca	10200
	tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcatcg	10260
60	ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat	10320
	ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctaga gcaagacggt	10380
65	tcccgttgaa tatggctcat aacaccctt gtattactgt ttatgtaagc agacagtttt	10440
	attgttcag atgatattt tttatcttgt gcaatgtaac atcagagatt ttgagacaca	10500
	acaattggtc gacggatcc	10519
70		

<210> 13
 <211> 11400
 <212> ДНК
 <213> штучна послідовність

5

<220>
 <223> pGM412 плазміда

<400> 13
 10 ggtacctcaa tattggccat tagccatatt attcattggt tatatagcat aaatcaatat 60
 tggctattgg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc 120
 atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat 180
 15 tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 240
 tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 300
 20 tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 360
 aactgcccac ttggcagtac atcaagtgtg tcatatgcc aagtcgcccc ctattgacgt 420
 caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttac gggactttcc 480
 25 tacttggcag tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca 540
 gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 600
 30 tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 660
 caactgcatg cggccgcccc gttgacgcaa atgggcggta ggcgtgtacg gtgggaggtc 720
 tatataagca gagctcgctg gcttgtaact cagtctctta ctaggagacc agcttgagcc 780
 35 tgggtgttcg ctggttagcc taacctgggt ggccaccagg ggtaaggact ccttggttta 840
 gaaagctaata aaacttgcct gcattagagc ttatctgagt caagtgtcct cattgacgcc 900
 40 tcaactctct gaacgggaat cttccttact gggttctctc tctgaccagc gcgagagaaa 960
 ctccagcagt ggcgccccga cagggacttg agtgagagtg taggcacgta cagctgagaa 1020
 ggcgtcggac gcgaaggaag cgcggggtgc gacgcgacca agaaggagac ttggtgagta 1080
 45 ggcttctcga gtgccgggaa aaagctcgag cctagttaga ggactaggag aggccgtagc 1140
 cgtaactact ctgggcaagt agggcaggcg gtgggtacgc aatgggggag gctacctcag 1200
 50 cactaaatag gagacaatta gaccaatttg agaaaatacg acttcgcccc aacggaaaga 1260
 aaaagtagca aattaaacat ttaatatggg caggcaagga gatggagcgc ttcggcctcc 1320
 atgagagggt gttggagaca gaggaggggt gtaaaagaat catagaagtc ctctaccccc 1380
 55 tagaaccaac aggatcggag ggcttaaaaa gtctgttcaa tcttgtgtgc gtgctatatt 1440
 gcttgcacaa ggaacagaaa gtgaaagaca cagaggaagc agtagcaaca gtaagacaac 1500
 60 actgccatct agtggaaaaa gaaaaaagt caacagagac atctagtggg caaaagaaaa 1560
 atgacaaggg aatagcagcg ccacctggtg gcagtcagaa tttccagcg caacaacaag 1620
 gaaatgcctg ggtacatgta cccttgtcac cgcgcacctt aaatgcgtgg gtaaaagcag 1680
 65 tagaggagaa aaaatttgga gcagaaatag taccatgtt tcaagcccta tcgaattccc 1740
 gtttgtgcta gggttcttag gcttcttggg ggctgctgga actgcaatgg gagcagcggc 1800
 70 gacagccctg acggtccagt ctcagcattt gcttgcctgg atactgcagc agcagaagaa 1860

	tctgctggcg	gctgtggagg	ctcaacagca	gatgttgaag	ctgaccattt	ggggtgttaa	1920
5	aaacctcaat	gcccgcgtca	cagcccttga	gaagtaccta	gaggatcagg	cacgactaaa	1980
	ctcctggggg	tgcgcatgga	aacaagtatg	tcataccaca	gtggagtggc	cctggacaaa	2040
	tcggactccg	gattggcaaa	atatgacttg	gttggagtgg	gaaagacaaa	tagctgattt	2100
10	ggaaagcaac	attacgagac	aattagtga	ggctagagaa	caagaggaaa	agaatctaga	2160
	tgcctatcag	aagttaacta	gttggtcaga	tttctgggtc	tggttcgatt	tctcaaaatg	2220
15	gcttaacatt	ttaaaaatgg	gatttttagt	aatagtagga	ataataggg	taagattact	2280
	ttacacagta	tatggatgta	tagtgaggg	taggcaggga	tatgttcctc	tatctccaca	2340
	gatccatc	cgcggaatt	ttaaaagaaa	gggaggaata	gggggacaga	cttcagcaga	2400
20	gagactaatt	aatataataa	caacacaatt	agaaatacaa	catttacaaa	ccaaaattca	2460
	aaaaatttta	aatttttagag	ccgcggagat	ctcaatattg	gccattagcc	atattattca	2520
25	ttggttatat	agcataaatc	aatattggct	attggccatt	gcatacgttg	tatctatatc	2580
	ataatatgta	catttatatt	ggctcatgtc	caatatgacc	gccatgttgg	cattgattat	2640
	tgactagtta	ttaatagtaa	tcaattacgg	ggtcattagt	tcatagcccc	tatatggagt	2700
30	tccgcgttac	ataacttacg	gtaaattggcc	cgcttggtcg	accgccaac	gacccccgcc	2760
	cattgacgtc	aataatgacg	tatgttcccc	tagtaacgcc	aatagggact	ttccattgac	2820
35	gtcaatgggt	ggagtattta	cggtaaactg	cccacttggc	agtacatcaa	gtgtatcata	2880
	tgccaagtcc	gccccctatt	gacgtcaatg	acggtaaatg	gcccgcctgg	cattatgccc	2940
	agtacatgac	cttacgggac	tttctacttt	ggcagtacat	ctacgtatta	gtcatcgcta	3000
40	ttaccatgg	gatgcgggtt	tggcagtaca	ccaatggg	tgatagcgg	tttgactcac	3060
	ggggatttcc	aagtctccac	cccattgacg	tcaatgggag	tttgttttgg	caccaaatac	3120
45	aacgggactt	tccaaaatgt	cgtaataacc	ccgccccgtt	gacgcaaata	ggcggtaggc	3180
	gtgtacgg	ggaggtctat	ataagcagag	ctcgtttagt	gaaccgtcag	atcactagaa	3240
	gctttattgc	ggtagtttat	cacagttaaa	ttgctaacgc	agtcagtgtc	tctgacacaa	3300
50	cagtctcgaa	cttaagctgc	agaagtgg	cgtgaggcac	tgggcaggct	agccaccaat	3360
	gcagattgag	ctgagcacct	gcttcttcct	gtgcctgtcg	aggttctgtc	tctctgccac	3420
55	caggagatac	tacctggggg	ctgtggagct	gagctgggac	tacatgcagt	ctgacctggg	3480
	ggagctgcct	gtggatgcca	ggttcccccc	cagagtgtcc	aagagcttcc	ccttcaacac	3540
	ctctgtggtg	tacaagaaga	ccctgtttgt	ggagttcact	gaccacctgt	tcaacattgc	3600
60	caagcccagg	ccccctgga	tgggcctgtc	gggccccacc	atccaggctg	aggtgtatga	3660
	cactgtggtg	atcacctga	agaacatggc	cagccaccct	gtgagcctgc	atgctgtggg	3720
65	ggtgagctac	tgaaggcct	ctgagggggc	tgagtatgat	gaccagacca	gccagaggga	3780
	gaaggaggat	gacaagggtg	tccctggggg	cagccacacc	tatgtgtggc	aggtgtctgaa	3840
	ggagaatggc	cccatggcct	ctgacccctt	gtgcctgacc	tacagctacc	tgagccatgt	3900
70	ggacctgg	aaggacctga	actctggcct	gattggggcc	ctgctggtgt	gcagggaggg	3960

	cagcctggcc	aaggagaaga	cccagaccct	gcacaagttc	atcctgctgt	ttgctgtggt	4020
5	tgatgagggc	aagagctggc	actctgaaac	caagaacagc	ctgatgcagg	acagggatgc	4080
	tgccctctgc	agggcctggc	ccaagatgca	cactgtgaat	ggctatgtga	acaggagcct	4140
	gcctggcctg	attggctgcc	acaggaagtc	tgtgtactgg	catgtgattg	gcatgggcac	4200
10	caccctgag	gtgcacagca	tcttcctgga	gggccacacc	ttcctgggtca	ggaaccacag	4260
	gcaggccagc	ctggagatca	gccccatcac	cttcctgact	gccagaccc	tgctgatgga	4320
	cctgggccag	ttcctgctgt	tctgccacat	cagcagccac	cagcatgatg	gcatggaggc	4380
15	ctatgtgaag	gtggacagct	gccctgagga	gccccagctg	aggatgaaga	acaatgagga	4440
	ggctgaggac	tatgatgatg	acctgactga	ctctgagatg	gatgtggtga	ggtttgatga	4500
20	tgacaacagc	cccagcttca	tccagatcag	gtctgtggcc	aagaagcacc	ccaagacctg	4560
	ggtgcactac	attgctgctg	aggaggagga	ctgggactat	gccccctgg	tgctggcccc	4620
	tgatgacagg	agctacaaga	gccagtacct	gaacaatggc	ccccagagga	ttggcaggaa	4680
25	gtacaagaag	gtcaggttca	tggcctacac	tgatgaaacc	ttcaagacca	gggaggccat	4740
	ccagcatgag	tctggcatcc	tgggccccct	gctgtatggg	gaggtggggg	acaccctgct	4800
30	gatcatcttc	aagaaccagg	ccagcaggcc	ctacaacatc	tacccccatg	gcatcactga	4860
	tgtgaggccc	ctgtacagca	ggaggctgcc	caagggggtg	aagcacctga	aggacttccc	4920
	catcctgcct	ggggagatct	tcaagtacaa	gtggactgtg	actgtggagg	atggccccac	4980
35	caagtctgac	cccaggtgcc	tgaccagata	ctacagcagc	tttgtgaaca	tggagagggga	5040
	cctggcctct	ggcctgattg	gccccctgct	gatctgctac	aaggagtctg	tggaccagag	5100
40	gggcaaccag	atcatgtctg	acaagaggaa	tgtgatcctg	ttctctgtgt	ttgatgagaa	5160
	caggagctgg	tacctgactg	agaacatcca	gaggttcctg	cccaaccctg	ctggggtgca	5220
	gctggaggac	cctgagttcc	aggccagcaa	catcatgcac	agcatcaatg	gctatgtggt	5280
45	tgacagcctg	cagctgtctg	tgtgcctgca	tgaggtggcc	tactggtaca	tcctgagcat	5340
	tggggccag	actgacttcc	tgtctgtggt	cttctctggc	tacaccttca	agcacaagat	5400
50	ggtgtatgag	gacaccctga	ccctgttccc	cttctctggg	gagactgtgt	tcatgagcat	5460
	ggagaaccct	ggcctgtgga	ttctgggctg	ccacaactct	gacttcagga	acaggggcat	5520
	gactgccctg	ctgaaagtct	ccagctgtga	caagaacact	ggggactact	atgaggacag	5580
55	ctatgaggac	atctctgcct	acctgctgag	caagaacaat	gccattgagc	ccaggagctt	5640
	cagccagaac	agcaggcacc	ccagcaccag	gcagaagcag	ttcaatgcca	ccaccatccc	5700
60	tgagaatgac	atagagaaga	cagacccatg	gtttgcccac	cggaccccca	tgcccaagat	5760
	ccagaatgtg	agcagctctg	acctgctgat	gctgctgagg	cagagcccca	ccccccatgg	5820
	cctgagcctg	tctgacctgc	aggaggccaa	gtatgaaacc	ttctctgatg	accccagccc	5880
65	tggggccatt	gacagcaaca	acagcctgtc	tgagatgacc	cacttcaggc	cccagctgca	5940
	ccactctggg	gacatggtgt	tcacccttga	gtctggcctg	cagctgaggc	tgaatgagaa	6000
70	gctgggcacc	actgctgcca	ctgagctgaa	gaagctggac	ttcaaagtct	ccagcaccag	6060

	caacaacctg	atcagcacca	tcccctctga	caacctggct	gctggcactg	acaacaccag	6120
5	cagcctgggc	ccccccagca	tgcctgtgca	ctatgacagc	cagctggaca	ccaccctggt	6180
	tggcaagaag	agcagccccc	tgactgagtc	tggggggccc	ctgagcctgt	ctgaggagaa	6240
	caatgacagc	aagctgctgg	agtctggcct	gatgaacagc	caggagagca	gctggggcaa	6300
10	gaatgtgagc	agcagggaga	tcaccaggac	caccctgcag	tctgaccagg	aggagattga	6360
	ctatgatgac	accatctctg	tggagatgaa	gaaggaggac	tttgacatct	acgacgagga	6420
	cgagaaccag	agccccagga	gcttccagaa	gaagaccagg	cactacttca	ttgctgctgt	6480
15	ggagaggctg	tgggactatg	gcatgagcag	cagcccccat	gtgctgagga	acagggccca	6540
	gtctggctct	gtgccccagt	tcaagaaggt	ggtgttccag	gagttcactg	atggcagctt	6600
20	caccagccc	ctgtacagag	gggagctgaa	tgagcacctg	ggcctgctgg	gccctacat	6660
	cagggctgag	gtggaggaca	acatcatggt	gaccttcagg	aaccaggcca	gcaggcccta	6720
	cagcttctac	agcagcctga	tcagctatga	ggaggaccag	aggcaggggg	ctgagcccag	6780
25	gaagaacttt	gtgaagccca	atgaaaccaa	gacctacttc	tggaaggtgc	agcaccacat	6840
	ggccccacc	aaggatgagt	ttgactgcaa	ggcctgggcc	tacttctctg	atgtggacct	6900
30	ggagaaggat	gtgcactctg	gcctgattgg	ccccctgctg	gtgtgccaca	ccaacaccct	6960
	gaaccctgcc	catggcaggc	aggtgactgt	gcaggagttt	gccctgttct	tcaccatctt	7020
	tgatgaaacc	aagagctggg	acttcactga	gaacatggag	aggaactgca	gggccccctg	7080
35	caacatccag	atggaggacc	ccaccttcaa	ggagaactac	aggttccatg	ccatcaatgg	7140
	ctacatcatg	gacaccctgc	ctggcctggg	gatggcccag	gaccagagga	tcaggtggta	7200
40	cctgctgagc	atgggcagca	atgagaacat	ccacagcatc	cacttctctg	gccatgtggt	7260
	cactgtgagg	aagaaggagg	agtacaagat	ggcctgttac	aacctgtacc	ctggggtggt	7320
	tgagactgtg	gagatgctgc	ccagcaaggc	tggcatctgg	aggggtggagt	gcctgattgg	7380
45	ggagcacctg	catgctggca	tgagcacctt	gttcctgggt	tacagcaaca	agtgccagac	7440
	ccccctgggc	atggcctctg	gccacatcag	ggacttccag	atcactgcct	ctggccagta	7500
50	tggccagtgg	gcccccaagc	tggccaggct	gcactactct	ggcagcatca	atgcctggag	7560
	caccaaggag	cccttcagct	ggatcaaggt	ggacctgctg	gcccccatga	tcatccatgg	7620
	catcaagacc	cagggggcca	ggcagaagtt	cagcagcctg	tacatcagcc	agttcatcat	7680
55	catgtacagc	ctggatggca	agaagtggca	gacctacagg	ggcaacagca	ctggcacctt	7740
	gatggtgttc	tttggcaatg	tggacagctc	tggcatcaag	cacaacatct	tcaaccccc	7800
60	catcattgcc	agatacatca	ggctgcaccc	caccacttac	agcatcagga	gcaccctgag	7860
	gatggagctg	atgggctgtg	acctgaacag	ctgcagcatg	cccctgggca	tggagagcaa	7920
	ggccatctct	gatgcccaga	tactgccag	cagctacttc	accaacatgt	ttgccacctg	7980
65	gagccccagc	aaggccaggc	tgcacctgca	gggcaggagc	aatgcctgga	ggccccaggt	8040
	caacaacccc	aaggagtggc	tgcaggtgga	cttccagaag	accatgaagg	tgactggggt	8100
70	gaccacccag	ggggtgaaga	gcctgctgac	cagcatgtat	gtgaaggagt	tcctgatcag	8160

	cagcagccag	gatggccacc	agtggaccct	gttcttccag	aatggcaagg	tgaaggtggt	8220
5	ccagggcaac	caggacagct	tcaccctgt	ggtgaacagc	ctggaccccc	ccctgctgac	8280
	cagatacctg	aggattcacc	cccagagctg	ggtgcaccag	attgccctga	ggatggaggt	8340
	gctgggctgt	gaggcccagg	acctgtactg	agcggccgcg	ggcccaatca	acctctggat	8400
10	tacaaaattt	gtgaaagatt	gactgggtatt	cttaactatg	ttgctccttt	tacgctatgt	8460
	ggatacgctg	ctttaatgcc	tttgtatcat	gctattgctt	cccgtatggc	tttcattttc	8520
	tcctccttgt	ataaatcctg	gttgctgtct	ctttatgagg	agttgtggcc	cgttgtcagg	8580
15	caacgtggcg	tgggtgtgcac	tgtgtttgct	gacgcaaccc	ccactggttg	gggcattgcc	8640
	accacctgtc	agctcctttc	cgggactttc	gctttccccc	tccctattgc	cacggcggaa	8700
20	ctcatcgccg	cctgccttgc	ccgctgctgg	acaggggctc	ggctgttggg	cactgacaat	8760
	tccgtggtgt	tgtcggggaa	atcatcgctc	tttccttggc	tgctcgcctg	tgttgccacc	8820
	tggattctgc	gcgggacgtc	cttctgtac	gtcccttcgg	ccctcaatcc	agcggacctt	8880
25	ccttcccgcg	gcctgctgcc	ggctctgcgg	cctcttcgcg	gtcttcgcct	tcgccctcag	8940
	acgagtcgga	tctccctttg	ggccgcctcc	ccgcaagctt	cgcacttttt	aaaagaaaag	9000
30	ggaggactgg	atgggattta	ttactccgat	aggacgctgg	cttgtaactc	agtctcttac	9060
	taggagacca	gcttgagcct	gggtgttcgc	tggttagcct	aacctggttg	gccaccaggg	9120
	gtaaggactc	cttggcttag	aaagctaata	aacttgccctg	cattagagct	cttacgcgtc	9180
35	ccgggctcga	gatccgcata	tcaattagtc	agcaaccata	gtcccgcgcc	taactccgcc	9240
	catcccgcgc	ctaactccgc	ccagttccgc	ccattctccg	cccatgggt	gactaatfff	9300
40	ttttatttat	gcagaggccg	aggccgcctc	ggcctctgag	ctattccaga	agtagtgagg	9360
	aggctttttt	ggaggcctag	gcttttgcaa	aaagctaact	tgtttattgc	agcttataat	9420
	ggttacaaat	aaagcaatag	catcacaat	ttcacaata	aagcattttt	ttactgcat	9480
45	tctagtgtgt	gtttgtccaa	actcatcaat	gtatcttatc	atgtctgtcc	gcttcctcgc	9540
	tcactgactc	gctgcgctcg	gtcgttcggc	tgccggcagc	ggtatcagct	cactcaaagg	9600
50	cggtaatcag	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	9660
	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	9720
	gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	9780
55	gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	9840
	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcggt	gcgctttctc	9900
60	atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	9960
	tgcacgaacc	ccccgttcag	cccgaccgct	gcgccttatc	cggtaactat	cgtcttgagt	10020
	ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	10080
65	gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	10140
	ctagaagaac	agtatttggt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	10200
70	ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cgggtggtttt	tttgtttgca	10260

	agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg	10320
	ggctctgacgc tcagtggaaac gaaaactcac gttaagggat tttggtcattg agattatcaa	10380
5	aaaggatcctt cacctagatc cttttaaatt aaaaatgaag ttttaaataca atctaaagta	10440
	tatatgagta aacttgggtct gacagttaga aaaactcatc gagcatcaaa tgaaactgca	10500
10	atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc tgtaatgaag	10560
	gagaaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg tctgcgattc	10620
	cgactcgtcc aacatcaata caacctatta atttcccctc gtcaaaaata aggttatcaa	10680
15	gtgagaaaatc accatgagtg acgactgaat ccggtgagaa tggcaacagc ttatgcattt	10740
	ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca ctcgcatcaa	10800
20	ccaaaccgtt attcattcgt gattgctgctt gagcgcagacg aaatacgcga tcgctgttaa	10860
	aaggacaatt acaaacagga atcgaatgca accggcgcag gaacactgcc agcgcaccaa	10920
	caatatttttc acctgaatca ggatattctt ctaataacctg gaatgctgtt tttccgggga	10980
25	tcgcagtggg gagtaaccat gcatcatcag gagtacggat aaaatgcttg atggtcggaa	11040
	gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca tcattggcaa	11100
30	cgctaccttt gccatgtttc agaaacaact ctggcgcgacg gggcttccca tacaatcgat	11160
	agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca tataaatcag	11220
	catccatggt ggaatttaat cgcggcctag agcaagacgt ttcccgttga atatggctca	11280
35	taacaccctt tgtattactg tttatgtaag cagacagttt tattgttcat gatgatatat	11340
	ttttatcttg tgcaatgtaa catcagagat tttgagacac aacaattggg cgacggatcc	11400
40	<210> 14	
	<211> 11108	
	<212> ДНК	
	<213> штучна послідовність	
45	<220>	
	<223> pGM414 плазміда	
	<400> 14	
50	ggtacctcaa tattggccat tagccatatt attcattggg tatatagcat aaatcaatat	60
	tggctatttg ccattgcata cgttgtatct atatcataat atgtacattt atattggctc	120
	atgtccaata tgaccgccat gttggcattg attattgact agttattaat agtaatcaat	180
55	tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa	240
	tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt	300
60	tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	360
	aactgcccac ttggcagtac atcaagtgtg tcatatgccg agtccgcccc ctattgacgt	420
	caatgacggg aaatggcccc cctggcatta tgcccagtac atgaccttac gggactttcc	480
65	tacttggcag tacatctacg tatttagtcat cgctattacc atggtgatgc ggttttggca	540
	gtacaccaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat	600
70	tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa	660

	caactgcgat	cgcccgcccc	gttgacgcaa	atgggaggta	ggcgtgtacg	gtgggaggtc	720
5	tatataagca	gagctcgctg	gcttgtaact	cagtctctta	ctaggagacc	agcttgagcc	780
	tgggtgttcg	ctgggttagcc	taacctgggt	ggccaccagg	ggtaaggact	ccttggtcta	840
	gaaagcta	aaacttgctt	gcattagagc	ttatctgagt	caagtgtcct	cattgacgcc	900
10	tcactctctt	gaacgggaat	cttccttact	gggttctctc	tctgaccag	gcgagagaaa	960
	ctccagcagt	ggcgcccgaa	cagggacttg	agtgagagtg	taggcacgta	cagctgagaa	1020
15	ggcgtcggac	gcgaaggaag	cgcggggtgc	gacgcgacca	agaaggagac	ttggtgagta	1080
	ggcttctcga	gtgccgggaa	aaagctcgag	cctagttaga	ggactaggag	agggcgtagc	1140
	cgtaactact	cttgggcaag	tagggcaggc	ggtgggtacg	caatgggggc	ggctacctca	1200
20	gcactaaata	ggagacaatt	agaccaat	gagaaaatac	gacttcgccc	gaacggaaag	1260
	aaaaagtacc	aaattaaaca	tttaatatgg	gcaggcaagg	agatggagcg	cttcggcctc	1320
25	catgagaggt	tgttgagagc	agaggagggg	tgtaaaagaa	tcatagaagt	cctctacccc	1380
	ctagaaccaa	caggatcgga	gggcttaaaa	agtctgttca	atcttgtgtg	cggtctatat	1440
	tgcttgacaa	aggaacagaa	agtgaagac	acagaggaag	cagtagcaac	agtaagacaa	1500
30	cactgccatc	tagtggaata	agaaaaagt	gcaacagaga	catctagtgg	acaaaagaaa	1560
	aatgacaagg	gaatagcagc	gccacctggg	ggcagtcaga	atcttcagc	gcaacaacaa	1620
35	ggaaatgcct	gggtacatgt	acccttgtca	ccgcgcacct	taaatgcgtg	ggtaaaagca	1680
	gtagaggaga	aaaaatttgg	agcagaaata	gtacccatgt	ttcaagccct	atcgaattcc	1740
	cgtttgtgct	agggttctta	ggcttcttgg	gggctgctgg	aactgcaatg	ggagcagcgg	1800
40	cgacagccct	gacggtccag	tctcagcatt	tgcttgctgg	gatactgcag	cagcagaaga	1860
	atctgctggc	ggctgtggag	gctcaacagc	agatgttgaa	gctgaccatt	tggggtgtta	1920
45	aaaacctcaa	tgccgcgctc	acagcccttg	agaagtacct	agaggatcag	gcacgactaa	1980
	actcctgggg	gtgcgcatgg	aaacaagtat	gtcataccac	agtggagtgg	ccctggacaa	2040
	atcggactcc	ggattggcaa	aatatgactt	ggttggagtg	ggaaagacaa	atagctgatt	2100
50	tggaaagcaa	cattacgaga	caattagtga	aggctagaga	acaagaggaa	aagaatctag	2160
	atgcctatca	gaagttaact	agttggtcag	atctctggtc	ttggttcgat	ttctcaaaat	2220
55	ggcttaacat	tttaaaaatg	ggatttttag	taatagtagg	aataataggg	ttaagattac	2280
	tttacacagt	atatggatgt	atagtggagg	ttaggcaggg	atatgttcct	ctatctccac	2340
	agatccatat	ccgcggcaat	tttaaaagaa	agggaggaat	agggggacag	acttcagcag	2400
60	agagactaat	taatataata	acaacacaat	tagaaataca	acatttacaa	acaaaatttc	2460
	aaaaaatttt	aaattttaga	gccgcggaga	tctgttacat	aacttatggg	aatggcctg	2520
65	cctggctgac	tgcccaatga	cccctgcccc	atgatgtcaa	taatgatgta	tgttcccatg	2580
	taatgccaat	agggactttc	cattgatgtc	aatgggtgga	gtatttatgg	taactgcccc	2640
	cttggcagta	catcaagtgt	atcatatgcc	aagtatgccc	cctattgatg	tcaatgatgg	2700
70	taaatggcct	gcctggcatt	atgcccagta	catgacctta	tgggactttc	ctacttggca	2760

	gtacatctat	gtattagtca	ttgctattac	catgggaatt	cactagtgga	gaagagcatg	2820
5	cttgagggct	gagtggccct	cagtgggcag	agagcacatg	gcccacagtc	cctgagaagt	2880
	tggggggagg	ggtgggcaat	tgaactggtg	cctagagaag	gtggggcctg	ggtaaactgg	2940
	gaaagtgatg	tgggtgtactg	gctccacctt	tttccccagg	gtgggggaga	accatatata	3000
10	agtgcagtag	tctctgtgaa	cattcaagct	tctgccttct	ccctcctgtg	agtttgctag	3060
	ccaccaatgc	agattgagct	gagcacctgc	ttcttcctgt	gcctgctgag	gttctgcttc	3120
15	tctgccacca	ggagatacta	cctgggggct	gtggagctga	gctgggacta	catgcagtct	3180
	gacctggggg	agctgcctgt	ggatgccagg	ttcccccca	gagtggccaa	gagcttcccc	3240
	ttcaacacct	ctgtggtgta	caagaagacc	ctgtttgtgg	agttcactga	ccacctgttc	3300
20	aacattgcca	agcccaggcc	cccctggatg	ggcctgctgg	gccccaccat	ccaggctgag	3360
	gtgtatgaca	ctgtggtgat	caccctgaag	aacatggcca	gccaccctgt	gagcctgcat	3420
25	gctgtggggg	tgagctactg	gaaggcctct	gagggggctg	agtatgatga	ccagaccagc	3480
	cagagggaga	aggaggatga	caaggtgttc	cctgggggca	gccacaccta	tgtgtggcag	3540
	gtgctgaagg	agaatggccc	catggcctct	gacccctgt	gcctgaccta	cagctacctg	3600
30	agccatgtgg	acctggtgaa	ggacctgaac	tctggcctga	ttggggccct	gctggtgtgc	3660
	agggagggca	gcctggccaa	ggagaagacc	cagaccctgc	acaagtcat	cctgctgttt	3720
35	gctgtgtttg	atgaggggcaa	gagctggcac	tctgaaacca	agaacagcct	gatgcaggac	3780
	agggatgctg	cctctgccag	ggcctggccc	aagatgcaca	ctgtgaatgg	ctatgtgaac	3840
	aggagcctgc	ctggcctgat	tggctgccac	aggaagtctg	tgtactggca	tgtgattggc	3900
40	atgggcacca	cccctgaggt	gcacagcatc	ttcctggagg	gccacacctt	cctggtcagg	3960
	aaccacaggc	aggccagcct	ggagatcagc	cccatcacct	tcctgactgc	ccagaccctg	4020
45	ctgatggacc	tgggccagtt	cctgctgttc	tgccacatca	gcagccacca	gcatgatggc	4080
	atggaggcct	atgtgaaggt	ggacagctgc	cctgaggagc	cccagctgag	gatgaagaac	4140
	aatgaggagg	ctgaggacta	tgatgatgac	ctgactgact	ctgagatgga	tgtggtgagg	4200
50	tttgatgatg	acaacagccc	cagcttcatc	cagatcaggt	ctgtggccaa	gaagcacccc	4260
	aagacctggg	tgcactacat	tgctgctgag	gaggaggact	gggactatgc	ccccctggtg	4320
55	ctggcccctg	atgacaggag	ctacaagagc	cagtacctga	acaatggccc	ccagaggatt	4380
	ggcaggaagt	acaagaaggt	caggttcatg	gcctacactg	atgaaacctt	caagaccagg	4440
	gaggccatcc	agcatgagtc	tggcatcctg	ggccccctgc	tgtatgggga	ggtggggggac	4500
60	acctgctga	tcatcttcaa	gaaccaggcc	agcaggccct	acaacatcta	cccccatggc	4560
	atcactgatg	tgaggcccct	gtacagcagg	aggctgccc	agggggtgaa	gcacctgaag	4620
65	gacttcccca	tcctgcctgg	ggagatcttc	aagtacaagt	ggactgtgac	tgtggaggat	4680
	ggccccacca	agtctgacct	caggtgcctg	accagatact	acagcagctt	tgtgaacatg	4740
	gagagggacc	tggcctctgg	cctgattggc	cccctgctga	tctgctacaa	ggagtctgtg	4800
70	gaccagaggg	gcaaccagat	catgtctgac	aagaggaatg	tgatcctggt	ctctgtgttt	4860

	gatgagaaca	ggagctggta	cctgactgag	aacatccaga	ggttcctgcc	caaccctgct	4920
5	ggggtgcagc	tggaggaccc	tgagttccag	gccagcaaca	tcatgcacag	catcaatggc	4980
	tatgtgtttg	acagcctgca	gctgtctgtg	tgcctgcatg	agggtggccta	ctggtacatc	5040
	ctgagcattg	gggcccagac	tgacttcctg	tctgtgttct	tctctggcta	caccttcaag	5100
10	cacaagatgg	tgtatgagga	caccctgacc	ctgttcccct	tctctgggga	gactgtgttc	5160
	atgagcatgg	agaaccctgg	cctgtggatt	ctgggctgcc	acaactctga	cttcagggaac	5220
15	aggggcatga	ctgccctgct	gaaagtctcc	agctgtgaca	agaacactgg	ggactactat	5280
	gaggacagct	atgaggacat	ctctgcctac	ctgctgagca	agaacaatgc	cattgagccc	5340
	aggagcttca	gccagaacag	caggcacccc	agcaccaggc	agaagcagtt	caatgccacc	5400
20	accatcccctg	agaatgacat	agagaagaca	gacccatggt	ttgcccaccg	gacccccatg	5460
	cccaagatcc	agaatgtgag	cagctctgac	ctgctgatgc	tgctgaggca	gagccccacc	5520
25	ccccatggcc	tgagcctgtc	tgacctgcag	gaggccaagt	atgaaacctt	ctctgatgac	5580
	cccagcccctg	gggccattga	cagcaacaac	agcctgtctg	agatgaccca	cttcaggccc	5640
	cagctgcacc	actctgggga	catggtgttc	acccctgagt	ctggcctgca	gctgaggctg	5700
30	aatgagaagc	tgggcaccac	tgctgccact	gagctgaaga	agctggactt	caaagtctcc	5760
	agcaccagca	acaacctgat	cagcaccatc	ccctctgaca	acctggctgc	tggcactgac	5820
35	aacaccagca	gcctggggccc	ccccagcatg	cctgtgcact	atgacagcca	gctggacacc	5880
	accctgtttg	gcaagaagag	cagccccctg	actgagtctg	ggggccccct	gagcctgtct	5940
	gaggagaaca	atgacagcaa	gctgctggag	tctggcctga	tgaacagcca	ggagagcagc	6000
40	tggggcaaga	atgtgagcag	caggggagatc	accaggacca	ccctgcagtc	tgaccaggag	6060
	gagattgact	atgatgacac	catctctgtg	gagatgaaga	aggaggactt	tgacatctac	6120
45	gacgaggacg	agaaccagag	ccccaggagc	ttccagaaga	agaccaggca	ctacttcatt	6180
	gctgctgtgg	agaggctgtg	ggactatggc	atgagcagca	gcccccatgt	gctgaggaac	6240
	aggggcccagt	ctggctctgt	gccccagttc	aagaagggtg	tggtccagga	gttactgat	6300
50	ggcagcttca	cccagcccct	gtacagaggg	gagctgaatg	agcacctggg	cctgctgggc	6360
	ccctacatca	gggctgaggt	ggaggacaac	atcatggtga	ccttcaggaa	ccaggccagc	6420
55	aggccctaca	gcttctacag	cagcctgatc	agctatgagg	aggaccagag	gcagggggct	6480
	gagcccagga	agaactttgt	gaagcccaat	gaaaccaaga	cctacttctg	gaagggtcag	6540
	caccacatgg	cccccaccaa	ggatgagttt	gactgcaagg	cctgggccta	cttctctgat	6600
60	gtggacctgg	agaaggatgt	gcactctggc	ctgattggcc	ccctgctggt	gtgccacacc	6660
	aacaccctga	accctgccc	tggcaggcag	gtgactgtgc	aggagtttgc	cctgttcttc	6720
65	accatctttg	atgaaaccaa	gagctggtac	ttactgaga	acatggagag	gaactgcagg	6780
	gccccctgca	acatccagat	ggaggacccc	accttcaagg	agaactacag	gttccatgcc	6840
	atcaatggct	acatcatgga	caccctgcct	ggcctggtga	tggcccagga	ccagaggatc	6900
70	agggtgtacc	tgctgagcat	gggcagcaat	gagaacatcc	acagcatcca	cttctctggc	6960

	catgtgttca	ctgtgaggaa	gaaggaggag	tacaagatgg	ccctgtacaa	cctgtaccct	7020
	ggggtgtttg	agactgtgga	gatgctgccc	agcaaggctg	gcatctggag	ggtggagtgc	7080
5	ctgattgggg	agcacctgca	tgctggcatg	agcaccctgt	tcctgggtgta	cagcaacaag	7140
	tgccagaccc	ccctgggcat	ggcctctggc	cacatcaggg	acttccagat	cactgcctct	7200
10	ggccagtatg	gccagtgggc	ccccaagctg	gccaggctgc	actactctgg	cagcatcaat	7260
	gcctggagca	ccaaggagcc	cttcagctgg	atcaagggtg	acctgctggc	ccccatgatc	7320
	atccatggca	tcaagaccca	ggggggccagg	cagaagttca	gcagcctgta	catcagccag	7380
15	ttcatcatca	tgtacagcct	ggatggcaag	aagtggcaga	cctacagggg	caacagcact	7440
	ggcaccctga	tggtgttctt	tggcaatgtg	gacagctctg	gcatcaagca	caacatcttc	7500
20	aaccccccca	tcattgccag	atacatcagg	ctgcacccca	cccactacag	catcaggagc	7560
	accctgagga	tggagctgat	gggctgtgac	ctgaacagct	gcagcatgcc	cctgggcatg	7620
	gagagcaagg	ccatctctga	tgcccagatc	actgccagca	gctacttcac	caacatgttt	7680
25	gccacctgga	gccccagcaa	ggccaggctg	cacctgcagg	gcaggagcaa	tgccctggagg	7740
	ccccagggtca	acaaccccaa	ggagtggctg	caggtggact	tccagaagac	catgaagggtg	7800
30	actgggggtga	ccaccagggg	ggtgaagagc	ctgctgacca	gcatgtatgt	gaaggagttc	7860
	ctgatcagca	gcagccagga	tggccaccag	tggaccctgt	tcttccagaa	tggcaagggtg	7920
	aagggtgttc	agggcaacca	ggacagcttc	accctgtgtg	tgaacagcct	ggaccccccc	7980
35	ctgctgacca	gatacctgag	gattcacccc	cagagctggg	tgccacagat	tgccctgagg	8040
	atggaggtgc	tgggctgtga	ggcccaggac	ctgtactgag	cggccgcggg	cccaatcaac	8100
40	ctctggatta	caaaatttgt	gaaagattga	ctggtattct	taactatggt	gctcctttta	8160
	cgctatgtgg	atacgtgct	ttaatgcctt	tgtatcatgc	tattgcttcc	cgtatggctt	8220
	tcattttctc	ctccttgat	aaatcctggt	tgctgtctct	ttatgaggag	ttgtggcccg	8280
45	ttgtcaggca	acgtggcgtg	gtgtgcaactg	tgtttgctga	cgcaaccccc	actggttggg	8340
	gcattgccac	cacctgtcag	ctcctttccg	ggactttcgc	tttccccctc	cctattgccca	8400
50	cggcggaact	catcgccgcc	tgccctgccc	gctgctggac	aggggctcgg	ctggtgggca	8460
	ctgacaattc	cgtgggtgtg	tcggggaaat	catcgtcctt	tccttggtg	ctcgcctgtg	8520
	ttgccacctg	gattctgcgc	gggacgtcct	tctgctacgt	cccttcggcc	ctcaatccag	8580
55	cggaccttcc	ttcccgcggc	ctgctgccgg	ctctgcggcc	tcttccgcgt	cttcgccttc	8640
	gccctcagac	gagtcggatc	tccctttggg	cgcctcccc	gcaagcttcg	cactttttaa	8700
60	aagaaaagg	aggactggat	gggatttatt	actccgatag	gacgctggct	tgtaactcag	8760
	tctcttacta	ggagaccagc	ttgagcctgg	gtgttcgctg	gttagcctaa	cctggttggc	8820
	caccaggggt	aaggactcct	tggcttagaa	agctaataaa	cttgccctgca	ttagagctct	8880
65	tacgcgtccc	gggctcgaga	tccgcattct	aattagtcag	caaccatagt	cccgccccta	8940
	actccgccc	tcccggccc	aactccgccc	agttccgccc	attctccgccc	ccatggctga	9000
70	ctaatttttt	ttatttatgc	agaggccgag	gccgcctcgg	cctctgagct	attccagaag	9060

	tagtgaggag	gcttttttgg	aggcctaggc	ttttgcaaaa	agctaacttg	tttattgcag	9120
5	cttataatgg	ttacaaataa	agcaatagca	tcacaaatth	cacaaataaa	gcattttttt	9180
	cactgcattc	tagttgtggt	ttgtccaaac	tcacaaatth	atcttatcat	gtctgtccgc	9240
	ttcctcgctc	actgactcgc	tgcgctcgg	cgttcggtg	cggcgagcgg	tatcagctca	9300
10	ctcaaaggcg	gtaatacgg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	9360
	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	9420
15	taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggaggcgaaa	9480
	cccgcagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	9540
	tgttccgacc	ctgccgctta	ccgataacct	gtccgccttt	ctcccttcgg	gaagcggtgc	9600
20	gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcgttc	gtccaagct	9660
	gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	9720
25	tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	9780
	gattagcaga	gcgaggatg	taggcgggtg	tacagagttc	ttgaagtgg	ggcctaacta	9840
	cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	9900
30	aaaaagagtt	gtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	9960
	tgtttgcaag	cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	10020
35	ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggaatt	tggatcatgag	10080
	attatcaaaa	aggatcttca	cctagatcct	tttaaatata	aaatgaagtt	ttaaatcaat	10140
	ctaaagtata	tatgagtaaa	cttgggtctga	cagttagaaa	aactcatcga	gcatcaaatg	10200
40	aaactgcaat	ttattcatat	caggattatc	aataccatat	ttttgaaaaa	gccgtttctg	10260
	taatgaagga	gaaaactcac	cgaggcagtt	ccataggatg	gcaagatcct	ggtatcggtc	10320
45	tgcgattccg	actcgtccaa	catcaatata	acctattaat	ttcccctcgt	caaaaataag	10380
	gttatcaagt	gagaaatcac	catgagtgac	gactgaatcc	ggtgagaatg	gcaacagctt	10440
	atgcatttct	ttccagactt	gttcaacagg	ccagccatta	cgctcgatcat	caaaatcact	10500
50	cgcatcaacc	aaaccgttat	tcattcgtga	ttgcgcctga	gcgagacgaa	atacgcatc	10560
	gctgttaaaa	ggacaattac	aaacaggaat	cgaatgcaac	cggcgcagga	acactgccag	10620
55	cgcatcaaca	atattttcac	ctgaatcagg	atattcttct	aatacctgga	atgctgtttt	10680
	tccggggatc	gcagtgggtg	gtaaccatgc	atcatcagga	gtacggataa	aatgcttgat	10740
	ggtcggaaga	ggcataaatt	ccgtcagcca	gtttagtctg	accatctcat	ctgtaacatc	10800
60	attggcaacg	ctacctttgc	catgtttcag	aaacaactct	ggcgcatcgg	gcttcccata	10860
	caatcgatag	attgtcgcac	ctgattgccc	gacattatcg	cgagcccat	tatacccata	10920
65	taaatcagca	tccatgttgg	aatttaaatc	cggcctagag	caagacgttt	cccgttgaat	10980
	atggctcata	acacccttg	tattactgtt	tatgtaagca	gacagtttta	ttgttcatga	11040
	tgatatatth	ttatcttggt	caatgtaaca	tcagagatth	tgagacacia	caattgggtc	11100
70	acggatcc						11108

5	<210>	15	
	<211>	1257	
	<212>	ДНК	
	<213>	штучна послідовність	
10	<220>		
	<223>	кодон-оптимізований збіднений CpG трансген A1AT	
	<400>	15	
	atgccagct	ctgtgtcctg	gggcattctg
15	gtgtccctgg	ctgaggaccc	tcagggggat
	gaccaggacc	acccacacct	caacaagatc
	ctgtacagac	agctggccca	ccagagcaac
	attgccacag	cctttgccat	gctgagcctg
20	ctggaaggcc	tgaacttcaa	cctgacagag
	caggaactgc	tgagaaccct	gaaccagcca
	gggctgttcc	tgtctgaggg	cctgaagctg
	ctgtaccact	ctgaggcctt	cacagtgaac
25	atcaatgact	atgtggaaaa	gggcacccag
	gacagggaca	ctgtgtttgc	ccttgtgaac
	ccctttgaag	tgaaggacac	tgaggaagag
	aaggtgcaa	tgatgaagag	actggggatg
30	agctgggtgc	tgctgatgaa	gtacctgggc
	gagggcaagc	tgacgacact	ggaaaatgag
	gaaaatgagg	acagaagatc	tgccagcctg
	tatgacctga	agtctgtgct	gggacagctg
35	gacctgagtg	gagtgacaga	ggaagcccct
	gtgctgacca	ttgatgagaa	gggcacagag
	cccatgtcca	tccccccaga	agtgaagttc
	cagaacacca	agagccccct	gttcatgggc
40	gagggcaagc	tgacgacact	ggaaaatgag
	gaaaatgagg	acagaagatc	tgccagcctg
	tatgacctga	agtctgtgct	gggacagctg
	gacctgagtg	gagtgacaga	ggaagcccct
45	gtgctgacca	ttgatgagaa	gggcacagag
	cccatgtcca	tccccccaga	agtgaagttc
	cagaacacca	agagccccct	gttcatgggc
	aaggttgatg	agctgagcac	ctgcttcttc
50	atgcagattg	agctgagcac	ctgcttcttc
	accaggagat	actacctggg	ggctgtggag
	ggggagctgc	ctgtggatgc	caggttcccc
	acctctgtgg	tgtacaagaa	gaccctgttt
55	<210>	16	
	<211>	5013	
	<212>	ДНК	
	<213>	штучна послідовність	
60	<220>		
	<223>	кодон-оптимізований збіднений CpG трансген FVIII (N6)	
	<400>	16	
	atgcagattg	agctgagcac	ctgcttcttc
65	accaggagat	actacctggg	ggctgtggag
	ggggagctgc	ctgtggatgc	caggttcccc
	acctctgtgg	tgtacaagaa	gaccctgttt
	gtgagattga	acccaccca	gaaatga
70	atgcagattg	agctgagcac	ctgcttcttc
	accaggagat	actacctggg	ggctgtggag
	ggggagctgc	ctgtggatgc	caggttcccc
	acctctgtgg	tgtacaagaa	gaccctgttt

	gccaagccca	ggccccctg	gatgggcctg	ctgggccccca	ccatccaggc	tgaggtgtat	300
5	gacactgtgg	tgatcaccct	gaagaacatg	gccagccacc	ctgtgagcct	gcatgctgtg	360
	ggggtgagct	actggaaggc	ctctgagggg	gctgagtatg	atgaccagac	cagccagagg	420
	gagaaggagg	atgacaaggt	gttccctggg	ggcagccaca	cctatgtgtg	gcaggtgctg	480
10	aaggagaatg	gccccatggc	ctctgacccc	ctgtgcctga	cctacagcta	cctgagccat	540
	gtggacctgg	tgaaggacct	gaactctggc	ctgattgggg	ccctgctggt	gtgcagggag	600
	ggcagcctgg	ccaaggagaa	gacccagacc	ctgcacaagt	tcatacctgct	gtttgctgtg	660
15	tttgatgagg	gcaagagctg	gcaactctgaa	accaagaaca	gcctgatgca	ggacagggat	720
	gctgcctctg	ccagggcctg	gccaagatg	cacactgtga	atggctatgt	gaacaggagc	780
20	ctgcctggcc	tgattggctg	ccacaggaag	tctgtgtact	ggcatgtgat	tggcatgggc	840
	accacccctg	aggtgcacag	catcttcctg	gagggccaca	ccttcctggt	caggaaccac	900
	aggcaggcca	gcctggagat	cagccccatc	accttcctga	ctgcccagac	cctgctgatg	960
25	gacctgggcc	agttcctgct	gttctgccac	atcagcagcc	accagcatga	tggcatggag	1020
	gcctatgtga	aggtggacag	ctgccctgag	gagccccagc	tgaggatgaa	gaacaatgag	1080
30	gaggctgagg	actatgatga	tgacctgact	gactctgaga	tggatgtggt	gaggtttgat	1140
	gatgacaaca	gccccagctt	catccagatc	aggtctgttg	ccaagaagca	ccccaagacc	1200
	tgggtgcact	acattgctgc	tgaggaggag	gactgggact	atgccccctt	ggtgctggcc	1260
35	cctgatgaca	ggagctacaa	gagccagtag	ctgaacaatg	gccccagag	gattggcagg	1320
	aagtacaaga	aggtcagggt	catggcctac	actgatgaaa	ccttcaagac	cagggaggcc	1380
40	atccagcatg	agtctggcat	cctgggcccc	ctgctgtatg	gggaggtggg	ggacaccctg	1440
	ctgatcatct	tcaagaacca	ggccagcagg	ccctacaaca	tctacccccca	tggcatcact	1500
	gatgtgaggc	ccctgtacag	caggaggctg	cccaaggggg	tgaagcacct	gaaggacttc	1560
45	cccatcctgc	ctggggagat	cttcaagtac	aagtggactg	tgactgtgga	ggatggcccc	1620
	accaagtctg	accccagggtg	cctgaccaga	tactacagca	gctttgtgaa	catggagagg	1680
50	gacctggcct	ctggcctgat	tggccccctg	ctgatctgct	acaaggagtc	tgtggaccag	1740
	aggggcaacc	agatcatgtc	tgacaagagg	aatgtgatcc	tgttctctgt	gtttgatgag	1800
	aacaggagct	ggtacctgac	tgagaacatc	cagagggttc	tgcccaaccc	tgctgggggtg	1860
55	cagctggagg	accctgagtt	ccaggccagc	aacatcatgc	acagcatcaa	tggctatgtg	1920
	tttgacagcc	tgcagctgtc	tgtgtgcctg	catgagggtg	cctactggta	catcctgagc	1980
60	attggggccc	agactgactt	cctgtctgtg	ttcttctctg	gctacacctt	caagcacaag	2040
	atggtgtatg	aggacaccct	gaccctgttc	cccttctctg	gggagactgt	gttcatgagc	2100
	atggagaacc	ctggcctgtg	gattctgggc	tgccacaact	ctgacttcag	gaacaggggc	2160
65	atgactgccc	tgctgaaagt	ctccagctgt	gacaagaaca	ctggggacta	ctatgaggac	2220
	agctatgagg	acatctctgc	ctacctgctg	agcaagaaca	atgccattga	gcccaggagc	2280
70	ttcagccaga	acagcaggca	ccccagcacc	aggcagaagc	agttcaatgc	caccaccatc	2340

	cctgagaatg	acatagagaa	gacagaccca	tggtttgccc	accggacccc	catgcccag	2400
5	atccagaatg	tgagcagctc	tgacctgctg	atgctgctga	ggcagagccc	caccccccat	2460
	ggcctgagcc	tgtctgacct	gcaggaggcc	aagtatgaaa	ccttctctga	tgacccagc	2520
	cctggggcca	ttgacagcaa	caacagcctg	tctgagatga	cccacttcag	gccccagctg	2580
10	caccactctg	gggacatggg	gttcacccct	gagtctggcc	tgcagctgag	gctgaatgag	2640
	aagctgggca	ccactgctgc	cactgagctg	aagaagctgg	acttcaaagt	ctccagcacc	2700
15	agcaacaacc	tgatcagcac	catccccctt	gacaacctgg	ctgctggcac	tgacaacacc	2760
	agcagcctgg	gccccccag	catgcctgtg	cactatgaca	gccagctgga	caccaccctg	2820
	tttggcaaga	agagcagccc	cctgactgag	tctggggggc	ccctgagcct	gtctgaggag	2880
20	aacaatgaca	gcaagctgct	ggagtctggc	ctgatgaaca	gccaggagag	cagctggggc	2940
	aagaatgtga	gcagcaggga	gatcaccagg	accaccctgc	agtctgacca	ggaggagatt	3000
25	gactatgatg	acaccatctc	tgtggagatg	aagaaggagg	actttgacat	ctacgacgag	3060
	gacgagaacc	agagccccag	gagcttccag	aagaagacca	ggcactactt	cattgctgct	3120
	gtggagaggc	tgtgggacta	tggcatgagc	agcagccccc	atgtgctgag	gaacagggcc	3180
30	cagtctggct	ctgtgcccc	gttcaagaag	gtggtgttcc	aggagtccac	tgatggcagc	3240
	ttcaccagc	ccctgtacag	aggggagctg	aatgagcacc	tgggcctgct	gggcccctac	3300
	atcagggctg	aggtggagga	caacatcatg	gtgaccttca	ggaaccaggc	cagcaggccc	3360
35	tacagcttct	acagcagcct	gatcagctat	gaggaggacc	agaggcaggg	ggctgagccc	3420
	aggaagaact	ttgtgaagcc	caatgaaacc	aagacctact	tctggaaggt	gcagcaccac	3480
40	atggccccc	ccaaggatga	gtttgactgc	aaggcctggg	cctacttctc	tgatgtggac	3540
	ctggagaagg	atgtgcactc	tggcctgatt	ggccccctgc	tgggtgtgcca	caccaacacc	3600
	ctgaaccctg	cccatggcag	gcaggtgact	gtgcaggagt	ttgccctgtt	cttcaccatc	3660
45	tttgatgaaa	ccaagagctg	gtacttcact	gagaacatgg	agaggaactg	cagggccccc	3720
	tgcaacatcc	agatggagga	ccccaccttc	aaggagaact	acaggttcca	tgccatcaat	3780
50	ggctacatca	tggacaccct	gcctggcctg	gtgatggccc	aggaccagag	gatcaggtgg	3840
	tacctgctga	gcatgggcag	caatgagaac	atccacagca	tccacttctc	tggccatgtg	3900
55	ttcactgtga	ggaagaagga	ggagtacaag	atggccctgt	acaacctgta	ccctgggggtg	3960
	tttgagactg	tggagatgct	gcccagcaag	gctggcatct	ggaggggtgga	gtgcctgatt	4020
	ggggagcacc	tgcctgctgg	catgagcacc	ctgttcctgg	tgtacagcaa	caagtgccag	4080
60	acccccctgg	gcatggcctc	tggccacatc	agggacttcc	agatcactgc	ctctggccag	4140
	tatggccagt	gggcccccaa	gctggccagg	ctgcactact	ctggcagcat	caatgcctgg	4200
	agcaccaagg	agcccttcag	ctggatcaag	gtggacctgc	tggcccccat	gatcatccat	4260
65	ggcatcaaga	cccagggggc	caggcagaag	ttcagcagcc	tgtacatcag	ccagttcatc	4320
	atcatgtaca	gcctggatgg	caagaagtgg	cagacctaca	ggggcaacag	cactggcacc	4380
70	ctgatggtgt	tctttggcaa	tgtggacagc	tctggcatca	agcacaacat	cttcaacccc	4440

	cccatcattg ccagatacat caggctgcac cccacccact acagcatcag gagcacctg	4500
5	aggatggagc tgatgggctg tgacctgaac agctgcagca tgcccctggg catggagagc	4560
	aaggccatct ctgatgccca gatcactgcc agcagctact tcaccaacat gtttgccacc	4620
	tggagcccca gcaaggccag gctgcacctg cagggcagga gcaatgcctg gaggccccag	4680
10	gtcaacaacc ccaaggagtg gctgcagggtg gacttccaga agaccatgaa ggtgactggg	4740
	gtgaccaccc aggggggtgaa gagcctgctg accagcatgt atgtgaagga gttcctgatc	4800
	agcagcagcc aggatggcca ccagtggacc ctgttcttcc agaatggcaa ggtgaagggtg	4860
15	ttccagggca accaggacag cttcacccct gtggtgaaca gcctggaccc ccccctgctg	4920
	accagatacc tgaggattca cccccagagc tgggtgcacc agattgccct gaggatggag	4980
20	gtgctgggct gtgaggccca ggacctgtac tga	5013
	<210> 17	
	<211> 873	
25	<212> ДНК	
	<213> Цитомегаловірус	
	<400> 17	
30	ccgcggagat ctcaatattg gccattagcc atattattca ttggttatat agcataaatc	60
	aatattggct attggccatt gcatacgttg tatctatatc ataatatgta catttatatt	120
	ggctcatgtc caatatgacc gccatgtttg cattgattat tgactagtta ttaatagtaa	180
35	tcaattacgg ggtcattagt tcatagccca tatatggagt tccgcgttac ataacttacg	240
	gtaaatggcc cgcctggctg accgcccac gacccccgcc cattgacgtc aataatgacg	300
	tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt ggagtattta	360
40	cggtaaactg cccacttggc agtacatcaa gtgtatcata tgccaagtcc gccccctatt	420
	gacgtcaatg acggtaaatg gccgcctgg cattatgccc agtacatgac cttacgggac	480
45	tttcctactt ggcagtacat ctacgtatta gtcacgcta ttaccatggg gatgcgggtt	540
	tggcagtaca ccaatgggcg tggatagcgg tttgactcac ggggatttcc aagtctccac	600
	cccattgacg tcaatgggag tttgttttgg caccaaaatc aacgggactt tccaaaatgt	660
50	cgtaataacc ccgccccgtt gacgcaaagtg ggcggtaggc gtgtacgggt ggaggtctat	720
	ataagcagag ctcgttttagt gaaccgtcag atcactagaa gctttattgc ggtagtttat	780
55	cacagttaaa ttgctaacgc agtcagtgt tctgacacaa cagtctcgaa cttaaagctgc	840
	agaagttggc cgtgaggcac tgggcaggct agc	873
60	<210> 18	
	<211> 67	
	<212> ДНК	
	<213> штучна послідовність	
65	<220>	
	<223> праймер	
	<400> 18	
70	tcgagatgtg gtctgagtta aaaatcagga gcaacgacgg aggtgaagga ccagacgcca	60

	acgaccc	67
5	<210> 19 <211> 67 <212> ДНК <213> штучна послідовність	
10	<220> <223> праймер	
	<400> 19 ccgggggctcg ttggcgtctg gtccttcacc tccgtcgttg ctcttgattt ttaactcaga	60
15	ccacatc	67
20	<210> 20 <211> 57 <212> ДНК <213> штучна послідовність	
25	<220> <223> праймер	
	<400> 20 ccgggggaaag ggggtgcaac acatccatat ccagccatct ctacctgttt atggaca	57
30	<210> 21 <211> 58 <212> ДНК <213> штучна послідовність	
35	<220> <223> праймер	
40	<400> 21 accctctgtc cataaacagg tagagatggc tggatatgga tgtgttgac ccctttcc	58
45	<210> 22 <211> 39 <212> ДНК <213> штучна послідовність	
50	<220> <223> праймер	
	<400> 22 gggtaggtg gttgctgatt ctctattca cccagtggg	39
55	<210> 23 <211> 39 <212> ДНК <213> штучна послідовність	
60	<220> <223> праймер	
65	<400> 23 gatccsact gggatgaatga gagaatcagc aaccaccta	39
70	<210> 24 <211> 34 <212> ДНК <213> штучна послідовність	

	<220>				
	<223>	праймер			
	<400>	24			
5		gagactcgag atgtggtctg agttaaaaat cagg			34
	<210>	25			
	<211>	62			
10	<212>	ДНК			
	<213>	штучна послідовність			
	<220>				
	<223>	праймер			
15	<400>	25			
		agaggtagac cagtacgagt cacgtttgcc cctatcacca tccctaacc tctgtcataa			60
		ас			62
20					
	<210>	26			
	<211>	1257			
	<212>	ДНК			
25	<213>	штучна послідовність			
	<220>				
	<223>	комплементарная цепь кодон-оптимизированного обедненного			срг
		трансгена А1АТ			
30	<400>	26			
		tacgggtcga gacacaggac cccgtaagac gacgaccgac cggacacgac agaccacgga			60
		cacagggacc gactcctggg agtcccccta cgacgggtct tttgtctgtg gaggggtgga			120
35		ctggtcctgg tgggggtgga gttgttctag tgggggttg accgtctcaa acggaagtcg			180
		gacatgtctg tcgaccgggt ggtctcgttg tcgtggttgt agaaaaagtc gggacacagg			240
40		taacggtgtc ggaacggta cgactcggac ccgtggttcc gactgtgggt actactctag			300
		gaccttccgg acttgaagtt ggactgtctc tagggactcc gggcttaggt actcccgaag			360
		gtccttgacg actcttggga cttggtcggt ctgtcggctg acgtcgactg ttgtccgtta			420
45		cccgacaagg acagactccc ggacttcgac cacctgttca aagaccttct acacttcttc			480
		gacatggtga gactccgga gtgtcacttg aaaccctgt gtcttctccg gttctttgtc			540
50		tagttactga tacacctttt cccgtgggtc ccgttctaac acctggaaca ctttctcgac			600
		ctgtccctgt gacacaaacg ggaacacttg atgtagaaga agttcccgtt caccctctcc			660
		gggaaacttc acttcctgtg actccttctc ctgaaggtag acctgggttca ctggtgtcac			720
55		ttccacgggt actacttctc tgaccctac aagttatagg tcgtgacgtt ctttgactcg			780
		tcgaccacg acgactactt catggaccg ttacgatgtc ggtataagaa agacggacta			840
60		ctcccgttcg acgtcgtgga ctttttactc gactgggtac tgtagtagtg gtttaaagac			900
		cttttactcc tgtcttctag acggtcggac gtagacgggt tcgactcgta gtgtccgtgt			960
		atactggact tcagacacga ccctgtcgac ctttagtggt tccacaagtc gttaccccgt			1020
65		ctggactcac ctactgtct ctttcggga gacttcgaca gggtccgaca cgtgttccgt			1080
		cacgactggt aactactctt cccgtgtctc cgacgacccc ggtacaaaga ctttcggtag			1140
70		gggtacaggt aggggggtct tcacttcaag ttgttcggga aacacaagga ctactaactc			1200

gtcttgtggt tctcggggga caagtacccg ttccaacact tgggggtgggt ctttact 1257

5 <210> 27
 <211> 419
 <212> PRT
 <213> штучна послідовність

10 <220>
 <223> пептид, колируемый кодон-оптимизированным обедненным CpG трансгеном
 ALAT

15 <400> 27

Ala Glu Asp Pro Gln Gly Asp Ala Ala Gln Lys Thr Asp Thr Ser His
 1 5 10 15

His Asp Gln Asp His Pro Thr Phe Ala Glu Asp Pro Gln Gly Asp Ala
 20 20 25 30

Ala Gln Lys Thr Asp Thr Ser His His Asp Gln Asp His Pro Thr Phe
 25 35 40 45

Asn Lys Ile Thr Pro Asn Leu Ala Glu Phe Ala Phe Ser Leu Tyr Arg
 30 50 55 60

Gln Leu Ala His Gln Ser Asn Ser Thr Asn Ile Phe Phe Ser Pro Val
 65 70 75 80

Ser Ile Ala Thr Ala Phe Ala Met Leu Ser Leu Gly Thr Lys Ala Asp
 85 90 95

Thr His Asp Glu Ile Leu Glu Gly Leu Asn Phe Asn Leu Thr Glu Ile
 100 105 110

Pro Glu Ala Gln Ile His Glu Gly Phe Gln Glu Leu Leu Arg Thr Leu
 115 120 125

Asn Gln Pro Asp Ser Gln Leu Gln Leu Thr Thr Gly Asn Gly Leu Phe
 130 135 140

Leu Ser Glu Gly Leu Lys Leu Val Asp Lys Phe Leu Glu Asp Val Lys
 145 150 155 160

Lys Leu Tyr His Ser Glu Ala Phe Thr Val Asn Phe Gly Asp Thr Glu
 165 170 175

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

Lys Ile Val Asp Leu Val Lys Glu Leu Asp Arg Asp Thr Val Phe Ala
 195 200 205

Leu Val Asn Tyr Ile Phe Phe Lys Gly Lys Trp Glu Arg Pro Phe Glu
 210 215 220

70

	Val	Lys	Asp	Thr	Glu	Glu	Glu	Asp	Phe	His	Val	Asp	Gln	Val	Thr	Thr	
	225					230					235					240	
5	Val	Lys	Val	Pro	Met	Met	Lys	Arg	Leu	Gly	Met	Phe	Asn	Ile	Gln	His	
					245					250					255		
10	Cys	Lys	Lys	Leu	Ser	Ser	Trp	Val	Leu	Leu	Met	Lys	Tyr	Leu	Gly	Asn	
				260					265					270			
15	Ala	Thr	Ala	Ile	Phe	Phe	Leu	Pro	Asp	Glu	Gly	Lys	Leu	Gln	His	Leu	
			275					280					285				
20	Glu	Asn	Glu	Leu	Thr	His	Asp	Ile	Ile	Thr	Lys	Phe	Leu	Glu	Asn	Glu	
		290					295					300					
25	Asp	Arg	Arg	Ser	Ala	Ser	Leu	His	Leu	Pro	Lys	Leu	Ser	Ile	Thr	Gly	
	305					310					315					320	
30	Thr	Tyr	Asp	Leu	Lys	Ser	Val	Leu	Gly	Gln	Leu	Gly	Ile	Thr	Lys	Val	
					325					330					335		
35	Phe	Ser	Asn	Gly	Ala	Asp	Leu	Ser	Gly	Val	Thr	Glu	Glu	Ala	Pro	Leu	
				340					345					350			
40	Lys	Leu	Ser	Lys	Ala	Val	His	Lys	Ala	Val	Leu	Thr	Ile	Asp	Glu	Lys	
			355					360					365				
45	Gly	Thr	Glu	Ala	Ala	Gly	Ala	Met	Phe	Leu	Glu	Ala	Ile	Pro	Met	Ser	
		370					375					380					
50	Ile	Pro	Pro	Glu	Val	Lys	Phe	Asn	Lys	Pro	Phe	Val	Phe	Leu	Met	Ile	
	385					390					395					400	
55	Glu	Gln	Asn	Thr	Lys	Ser	Pro	Leu	Phe	Met	Gly	Lys	Val	Val	Asn	Pro	
					405					410					415		
60	Thr	Gln	Lys														
65	<210>	28															
	<211>	5013															
	<212>	ДНК															
	<213>	штучна послідовність															
70	<220>																
	<223>	комплементарная	цепь	кодон-оптимизированного	обедненного	СрG											
		трансгена FVIII (N6)															
75	<400>	28															
	tacgtctaac	tcgactcgtg	gacgaagaag	gacacggacg	actccaagac	gaagagacgg										60	
80	tggtcctcta	tgatggaccc	ccgacacctc	gactcgaccc	tgatgtacgt	cagactggac										120	
85	cccctcgacg	gacacctacg	gtccaagggg	gggtctcacg	ggttctcgaa	ggggaagttg										180	
90	tggagacacc	acatgttctt	ctgggacaaa	cacctcaagt	gactgggtgga	caagttgtaa										240	

	cggttcgggt	ccgggggggac	ctacccggac	gacccggggt	ggtaggtccg	actccacata	300
5	ctgtgacacc	actagtggga	cttcttgtac	cggtcggttg	gacactcgga	cgtacgacac	360
	ccccactcga	tgaccttccg	gagactcccc	cgactcatac	tactgggtctg	gtcgggtctcc	420
	ctcttcctcc	tactgtttcca	caagggaccc	ccgtcggtgt	ggatacacac	cgtccacgac	480
10	ttcctcttac	cgggggtaccg	gagactgggg	gacacggact	ggatgtcgat	ggactcggtg	540
	cacctggacc	acttcctgga	cttgagaccg	gactaaccac	gggacgacca	cacgtccctc	600
15	ccgtcggacc	ggttcctctt	ctgggtcttg	gacgtgttca	agtaggacga	caaacgacac	660
	aaactactcc	cgttctcgac	cgtgagactt	tggttcttgt	cggactacgt	cctgtcccta	720
	cgacggagac	ggtcccggac	cggtttctac	gtgtgacact	taccgataca	cttgtcctcg	780
20	gacggaccgg	actaaccgac	ggtgtccttc	agacacatga	ccgtacacta	accgtaccgg	840
	tggtggggac	tccacgtgtc	gtagaaggac	ctcccgggtg	ggaaggacca	gtccttggtg	900
25	tccgtccggt	cggacctcta	gtcggggtag	tggaaggact	gacgggtctg	ggacgactac	960
	ctggacccgg	tcaaggacga	caagacggtg	tagtcgtcgg	tggtcgtact	accgtacctc	1020
	cggatacact	tccacctgtc	gacgggactc	ctcggggctg	actcctactt	cttgttactc	1080
30	ctccgactcc	tgatactact	actggactga	ctgagactct	acctacacca	ctccaaacta	1140
	ctactgttgt	cgggggtcgaa	gtaggtctag	tccagacacc	ggttcttcgt	ggggttcttg	1200
35	accacagtga	tgtaacgacg	actcctcctc	ctgaccctga	tacgggggga	ccacgaccgg	1260
	ggactactgt	cctcgatgtt	ctcggtcatt	gacttggtac	cgggggtctc	ctaaccgtcc	1320
	ttcatgttct	tccagtccaa	gtaccggatg	tgactacttt	ggaagtcttg	gtccctccgg	1380
40	taggtcgtac	tcagaccgta	ggacccgggg	gacgacatac	ccctccaccc	cctgtgggac	1440
	gactagtaga	agttcttggt	ccggtcgtcc	gggatgttgt	agatgggggt	accgtagtga	1500
45	ctacactccg	gggacatgtc	gtcctccgac	gggttcccc	acttcgtgga	cttcctgaag	1560
	gggtaggacg	gaccctcta	gaagttcatg	ttcacctgac	actgacacct	cctaccgggg	1620
	tggttcagac	tggggtccac	ggactggtct	atgatgtcgt	cgaaacactt	gtacctctcc	1680
50	ctggaccgga	gaccggacta	accgggggac	gactagacga	tgttcctcag	acacctggtc	1740
	tccccgttgg	tctagtacag	actgttctcc	ttacactagg	acaagagaca	caaactactc	1800
55	ttgtcctcga	ccatggactg	actcttgtag	gtctccaagg	acgggttggg	acgacccac	1860
	gtcgacctcc	tgggactcaa	ggtccggtcg	ttgtagtacg	tgctgtagtt	accgatacac	1920
	aaactgtcgg	acgtcgacag	acacacggac	gtactccacc	ggatgaccat	gtaggactcg	1980
60	taaccccggg	tctgactgaa	ggacagacac	aagaagagac	cgatgtggaa	gttcgtgttc	2040
	taccacatac	tcctgtggga	ctgggacaag	gggaagagac	ccctctgaca	caagtactcg	2100
65	tacctcttgg	gaccggacac	ctaagacccg	acggtgttga	gactgaagtc	cttgtccccg	2160
	tactgacggg	acgactttca	gaggtcgaca	ctgttcttgt	gaccctgat	gatactcctg	2220
	tcgatactcc	tgtagagacg	gatggacgac	tcgttcttgt	tacggtaact	cgggtcctcg	2280
70	aagtcggtct	tgtcgtccgt	ggggtcgtgg	tccgtcttcg	tcaagttacg	gtggtggtag	2340

	ggactcttac	tgtatctctt	ctgtctgggt	accaaacggg	tggcctgggg	gtacggggtc	2400
5	taggtcttac	actcgtcgag	actggacgac	tacgacgact	ccgtctcggg	gtggggggta	2460
	ccggactcgg	acagactgga	cgtcctccgg	ttcatacttt	ggaagagact	actggggctg	2520
	ggaccccgg	aactgtcgtt	gttgctcgac	agactctact	gggtgaagtc	cggggctcgac	2580
10	gtggtgagac	ccctgtacca	caagtgggga	ctcagaccgg	acgtcgactc	cgacttactc	2640
	ttcgacccgt	ggtgacgacg	gtgactcgac	ttcttcgacc	tgaagtttca	gaggtcgtgg	2700
	tcgttggttg	actagtcgtg	gtaggggaga	ctgttgacc	gacgaccgtg	actggtgtgg	2760
15	tcgtcggacc	cgggggggtc	gtacggacac	gtgatactgt	cggtcgacct	gtggtgggac	2820
	aaaccgttct	tctcgtcggg	ggactgactc	agacccccgg	gggactcggg	cagactcctc	2880
20	ttgttactgt	cgttcgacga	cctcagaccg	gactacttgt	cggtcctctc	gtcgaccccg	2940
	ttcttacact	cgtcgtccct	ctagtgggtc	tgggtgggacg	tcagactggg	cctcctctaa	3000
	ctgatactac	tgtggtagag	acacctctac	ttcttcctcc	tgaactgta	gatgctgctc	3060
25	ctgctcttgg	tctcgggggtc	ctcgaagggtc	ttcttctggt	ccgtgatgaa	gtaacgacga	3120
	cacctctccg	acaccctgat	accgtactcg	tcgtcggggg	tacacgactc	cttgtcccgg	3180
30	gtcagaccga	gacacggggg	caagtctctc	caccacaagg	tcctcaagtg	actaccgtcg	3240
	aagtgggtcg	gggacatgtc	tcccctcgac	ttactcgtgg	acccggacga	cccggggatg	3300
	tagtcccgc	tccacctcct	gttgtagtac	cactggaagt	ccttggtccg	gtcgtccggg	3360
35	atgtcgaaga	tgctcgtcga	ctagtcgata	ctcctcctgg	tctccgtccc	ccgactcggg	3420
	tccttcttga	aacacttcgg	gttactttgg	ttctggatga	agaccttcca	cgctcgtggtg	3480
40	taccgggggt	ggttcctact	caaactgacg	ttccggaccc	ggatgaagag	actacacctg	3540
	gacctcttcc	tacacgtgag	accggactaa	ccgggggacg	accacacggg	gtggttgtgg	3600
	gacttgggac	gggtaccgtc	cgtccactga	cacgtcctca	aacgggacaa	gaagtggtag	3660
45	aaactacttt	ggttctcgac	catgaagtga	ctcttgtagc	tctccttgac	gtcccggggg	3720
	acgttgtagg	tctacctcct	ggggtggaag	ttcctcttga	tgtccaagggt	acggtagtta	3780
50	ccgatgtagt	acctgtggga	cggaccggac	cactaccggg	tcctggtctc	ctagtccacc	3840
	atggacgact	cgtaccctgc	gttactcttg	taggtgtcgt	aggatgaagag	accggtacac	3900
	aagtgacact	ccttcttcct	cctcatgttc	taccgggaca	tgttgacat	gggacccac	3960
55	aaactctgac	acctctacga	cgggtcgttc	cgaccgtaga	cctccacact	cacggactaa	4020
	cccctcgtgg	acgtacgacc	gtactcgtgg	gacaaggacc	acatgtcggt	gttcacgggtc	4080
60	tgggggggac	cgtaccggag	accggtgtag	tccctgaagg	tctagtacg	gagaccgggtc	4140
	ataccgggtca	cccgggggtt	cgaccggtcc	gacgtgatga	gaccgtcgta	gttacggacc	4200
	tcgtgggttc	tcgggaagtc	gacctagtcc	cacctggacg	accgggggta	ctagtaggta	4260
65	ccgtagtctt	gggtccccc	gtccgtcttc	aagtcgtcgg	acatgtagtc	ggtcaagtag	4320
	tagtacatgt	cggacctacc	gttcttcacc	gtctggatgt	ccccgttggtc	gtgaccgtgg	4380
70	gactaccaca	agaaaccgtt	acacctgtcg	agaccgtagt	tcgtgttgta	gaagttgggg	4440

5 gggtagtaac ggtctatgta gtccgacgtg ggggtgggtga tgctgtagtc ctcgtgggac 4500
 tcctacctcg actacccgac actggacttg tcgacgtcgt acgggggaccc gtacctctcg 4560
 10 ttccggtaga gactacgggt ctagtgacgg tcgtcgatga agtggttgta caaacgggtg 4620
 acctcgggggt cgttcgggtc cgacgtggac gtcccgtcct cgttacggac ctccgggggtc 4680
 15 cagttgttgg ggttcctcac cgacgtccac ctgaaggctt tctggtactt ccaactgaccc 4740
 cactggtggg tccccactt ctcggacgac tggtcgtaca tacacttcct caaggactag 4800
 tcgtcgtcgg tcctaccggt ggtcacctgg gacaagaagg tcttaccgtt ccacttccac 4860
 20 aaggtcccgt tggtcctgtc gaagtgggga caccacttgt cggacctggg gggggacgac 4920
 tggctctatgg actcctaagt gggggtctcg acccacgtgg tctaacggga ctctacctc 4980
 25 cacgacccga cactccgggt cctggacatg act 5013

25 <210> 29
 <211> 1670
 <212> PRT
 <213> штучна послідовність
 <220>
 30 <223> пептид, який кодується кодоно-оптимізованим збідненим CpG трансгеном
 FVIII (N6)
 <400> 29

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

	145		150		155		160									
5	Lys	Glu	Asn	Gly	Pro 165	Met	Ala	Ser	Asp	Pro 170	Leu	Cys	Leu	Thr	Tyr 175	Ser
10	Tyr	Leu	Ser	His 180	Val	Asp	Leu	Val	Lys 185	Asp	Leu	Asn	Ser	Gly 190	Leu	Ile
15	Gly	Ala	Leu 195	Leu	Val	Cys	Arg	Glu 200	Gly	Ser	Leu	Ala	Lys 205	Glu	Lys	Thr
20	Gln	Thr 210	Leu	His	Lys	Phe	Ile 215	Leu	Leu	Phe	Ala	Val 220	Phe	Asp	Glu	Gly
25	Lys 225	Ser	Trp	His	Ser	Glu 230	Thr	Lys	Asn	Ser	Leu 235	Met	Gln	Asp	Arg	Asp 240
30	Ala	Ala	Ser	Ala	Arg 245	Ala	Trp	Pro	Lys	Met 250	His	Thr	Val	Asn	Gly 255	Tyr
35	Val	Asn	Arg	Ser 260	Leu	Pro	Gly	Leu	Ile 265	Gly	Cys	His	Arg	Lys 270	Ser	Val
40	Tyr	Trp	His 275	Val	Ile	Gly	Met	Gly 280	Thr	Thr	Pro	Glu	Val 285	His	Ser	Ile
45	Phe	Leu 290	Glu	Gly	His	Thr	Phe 295	Leu	Val	Arg	Asn	His 300	Arg	Gln	Ala	Ser
50	Leu 305	Glu	Ile	Ser	Pro	Ile 310	Thr	Phe	Leu	Thr	Ala 315	Gln	Thr	Leu	Leu	Met 320
55	Asp	Leu	Gly	Gln	Phe 325	Leu	Leu	Phe	Cys	His 330	Ile	Ser	Ser	His	Gln 335	His
60	Asp	Gly	Met	Glu 340	Ala	Tyr	Val	Lys	Val 345	Asp	Ser	Cys	Pro	Glu 350	Glu	Pro
65	Gln	Leu	Arg 355	Met	Lys	Asn	Asn	Glu 360	Glu	Ala	Glu	Asp	Tyr 365	Asp	Asp	Asp
70	Leu	Thr 370	Asp	Ser	Glu	Met	Asp 375	Val	Val	Arg	Phe	Asp 380	Asp	Asp	Asn	Ser
	Pro 385	Ser	Phe	Ile	Gln	Ile 390	Arg	Ser	Val	Ala	Lys 395	Lys	His	Pro	Lys	Thr 400
	Trp	Val	His	Tyr	Ile 405	Ala	Ala	Glu	Glu	Glu 410	Asp	Trp	Asp	Tyr	Ala 415	Pro
	Leu	Val	Leu	Ala 420	Pro	Asp	Asp	Arg	Ser 425	Tyr	Lys	Ser	Gln	Tyr 430	Leu	Asn

	Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met
			435					440					445			
5	Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu
		450					455					460				
10	Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu
	465					470					475					480
15	Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro
					485					490					495	
20	His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys
				500					505					510		
25	Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe
			515					520					525			
30	Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp
		530					535					540				
35	Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg
	545					550					555					560
40	Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu
					565					570					575	
45	Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val
				580					585					590		
50	Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu
			595					600					605			
55	Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp
		610					615					620				
60	Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val
	625					630					635					640
65	Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp
					645					650					655	
70	Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe
				660					665					670		
75	Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr
			675					680					685			
80	Leu	Phe	Pro	Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro
		690					695					700				
85	Gly	Leu	Trp	Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly

	705				710					715					720	
5	Met	Thr	Ala	Leu	Leu 725	Lys	Val	Ser	Ser	Cys 730	Asp	Lys	Asn	Thr	Gly 735	Asp
10	Tyr	Tyr	Glu	Asp 740	Ser	Tyr	Glu	Asp	Ile 745	Ser	Ala	Tyr	Leu	Leu 750	Ser	Lys
15	Asn	Asn	Ala 755	Ile	Glu	Pro	Arg	Ser 760	Phe	Ser	Gln	Asn	Ser 765	Arg	His	Pro
20	Ser	Thr 770	Arg	Gln	Lys	Gln	Phe 775	Asn	Ala	Thr	Thr	Ile 780	Pro	Glu	Asn	Asp
25	Ile 785	Glu	Lys	Thr	Asp	Pro 790	Trp	Phe	Ala	His	Arg 795	Thr	Pro	Met	Pro	Lys 800
30	Ile	Gln	Asn	Val	Ser 805	Ser	Ser	Ser	Asp	Leu	Leu 810	Met	Leu	Leu	Arg	Gln 815
35	Pro	Thr	Pro	His 820	Gly	Leu	Ser	Leu	Ser 825	Asp	Leu	Gln	Glu	Ala 830	Lys	Tyr
40	Glu	Thr	Phe 835	Ser	Asp	Asp	Pro	Ser 840	Pro	Gly	Ala	Ile	Asp 845	Ser	Asn	Asn
45	Ser	Leu 850	Ser	Glu	Met	Thr	His 855	Phe	Arg	Pro	Gln	Leu 860	His	His	Ser	Gly
50	Asp 865	Met	Val	Phe	Thr	Pro 870	Glu	Ser	Gly	Leu	Gln 875	Leu	Arg	Leu	Asn	Glu 880
55	Lys	Leu	Gly	Thr	Thr 885	Ala	Ala	Thr	Glu	Leu 890	Lys	Lys	Leu	Asp	Phe 895	Lys
60	Val	Ser	Ser	Thr 900	Ser	Asn	Asn	Leu	Ile 905	Ser	Thr	Ile	Pro	Ser 910	Asp	Asn
65	Leu	Ala	Ala 915	Gly	Thr	Asp	Asn	Thr 920	Ser	Ser	Leu	Gly	Pro 925	Pro	Ser	Met
70	Pro	Val 930	His	Tyr	Asp	Ser	Gln 935	Leu	Asp	Thr	Thr	Leu 940	Phe	Gly	Lys	Lys
75	Ser 945	Ser	Pro	Leu	Thr	Glu 950	Ser	Gly	Gly	Pro	Leu 955	Ser	Leu	Ser	Glu	Glu 960
80	Asn	Asn	Asp	Ser	Lys 965	Leu	Leu	Glu	Ser	Gly 970	Leu	Met	Asn	Ser	Gln 975	Glu
85	Ser	Ser	Trp	Gly 980	Lys	Asn	Val	Ser	Ser 985	Arg	Glu	Ile	Thr	Arg 990	Thr	Thr

	Leu	Gln	Ser	Asp	Gln	Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val
			995					1000					1005			
5	Glu	Met	Lys	Lys	Glu	Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	
		1010					1015					1020				
10	Gln	Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	
		1025					1030					1035				
15	Ala	Ala	Val	Glu	Arg	Leu	Trp	Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	
		1040					1045					1050				
20	His	Val	Leu	Arg	Asn	Arg	Ala	Gln	Ser	Gly	Ser	Val	Pro	Gln	Phe	
		1055					1060					1065				
25	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr	Asp	Gly	Ser	Phe	Thr	Gln	
		1070					1075					1080				
30	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu	His	Leu	Gly	Leu	Leu	Gly	
		1085					1090					1095				
35	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser	Ser	Leu	Ile	
		1115					1120					1125				
40	Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg	Lys	Asn	
		1130					1135					1140				
45	Phe	Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val	Gln	
		1145					1150					1155				
50	His	His	Met	Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp	
		1160					1165					1170				
55	Ala	Tyr	Phe	Ser	Asp	Val	Asp	Leu	Glu	Lys	Asp	Val	His	Ser	Gly	
		1175					1180					1185				
60	Leu	Ile	Gly	Pro	Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	
		1190					1195					1200				
65	Ala	His	Gly	Arg	Gln	Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	
		1205					1210					1215				
70	Thr	Ile	Phe	Asp	Glu	Thr	Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	
		1220					1225					1230				
75	Glu	Arg	Asn	Cys	Arg	Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	
		1235					1240					1245				
80	Thr	Phe	Lys	Glu	Asn	Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	

	1250		1255		1260								
5	Met	Asp	Thr	Leu	Pro	Gly	Leu	Val	Met	Ala	Gln	Asp	Gln Arg Ile
	1265						1270					1275	
10	Arg	Trp	Tyr	Leu	Leu	Ser	Met	Gly	Ser	Asn	Glu	Asn	Ile His Ser
	1280						1285					1290	
15	Ile	His	Phe	Ser	Gly	His	Val	Phe	Thr	Val	Arg	Lys	Lys Glu Glu
	1295						1300					1305	
20	Tyr	Lys	Met	Ala	Leu	Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe Glu Thr
	1310						1315					1320	
25	Val	Glu	Met	Leu	Pro	Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val Glu Cys
	1325						1330					1335	
30	Leu	Ile	Gly	Glu	His	Leu	His	Ala	Gly	Met	Ser	Thr	Leu Phe Leu
	1340						1345					1350	
35	Val	Tyr	Ser	Asn	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala Ser Gly
	1355						1360					1365	
40	His	Ile	Arg	Asp	Phe	Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr Gly Gln
	1370						1375					1380	
45	Trp	Ala	Pro	Lys	Leu	Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser Ile Asn
	1385						1390					1395	
50	Ala	Trp	Ser	Thr	Lys	Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val Asp Leu
	1400						1405					1410	
55	Leu	Ala	Pro	Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly Ala Arg
	1415						1420					1425	
60	Gln	Lys	Phe	Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile Met Tyr
	1430						1435					1440	
65	Ser	Leu	Asp	Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn Ser Thr
	1445						1450					1455	
70	Gly	Thr	Leu	Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser Gly Ile
	1460						1465					1470	
75	Lys	His	Asn	Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr Ile Arg
	1475						1480					1485	
80	Leu	His	Pro	Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg Met Glu
	1490						1495					1500	
85	Leu	Met	Gly	Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu Gly Met
	1505						1510					1515	

	Glu	Ser	Lys	Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr
	1520						1525					1530			
5	Phe	Thr	Asn	Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu
	1535						1540					1545			
10	His	Leu	Gln	Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn
	1550						1555					1560			
15	Pro	Lys	Glu	Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val
	1565						1570					1575			
20	Thr	Gly	Val	Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met
	1580						1585					1590			
	Tyr	Val	Lys	Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly	His	Gln
	1595						1600					1605			
25	Trp	Thr	Leu	Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly
	1610						1615					1620			
30	Asn	Gln	Asp	Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro
	1625						1630					1635			
35	Leu	Leu	Thr	Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His
	1640						1645					1650			
40	Gln	Ile	Ala	Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp
	1655						1660					1665			
	Leu	Tyr													
	1670														
45	<210> 30														
	<211> 4425														
	<212> ДНК														
	<213> штучна послідовність														
50	<220>														
	<223> кодон-оптимізований збіднений CpG трансген FVIII (V3)														
	<400> 30														
55	atgcagattg	agctgagcac	ctgcttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc									60
	accaggagat	actacctggg	ggctgtggag	ctgagctggg	actacatgca	gtctgacctg									120
	ggggagctgc	ctgtggatgc	caggttcccc	cccagagtgc	ccaagagctt	ccccttcaac									180
60	acctctgtgg	tgtacaagaa	gaccctgttt	gtggagttca	ctgaccacct	gttcaacatt									240
	gccaagccca	ggccccctg	gatgggcctg	ctgggccccca	ccatccaggc	tgagggtgat									300
65	gacactgtgg	tgatcaccct	gaagaacatg	gccagccacc	ctgtgagcct	gcatgctgtg									360
	ggggtgagct	actggaaggc	ctctgagggg	gctgagtatg	atgaccagac	cagccagagg									420
	gagaaggagg	atgacaaggt	gttccttggg	ggcagccaca	cctatgtgtg	gcagggtgctg									480
70															

	aaggagaatg	gccccatggc	ctctgacccc	ctgtgcctga	cctacagcta	cctgagccat	540
	gtggacctgg	tgaaggacct	gaactctggc	ctgattgggg	ccctgctggt	gtgcagggag	600
5	ggcagcctgg	ccaaggagaa	gacccagacc	ctgcacaagt	tcatacctgct	gtttgctgtg	660
	tttgatgagg	gcaagagctg	gcactctgaa	accaagaaca	gcctgatgca	ggacagggat	720
	gctgcctctg	ccagggcctg	gccaagatg	cacactgtga	atggctatgt	gaacaggagc	780
10	ctgcctggcc	tgattggctg	ccacaggaag	tctgtgtact	ggcatgtgat	tggcatgggc	840
	accacccctg	aggtgcacag	catcttcctg	gagggccaca	ccttcctggt	caggaaccac	900
15	aggcaggcca	gcctggagat	cagccccatc	accttcctga	ctgcccagac	cctgctgatg	960
	gacctgggcc	agttcctgct	gttctgccac	atcagcagcc	accagcatga	tggcatggag	1020
	gcctatgtga	aggtggacag	ctgccctgag	gagccccagc	tgaggatgaa	gaacaatgag	1080
20	gaggctgagg	actatgatga	tgacctgact	gactctgaga	tggatgtggt	gaggtttgat	1140
	gatgacaaca	gccccagctt	catccagatc	aggtctgttg	ccaagaagca	ccccaagacc	1200
25	tgggtgcact	acattgctgc	tgaggaggag	gactgggact	atgccccctt	ggtgctggcc	1260
	cctgatgaca	ggagctacaa	gagccagtac	ctgaacaatg	gccccagag	gattggcagg	1320
	aagtacaaga	aggtcaggtt	catggcctac	actgatgaaa	ccttcaagac	cagggaggcc	1380
30	atccagcatg	agtctggcat	cctgggcccc	ctgctgtatg	gggaggtggg	ggacaccctg	1440
	ctgatcatct	tcaagaacca	ggccagcagg	ccctacaaca	tctaccccca	tggcatcact	1500
35	gatgtgaggc	ccctgtacag	caggaggctg	cccaaggggg	tgaagcacct	gaaggacttc	1560
	cccatcctgc	ctggggagat	cttcaagtac	aagtggactg	tgactgtgga	ggatggcccc	1620
	accaagtctg	accccaggtg	cctgaccaga	tactacagca	gctttgtgaa	catggagagg	1680
40	gacctggcct	ctggcctgat	tggccccctg	ctgatctgct	acaaggagtc	tgtggaccag	1740
	aggggcaacc	agatcatgtc	tgacaagagg	aatgtgatcc	tgttctctgt	gtttgatgag	1800
45	aacaggagct	ggtacctgac	tgagaacatc	cagaggttcc	tgcccaaccc	tgctgggggtg	1860
	cagctggagg	accctgagtt	ccaggccagc	aacatcatgc	acagcatcaa	tggctatgtg	1920
	tttgacagcc	tgacagctgc	tgtgtgcctg	catgagggtg	cctactggta	catcctgagc	1980
50	attggggccc	agactgactt	cctgtctgtg	ttcttctctg	gctacacctt	caagcacaag	2040
	atgggtgtatg	aggacaccct	gaccctgttc	cccttctctg	gggagactgt	gttcatgagc	2100
55	atggagaacc	ctggcctgtg	gattctgggc	tgccacaact	ctgacttcag	gaacaggggc	2160
	atgactgccc	tgctgaaagt	ctccagctgt	gacaagaaca	ctggggacta	ctatgaggac	2220
	agctatgagg	acatctctgc	ctacctgctg	agcaagaaca	atgccattga	gcccaggagc	2280
60	ttcagccaga	atgccactaa	tgtgtctaac	aacagcaaca	ccagcaatga	cagcaatgtg	2340
	tctccccag	tgctgaagag	gcaccagagg	gagatcacca	ggaccaccct	gcagtctgac	2400
65	caggaggaga	ttgactatga	tgacaccatc	tctgtggaga	tgaagaagga	ggactttgac	2460
	atctacgacg	aggacgagaa	ccagagcccc	aggagcttcc	agaagaagac	caggcactac	2520
70	ttcattgctg	ctgtggagag	gctgtgggac	tatggcatga	gcagcagccc	ccatgtgctg	2580

	aggaacaggg	cccagtctgg	ctctgtgccc	cagttcaaga	aggtggtggt	ccaggagttc	2640
	actgatggca	gcttcaccca	gccctgtac	agaggggagc	tgaatgagca	cctgggcctg	2700
5	ctggggcccct	acatcagggc	tgaggtggag	gacaacatca	tggtgacctt	caggaaccag	2760
	gccagcaggc	cctacagctt	ctacagcagc	ctgatcagct	atgaggagga	ccagaggcag	2820
	ggggctgagc	ccaggaagaa	ctttgtgaag	cccaatgaaa	ccaagaccta	cttctggaag	2880
10	gtgcagcacc	acatggcccc	caccaaggat	gagtttgact	gcaaggcctg	ggcctacttc	2940
	tctgatgtgg	acctggagaa	ggatgtgcac	tctggcctga	ttggccccct	gctggtgtgc	3000
15	cacaccaaca	ccctgaacct	tgcccatggc	aggcaggtga	ctgtgcagga	gtttgccctg	3060
	ttcttcacca	tctttgatga	aaccaagagc	tggtacttca	ctgagaacat	ggagaggaac	3120
	tgcagggccc	cctgcaacat	ccagatggag	gacccacact	tcaaggagaa	ctacaggttc	3180
20	catgccatca	atggctacat	catggacacc	ctgcctggcc	tggtgatggc	ccaggaccag	3240
	aggatcaggt	ggtacctgct	gagcatgggc	agcaatgaga	acatccacag	catccacttc	3300
25	tctggccatg	tgttcactgt	gaggaagaag	gaggagtaca	agatggccct	gtacaacctg	3360
	taccctgggg	tgtttgagac	tgtggagatg	ctgcccagca	aggctggcat	ctggaggggtg	3420
	gagtgcctga	ttggggagca	cctgcatgct	ggcatgagca	ccctgttcct	ggtgtacagc	3480
30	aacaagtgcc	agacccccct	gggcatggcc	tctggccaca	tcagggactt	ccagatcact	3540
	gcctctggcc	agtatggcca	gtgggcccc	aagctggcca	ggctgcacta	ctctggcagc	3600
35	atcaatgcct	ggagcaccaa	ggagcccttc	agctggatca	aggtggacct	gctggcccc	3660
	atgatcatcc	atggcatcaa	gacccagggg	gccaggcaga	agttcagcag	cctgtacatc	3720
	agccagttca	tcatcatgta	cagcctggat	ggcaagaagt	ggcagaccta	caggggcaac	3780
40	agcactggca	ccctgatggg	gttctttggc	aatgtggaca	gctctggcat	caagcacaac	3840
	atcttcaacc	cccccatcat	tgccagatac	atcaggctgc	acccacacca	ctacagcatc	3900
45	aggagcacc	tgaggatgga	gctgatgggc	tgtgacctga	acagctgcag	catgcccctg	3960
	ggcatggaga	gcaaggccat	ctctgatgcc	cagatcactg	ccagcagcta	cttcaccaac	4020
	atgtttgcc	cctggagccc	cagcaaggcc	aggctgcacc	tgcaaggcag	gagcaatgcc	4080
50	tggaggcccc	aggtcaaaa	ccccaggag	tggtgcagg	tggaactcca	gaagaccatg	4140
	aaggtgactg	gggtgaccac	ccagggggtg	aagagcctgc	tgaccagcat	gtatgtgaag	4200
55	gagttcctga	tcagcagcag	ccaggatggc	caccagtgga	ccctgttctt	ccagaatggc	4260
	aaggtgaagg	tgttccaggg	caaccaggac	agcttcaccc	ctgtggtgaa	cagcctggac	4320
	ccccccctgc	tgaccagata	cctgaggatt	cacccccaga	gctgggtgca	ccagattgcc	4380
60	ctgaggatgg	aggtgctggg	ctgtgaggcc	caggacctgt	actga		4425

65 <210> 31
 <211> 4425
 <212> ДНК
 <213> штучна послідовність
 <220>

<223> комплементарний ланцюг кодон-оптимізованого збіденого cpg трансгена
FVIII
(V3)

5	<400> 31	tacgtctaac tcgactcgtg gacgaagaag gacacggacg actccaagac gaagagacgg	60
		tggtcctcta tgatggaccc ccgacacctc gactcgaccc tgatgtacgt cagactggac	120
10		cccctcgacg gacacctacg gtccaagggg ggggtctcacg ggttctcgaa ggggaagttg	180
		tggagacacc acatgttctt ctgggacaaa cacctcaagt gactgggtgga caagttgtaa	240
		cggttcgggt ccgggggggac ctacccggac gaccgggggt ggtaggtccg actccacata	300
15		ctgtgacacc actagtggga cttcttgtac cggtcgggtg gacactcgga cgtacgacac	360
		ccccactcga tgaccttccg gagactcccc cgactcatac tactgggtctg gtcgggtctcc	420
20		ctcttcctcc tactgttcca caagggaccc ccgtcgggtgt ggatacacac cgtccacgac	480
		ttcctcttac cgggggtaccg gagactgggg gacacggact ggatgtcgat ggactcggta	540
		cacctggacc acttcctgga cttgagaccg gactaaccac gggacgacca cacgtccctc	600
25		ccgtcggacc ggttcctctt ctgggtctg gacgtgttca agtaggacga caaacgacac	660
		aaactactcc cgttctcgac cgtgagactt tggttcttgt cggactacgt cctgtcccta	720
30		cgacggagac ggtcccgac cgggttctac gtgtgacact taccgataca cttgtcctcg	780
		gacggaccgg actaaccgac ggtgtccttc agacacatga ccgtacacta accgtaccgg	840
		tggtggggac tccacgtgtc gtagaaggac ctcccgggtgt ggaaggacca gtccttggtg	900
35		tccgtccggt cggacctcta gtcggggtag tggaaggact gacgggtctg ggacgactac	960
		ctggacccgg tcaaggacga caagacggtg tagtcgtcgg tggtcgtact accgtacctc	1020
40		cggatacact tccacctgtc gacgggactc ctcggggtcg actcctactt cttgttactc	1080
		ctccgactcc tgatactact actggactga ctgagactct acctacacca ctccaaacta	1140
		ctactgttgt cggggtcgaa gtaggtctag tccagacacc ggttcttcgt ggggttctgg	1200
45		accacgtga tgtaacgacg actcctcctc ctgaccctga tacgggggga ccacgaccgg	1260
		ggactactgt cctcgatgtt ctcggtcacg gacttggttac cgggggtctc ctaaccgtcc	1320
50		ttcatgttct tccagtcca gtaccgatg tgactacttt ggaagtctctg gtccctccgg	1380
		taggtcgtac tcagaccgta ggacccgggg gacgacatac ccctccaccc cctgtgggac	1440
		gactagtaga agttcttggt ccggtcgtcc gggatgttgt agatgggggt accgtagtga	1500
55		ctacactccg gggacatgtc gtcctccgac gggttcccc acttcgtgga cttcctgaag	1560
		gggtaggacg gaccctcta gaagttcatg ttcacctgac actgacacct cctaccgggg	1620
60		tggttcagac tgggggtccac ggactggtct atgatgtcgt cgaaacactt gtacctctcc	1680
		ctggaccgga gaccggacta accgggggac gactagacga tgttcctcag acacctggtc	1740
		tccccgttgg tctagtacag actgttctcc ttacactagg acaagagaca caaactactc	1800
65		ttgtcctcga ccatggactg actctttagt gtctccaagg acgggttggg acgacccac	1860
		gtcgacctcc tgggactcaa ggtccggtcg ttgtagtacg tgtcgtagtt accgatacac	1920
70		aaactgtcgg acgtcgacag acacacggac gtactccacc ggatgaccat gtaggactcg	1980

	taaccccggg	tctgactgaa	ggacagacac	aagaagagac	cgatgtggaa	gttcgtgttc	2040
5	taccacatac	tcctgtggga	ctgggacaag	gggaagagac	ccctctgaca	caagtactcg	2100
	tacctcttgg	gaccggacac	ctaagacccg	acggtgttga	gactgaagtc	cttgtccccg	2160
	tactgacggg	acgactttca	gaggtcgaca	ctgttcttgt	gaccctgat	gatactcctg	2220
10	tcgatactcc	tgtagagacg	gatggacgac	tcgttcttgt	tacggtaact	cgggtcctcg	2280
	aagtcggtct	tacggtgatt	acacagattg	ttgtcgttgt	ggtcgttact	gtcgttacac	2340
	agaggggggtc	acgacttctc	cgtggtctcc	ctctagtggg	cctggtgagg	cgtcagactg	2400
15	gtcctcctct	aactgatact	actgtggtag	agacacctct	acttcttcct	cctgaaactg	2460
	tagatgctgc	tcctgctctt	ggtctcgggg	tcctcgaagg	tcttcttctg	gtccgtgatg	2520
20	aagtaacgac	gacacctctc	cgacaccctg	ataccgtact	cgtcgtcggg	ggtacacgac	2580
	tccttgtccc	gggtcagacc	gagacacggg	gtcaagttct	tccaccacaa	ggtcctcaag	2640
	tgactaccgt	cgaagtgggt	cggggacatg	tctcccctcg	acttactcgt	ggacccggac	2700
25	gacccgggga	tgtagtcccg	actccacctc	ctgttgtagt	accactggaa	gtccttggtc	2760
	cggtcgtccg	ggatgtcgaa	gatgtcgtcg	gactagtcga	tactcctcct	ggtctccgtc	2820
30	ccccgactcg	ggtccttctt	gaaacacttc	gggttacttt	ggttctggat	gaagaccttc	2880
	cacgtcgtgg	tgtaccgggg	gtggttccta	ctcaaactga	cgttccggac	ccgatgaag	2940
	agactacacc	tggacctctt	cctacacgtg	agaccggact	aaccggggga	cgaccacacg	3000
35	gtgtgggttgt	gggacttggg	acgggtaccg	tccgtccact	gacacgtcct	caaacgggac	3060
	aagaagtggg	agaaactact	ttggttctcg	accatgaagt	gactcttgta	cctctccttg	3120
40	acgtcccggg	ggacgttgta	ggtctacctc	ctggggtgga	agttcctctt	gatgtccaag	3180
	gtacggtagt	taccgatgta	gtacctgtgg	gacggaccgg	accactaccg	ggtcctggtc	3240
	tcctagtcca	ccatggacga	ctcgtaccgg	tcgttactct	tgtaggtgtc	gtaggtgaag	3300
45	agaccggtac	acaagtgaca	ctccttcttc	ctcctcatgt	tctaccggga	catgttggtg	3360
	atgggacccc	acaaactctg	acacctctac	gacgggtcgt	tccgaccgta	gacctccac	3420
50	ctcacggact	aaccctcgt	ggacgtacga	ccgtactcgt	gggacaagga	ccacatgtcg	3480
	ttgttcacgg	tctggggggg	cccgtaccgg	agaccggtgt	agtccttgaa	ggtctagtga	3540
	cggagaccgg	tcataccggt	caccgggggg	ttcgaccggt	ccgacgtgat	gagaccgtcg	3600
55	tagttacgga	cctcgtgggt	cctcggaag	tcgacctagt	tccacctgga	cgaccggggg	3660
	tactagtagg	taccgtagtt	ctgggtcccc	cgggtccgtct	tcaagtcgtc	ggacatgtag	3720
60	tcgggtcaagt	agtagtacat	gtcggacctc	ccgttcttca	ccgtctggat	gtccccgttg	3780
	tcgtgaccgt	gggactacca	caagaaaccg	ttacacctgt	cgagaccgta	gttcgtgttg	3840
	tagaagttagg	gggggtagta	acgggtctatg	tagtccgacg	tggggtggtg	gatgtcgtag	3900
65	tcctcgtggg	actcctacct	cgactaccgg	acactggact	tgtcgacgtc	gtacggggac	3960
	ccgtacctct	cgttccggta	gagactacgg	gtctagtgtg	ggtcgtcgat	gaagtgggtg	4020
70	tacaaacggg	ggacctcggg	gtcgttccgg	tccgacgtgg	acgtcccgtc	ctcgttacgg	4080

acctccgggg tccagttggt ggggttcctc accgacgtcc acctgaaggt cttctggtac 4140
 5 ttccactgac cccactggtg ggtcccccac ttctcggacg actggtcgta catacacttc 4200
 ctcaaggact agtcgtcgtc ggtcctaccg gtggtcacct gggacaagaa ggtcttaccg 4260
 ttccacttcc acaagggtccc gttggtcctg tcgaagtggg gacaccactt gtcggacctg 4320
 10 ggggggggacg actggtctat ggactcctaa gtgggggtct cgaccacgt ggtctaacgg 4380
 gactcctacc tccacgaccc gacactccgg gtcctggaca tgact 4425

15 <210> 32
 <211> 1474
 <212> PRT
 <213> штучна послідовність

20 <220>
 <223> пептид, який кодується кодон-оптимізованим збідненим CpG трансгеном FVIII (V3)

<400> 32

25 Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
 1 5 10 15

30 Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser
 20 25 30

35 Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
 35 40 45

40 Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
 50 55 60

45 Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
 65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
 85 90 95

50 Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
 100 105 110

55 His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
 115 120 125

60 Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
 130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
 145 150 155 160

65 Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
 165 170 175

70 Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile

	180		185		190												
5	Gly	Ala	Leu 195	Leu	Val	Cys	Arg	Glu 200	Gly	Ser	Leu	Ala	Lys 205	Glu	Lys	Thr	
10	Gln	Thr 210	Leu	His	Lys	Phe	Ile 215	Leu	Leu	Phe	Ala	Val 220	Phe	Asp	Glu	Gly	
15	Lys 225	Ser	Trp	His	Ser	Glu 230	Thr	Lys	Asn	Ser	Leu 235	Met	Gln	Asp	Arg	Asp 240	
20	Ala	Ala	Ser	Ala	Arg 245	Ala	Trp	Pro	Lys	Met 250	His	Thr	Val	Asn	Gly 255	Tyr	
25	Val	Asn	Arg	Ser 260	Leu	Pro	Gly	Leu	Ile 265	Gly	Cys	His	Arg	Lys 270	Ser	Val	
30	Tyr	Trp	His 275	Val	Ile	Gly	Met	Gly 280	Thr	Thr	Pro	Glu	Val 285	His	Ser	Ile	
35	Phe	Leu 290	Glu	Gly	His	Thr	Phe 295	Leu	Val	Arg	Asn	His 300	Arg	Gln	Ala	Ser	
40	Leu 305	Glu	Ile	Ser	Pro	Ile 310	Thr	Phe	Leu	Thr	Ala 315	Gln	Thr	Leu	Leu	Met 320	
45	Asp	Leu	Gly	Gln	Phe 325	Leu	Leu	Phe	Cys	His 330	Ile	Ser	Ser	His	Gln 335	His	
50	Asp	Gly	Met	Glu 340	Ala	Tyr	Val	Lys	Val 345	Asp	Ser	Cys	Pro	Glu 350	Glu	Pro	
55	Gln	Leu	Arg 355	Met	Lys	Asn	Asn	Glu 360	Glu	Ala	Glu	Asp	Tyr 365	Asp	Asp	Asp	
60	Leu	Thr 370	Asp	Ser	Glu	Met	Asp 375	Val	Val	Arg	Phe	Asp 380	Asp	Asp	Asn	Ser	
65	Pro 385	Ser	Phe	Ile	Gln	Ile 390	Arg	Ser	Val	Ala	Lys 395	Lys	His	Pro	Lys	Thr 400	
70	Trp	Val	His	Tyr	Ile 405	Ala	Ala	Glu	Glu	Glu 410	Asp	Trp	Asp	Tyr	Ala 415	Pro	
	Leu	Val	Leu	Ala 420	Pro	Asp	Asp	Arg	Ser 425	Tyr	Lys	Ser	Gln	Tyr 430	Leu	Asn	
	Asn	Gly	Pro 435	Gln	Arg	Ile	Gly	Arg 440	Lys	Tyr	Lys	Lys	Val 445	Arg	Phe	Met	
	Ala	Tyr 450	Thr	Asp	Glu	Thr	Phe 455	Lys	Thr	Arg	Glu	Ala 460	Ile	Gln	His	Glu	

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480
 5 Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495
 10 His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510
 15 Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525
 20 Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540
 25 Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560
 30 Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575
 35 Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605
 40 Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620
 45 Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640
 50 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655
 55 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670
 60 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685
 65 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 70 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 65 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735
 70 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys

	740		745		750												
5	Asn	Asn	Ala 755	Ile	Glu	Pro	Arg	Ser 760	Phe	Ser	Gln	Asn	Ala 765	Thr	Asn	Val	
10	Ser	Asn 770	Asn	Ser	Asn	Thr	Ser 775	Asn	Asp	Ser	Asn	Val 780	Ser	Pro	Pro	Val	
15	Leu 785	Lys	Arg	His	Gln	Arg 790	Glu	Ile	Thr	Arg	Thr 795	Thr	Leu	Gln	Ser	Asp 800	
20	Gln	Glu	Glu	Ile	Asp 805	Tyr	Asp	Asp	Thr	Ile 810	Ser	Val	Glu	Met	Lys 815	Lys	
25	Glu	Asp	Phe	Asp 820	Ile	Tyr	Asp	Glu	Asp 825	Glu	Asn	Gln	Ser	Pro 830	Arg	Ser	
30	Phe	Gln	Lys 835	Lys	Thr	Arg	His	Tyr 840	Phe	Ile	Ala	Ala	Val 845	Glu	Arg	Leu	
35	Trp	Asp 850	Tyr	Gly	Met	Ser	Ser 855	Ser	Pro	His	Val	Leu 860	Arg	Asn	Arg	Ala	
40	Gln 865	Ser	Gly	Ser	Val	Pro 870	Gln	Phe	Lys	Lys	Val 875	Val	Phe	Gln	Glu	Phe 880	
45	Thr	Asp	Gly	Ser	Phe 885	Thr	Gln	Pro	Leu	Tyr 890	Arg	Gly	Glu	Leu	Asn 895	Glu	
50	His	Leu	Gly	Leu 900	Leu	Gly	Pro	Tyr	Ile 905	Arg	Ala	Glu	Val	Glu 910	Asp	Asn	
55	Ile	Met	Val 915	Thr	Phe	Arg	Asn	Gln 920	Ala	Ser	Arg	Pro	Tyr 925	Ser	Phe	Tyr	
60	Ser	Ser 930	Leu	Ile	Ser	Tyr	Glu 935	Glu	Asp	Gln	Arg	Gln 940	Gly	Ala	Glu	Pro	
65	Arg 945	Lys	Asn	Phe	Val	Lys 950	Pro	Asn	Glu	Thr	Lys 955	Thr	Tyr	Phe	Trp	Lys 960	
70	Val	Gln	His	His	Met 965	Ala	Pro	Thr	Lys	Asp 970	Glu	Phe	Asp	Cys	Lys 975	Ala	
75	Trp	Ala	Tyr	Phe 980	Ser	Asp	Val	Asp	Leu 985	Glu	Lys	Asp	Val	His 990	Ser	Gly	
80	Leu	Ile	Gly 995	Pro	Leu	Leu	Val	Cys 1000	His	Thr	Asn	Thr	Leu 1005	Asn	Pro	Ala	
85	His	Gly 1010	Arg	Gln	Val	Thr	Val 1015	Gln	Glu	Phe	Ala	Leu 1020	Phe	Phe	Thr		

	Ile	Phe 1025	Asp	Glu	Thr	Lys	Ser 1030	Trp	Tyr	Phe	Thr	Glu 1035	Asn	Met	Glu
5	Arg	Asn 1040	Cys	Arg	Ala	Pro	Cys 1045	Asn	Ile	Gln	Met	Glu 1050	Asp	Pro	Thr
10	Phe	Lys 1055	Glu	Asn	Tyr	Arg	Phe 1060	His	Ala	Ile	Asn	Gly 1065	Tyr	Ile	Met
15	Asp	Thr 1070	Leu	Pro	Gly	Leu	Val 1075	Met	Ala	Gln	Asp	Gln 1080	Arg	Ile	Arg
20	Trp	Tyr 1085	Leu	Leu	Ser	Met	Gly 1090	Ser	Asn	Glu	Asn	Ile 1095	His	Ser	Ile
25	His	Phe 1100	Ser	Gly	His	Val	Phe 1105	Thr	Val	Arg	Lys	Lys 1110	Glu	Glu	Tyr
	Lys	Met 1115	Ala	Leu	Tyr	Asn	Leu 1120	Tyr	Pro	Gly	Val	Phe 1125	Glu	Thr	Val
30	Glu	Met 1130	Leu	Pro	Ser	Lys	Ala 1135	Gly	Ile	Trp	Arg	Val 1140	Glu	Cys	Leu
35	Ile	Gly 1145	Glu	His	Leu	His	Ala 1150	Gly	Met	Ser	Thr	Leu 1155	Phe	Leu	Val
40	Tyr	Ser 1160	Asn	Lys	Cys	Gln	Thr 1165	Pro	Leu	Gly	Met	Ala 1170	Ser	Gly	His
	Ile	Arg 1175	Asp	Phe	Gln	Ile	Thr 1180	Ala	Ser	Gly	Gln	Tyr 1185	Gly	Gln	Trp
45	Ala	Pro 1190	Lys	Leu	Ala	Arg	Leu 1195	His	Tyr	Ser	Gly	Ser 1200	Ile	Asn	Ala
50	Trp	Ser 1205	Thr	Lys	Glu	Pro	Phe 1210	Ser	Trp	Ile	Lys	Val 1215	Asp	Leu	Leu
55	Ala	Pro 1220	Met	Ile	Ile	His	Gly 1225	Ile	Lys	Thr	Gln	Gly 1230	Ala	Arg	Gln
60	Lys	Phe 1235	Ser	Ser	Leu	Tyr	Ile 1240	Ser	Gln	Phe	Ile	Ile 1245	Met	Tyr	Ser
	Leu	Asp 1250	Gly	Lys	Lys	Trp	Gln 1255	Thr	Tyr	Arg	Gly	Asn 1260	Ser	Thr	Gly
65	Thr	Leu 1265	Met	Val	Phe	Phe	Gly 1270	Asn	Val	Asp	Ser	Ser 1275	Gly	Ile	Lys
70	His	Asn	Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu

UA 122127 C2

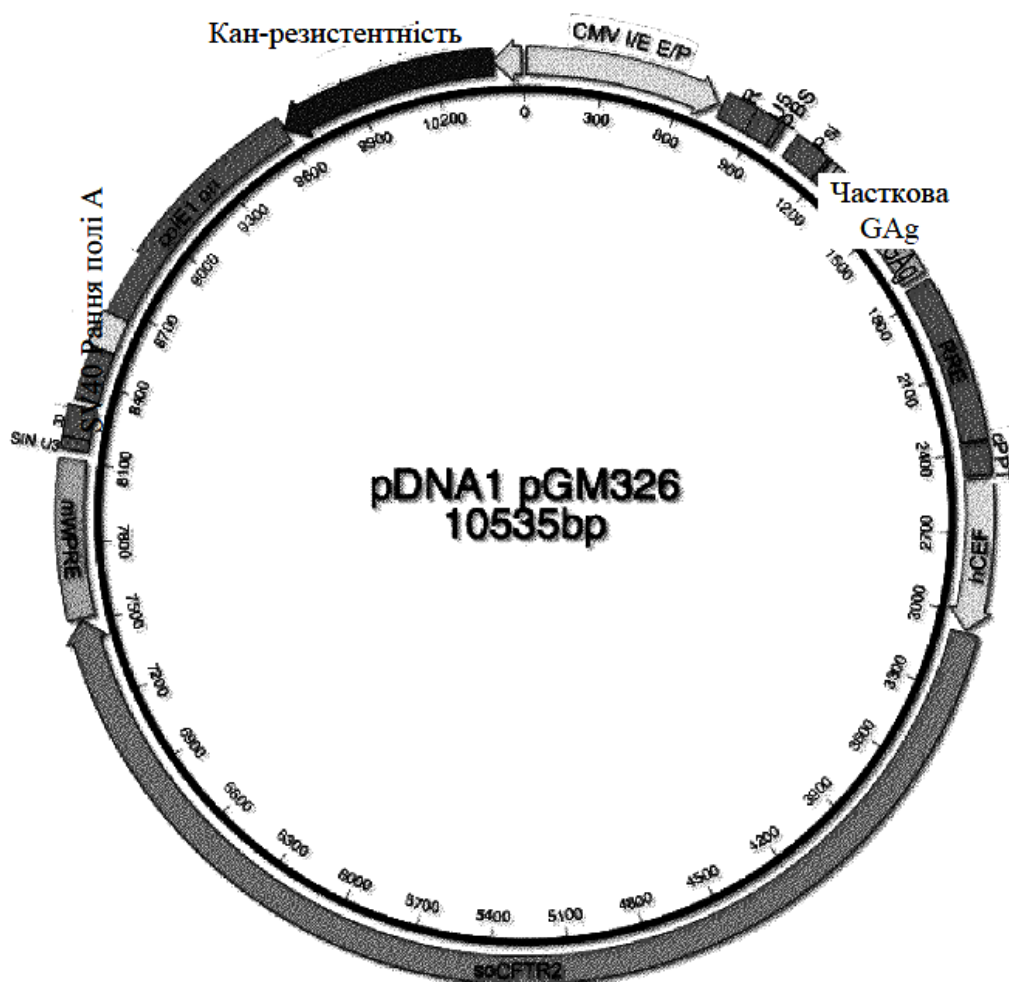
	1280	1285	1290	
5	His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu 1295 1300 1305			
10	Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu 1310 1315 1320			
15	Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe 1325 1330 1335			
20	Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His 1340 1345 1350			
25	Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro 1355 1360 1365			
30	Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr 1370 1375 1380			
35	Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr 1385 1390 1395			
40	Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp 1400 1405 1410			
45	Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn 1415 1420 1425			
50	Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu 1430 1435 1440			
55	Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln 1445 1450 1455			
60	Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu 1460 1465 1470			
65	Tyr			
70	<210> 33 <211> 873 <212> ДНК <213> Цитомегаловірус			
75	<400> 33 ggcgccctcta gagttataac cggtaatcgg tataataagt AACCAATATA tcgtatttag			60
80	TTATAACCga taaccggtaa CGTATGcaac atagatatag TATTATACat gtaaataaaa			120
85	cCGAGtacag GTtacttg cgGTACAacc GTAActAata actgatcaat AATTAtCAtt			180
90	aGTTaatgcc CCagtaatCa agtatcgggt ATAtAcCTCa aggCGcaatg tATTGaAtgc			240
95	CATttAccgg gcGGaCCgac TGgcGGgTTg CTgGGGgcGg gTAActGCag ttATTactgc			300

	atacaaggggt atcattgctg ttatccctga aaggtaactg cagttaccca cctcataaat	360
5	gccatttgac ggggtgaaccg tcatgtagtt cacatagtat acgggttcagg cgggggataa	420
	ctgcagttac tgccatttac cgggctgacc gtaatacggg tcatgtactg gaatgcctg	480
	aaaggatgaa ccgtcatgta gatgcataat cagtagcgat aatggtagca ctacgcaaaa	540
10	accgtcatgt gggtacccgc acctatcgcc aaactgagtg cccctaaagg ttcagaggtg	600
	gggtaactgc agttaccctc aaacaaaacc gtggttttag ttgccctgaa aggttttaca	660
15	gcattattgg ggcggggcaa ctgcgtttac ccgccatccg cacatgccac cctccagata	720
	tattcgtctc gagcaaatca cttggcagtc tagtgatctt cgaaataacg ccatcaaata	780
	gtgtcaattt aacgattgctg tcagtcacga agactgtgtt gtcagagctt gaattcgacg	840
20	tcttcaacca gcactccgtg acccgctcca tcg	873

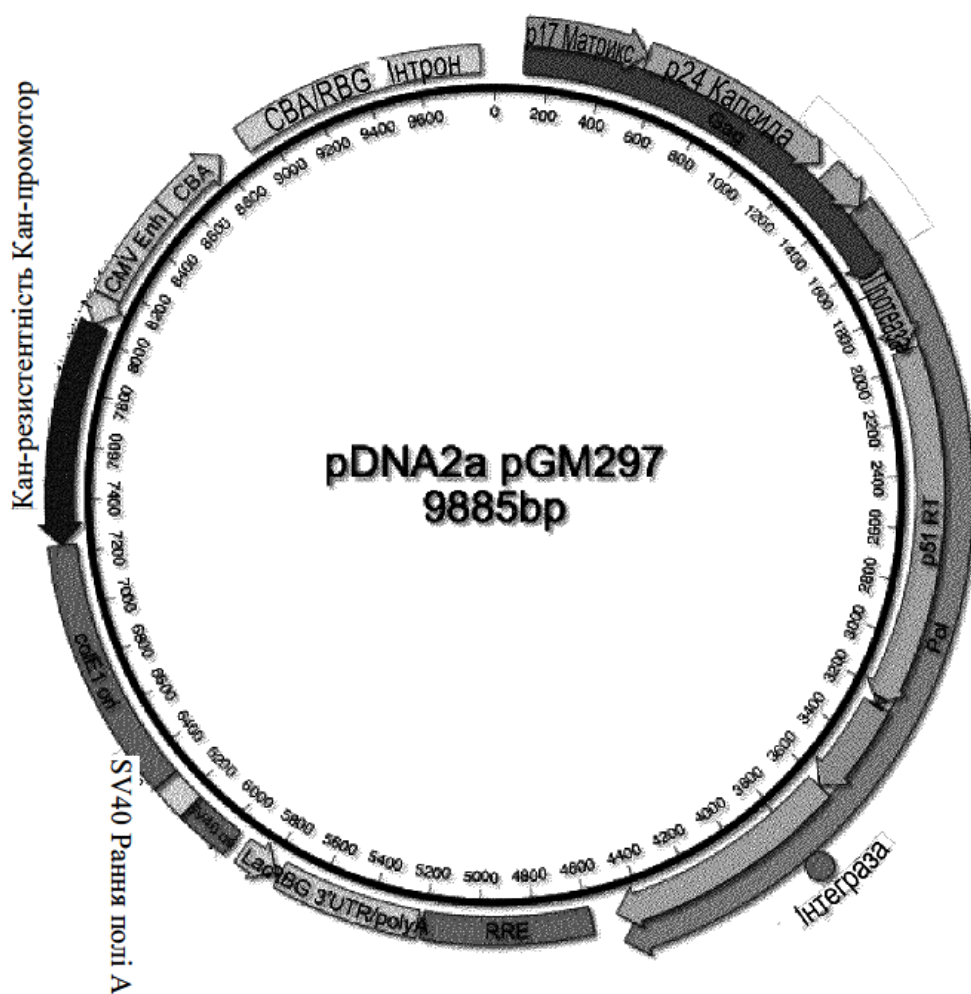
ФОРМУЛА ВИНАХОДУ

1. Лентивірусний вектор, псевдотипований гемаглютинін-нейрамінідазою (HN) і гібридними (F) білками з респіраторного параміксовірусу, причому вказаний лентивірусний вектор містить гібридний енхансер людського CMV/промотор EF1a (hCEF) і трансген, і причому в лентивірусному векторі відсутній будь-який інтрон, розташований між вказаним промотором і вказаним трансгеном.
2. Лентивірусний вектор за п. 1, де лентивірусний вектор вибирають із групи, яка складається з вектора вірусу імунодефіциту людини (HIV), вектора вірусу імунодефіциту мавп (SIV), вектора вірусу імунодефіциту котятих (FIV), вектора вірусу інфекційної анемії коней (EIAV) і вектора вірусу Вісна/Маєді.
3. Лентивірусний вектор за п. 1 або 2, де лентивірусний вектор являє собою вектор SIV.
4. Лентивірусний вектор за будь-яким з пп. 1-3, де респіраторний параміксовірус являє собою вірус Сендай.
5. Лентивірусний вектор за будь-яким з пп. 1-4, де трансген вибраний з:
 - (a) секретованого терапевтичного білка, у випадку необхідності альфа-1-антитрипсину (A1AT), фактора VIII, поверхнево-активного білка В (SFTPВ), фактора VII, фактора IX, фактора X, фактора XI, фактора фон Вілебранда, гранулоцитарно-макрофагального колонієстимулюючого фактора (GM-CSF) або моноклонального антитіла до інфекційного агента; або
 - (b) *CFTR*, *DNAH5*, *DNAH11*, *DNAI1* або *DNAI2*.
6. Лентивірусний вектор за будь-яким з пп. 1-5, де трансген кодує *CFTR*.
7. Лентивірусний вектор за будь-яким з пп. 1-5, де трансген кодує *A1AT*.
8. Лентивірусний вектор за будь-яким з пп. 1-5, де трансген кодує *FVIII*.
9. Спосіб одержання лентивірусу, визначеного у будь-якому з пп. 1-8, який включає наступні стадії:
 - (a) вирощування клітин у суспензії;
 - (b) трансфекція клітин з використанням однієї або декількох плазмід;
 - (c) додавання нуклеази;
 - (d) збирання лентивірусу;
 - (e) додавання трипсину; і
 - (f) очищення.
10. Спосіб за п. 9, де стадії (a)-(f) виконують послідовно.
11. Спосіб за п. 9 або 10, де клітини являють собою клітини HEK293T або клітини 293T/17.
12. Спосіб за пп. 9-11, де додавання нуклеази виконують на стадії до збирання.
13. Спосіб за пп. 9-12, де додавання трипсину виконують на стадії після збирання.
14. Спосіб за пп. 9-13, де стадія очищення включає стадію хроматографії.
15. Спосіб лікування захворювання, що включає введення лентивірусного вектора за будь-яким з пп. 1-8 пацієнту.
16. Спосіб за п. 15, де захворювання являє собою захворювання легень, вибране з: муковісцидозу (МВ), первинної циліарної дискінезії (ПЦД), недостатності поверхнево-активного білка В (ПАБ-В), недостатності альфа-1-антитрипсину (A1AT), легеневого альвеолярного протеїнозу (ЛАП) і хронічного обструктивного захворювання легень (ХОЗЛ).

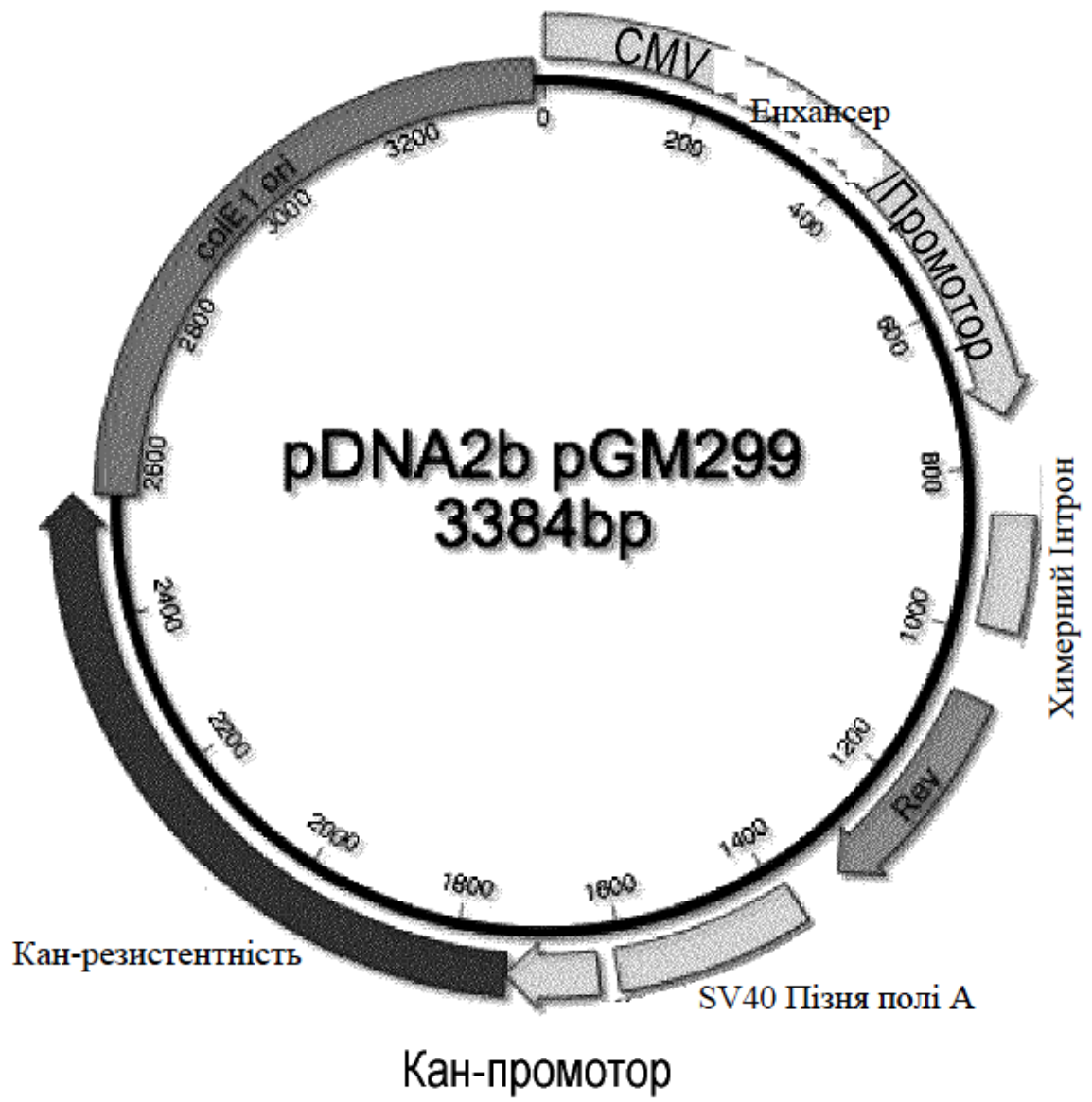
17. Спосіб за п. 15, де трансген являє собою фактор VIII, і захворювання являє собою гемофілію.
18. Лентівірусний вектор за будь-яким з пп. 1-6 для застосування в лікуванні МВ.
19. Лентівірусний вектор за будь-яким з пп. 1-5 або 7 для застосування в лікуванні недостатності A1AT.
20. Лентівірусний вектор за будь-яким одним з пп. 1-5 або 8 для застосування в лікуванні гемофілії.
21. Клітина-хазяїн, яка містить вектор за будь-яким з пп. 1-8.
22. Композиція, яка містить вектор за будь-яким з пп. 1-8, і фармацевтично прийнятний носій.



Фіг. 1А



Фіг. 1В



Фіг. 1С

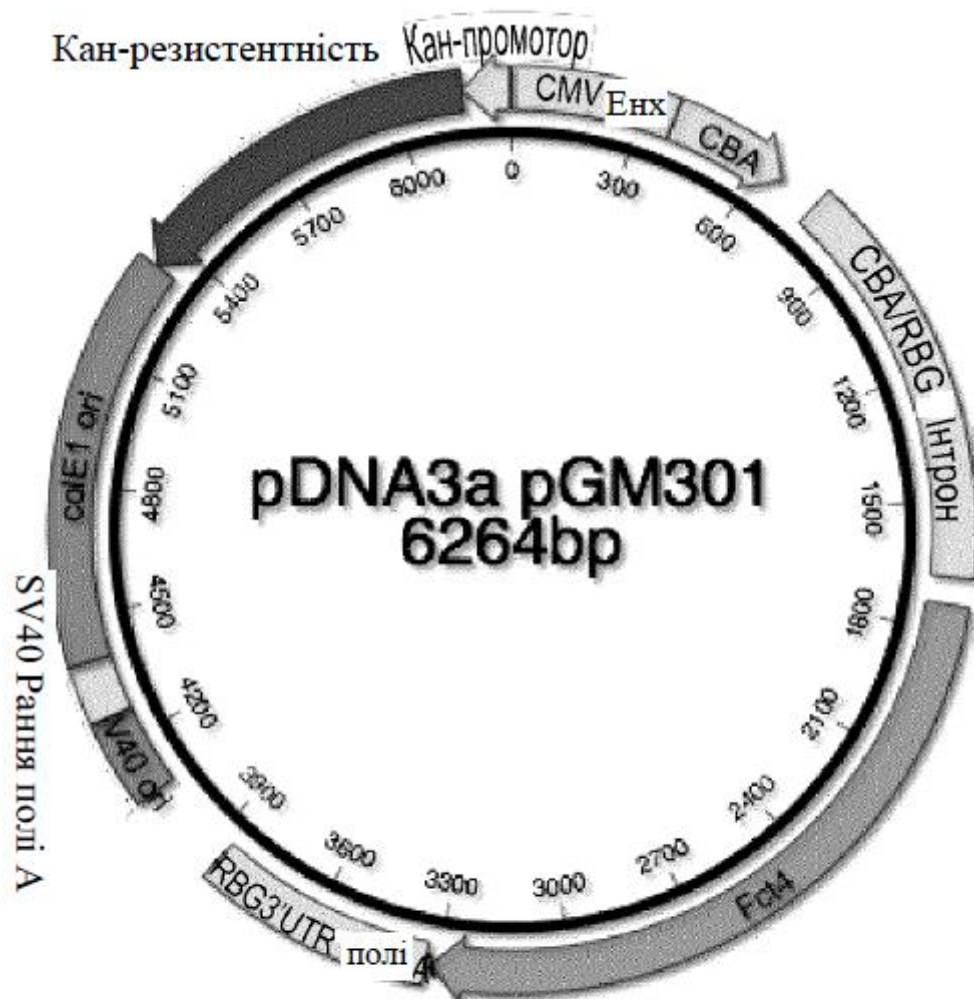
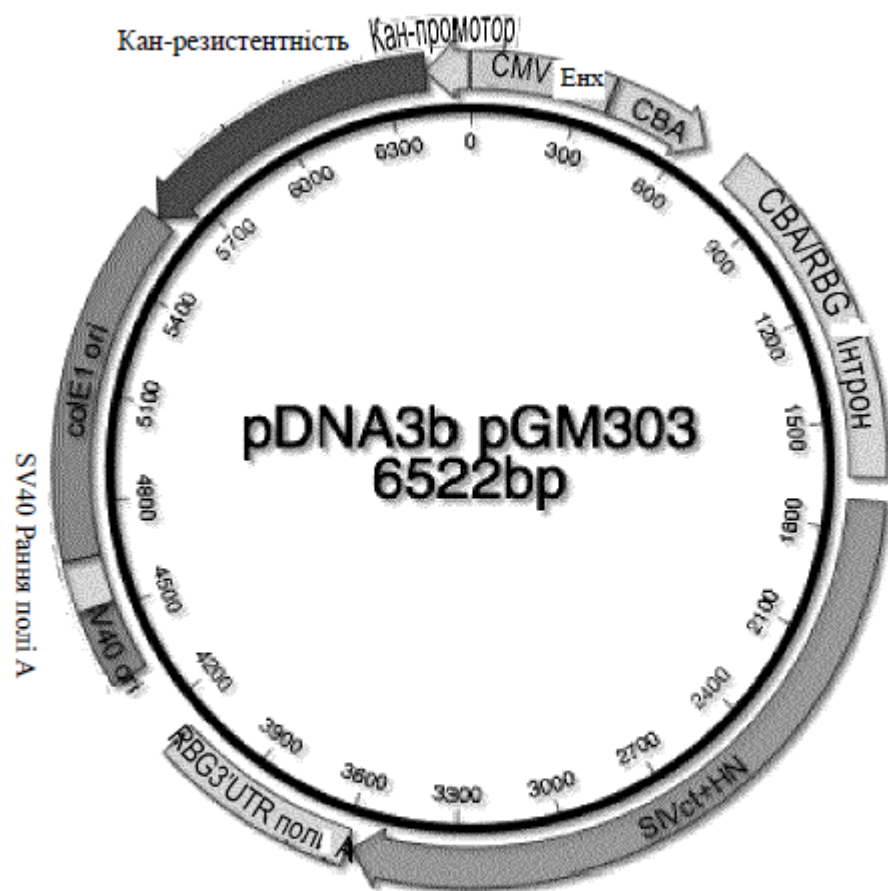
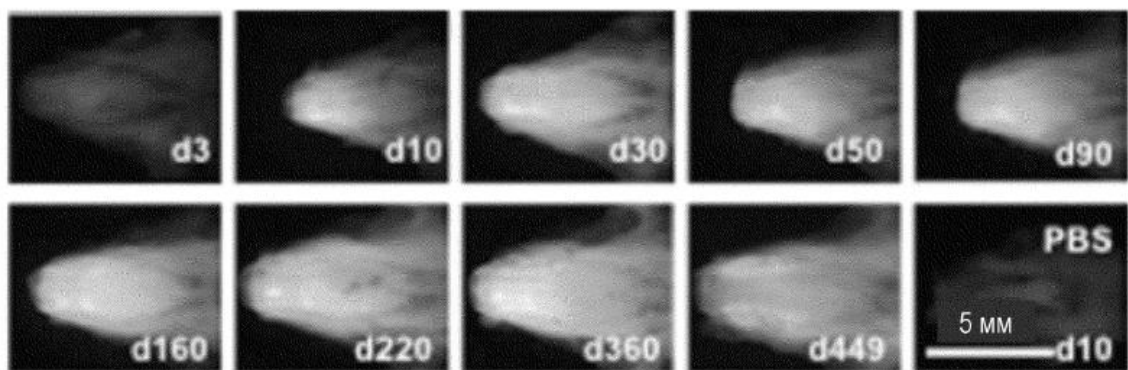


Fig. 1D



Фиг. 1Е



Фиг. 2

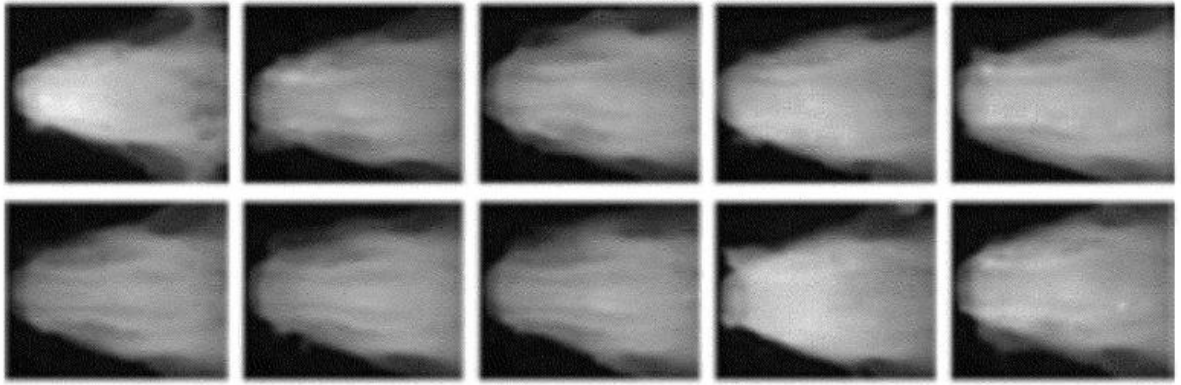


Fig. 3

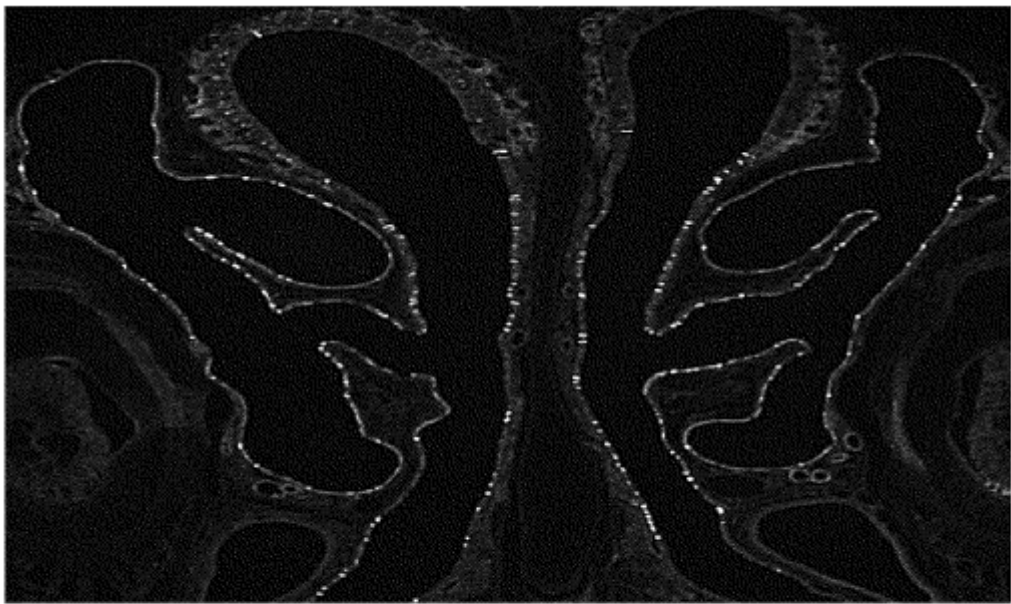


Fig. 4

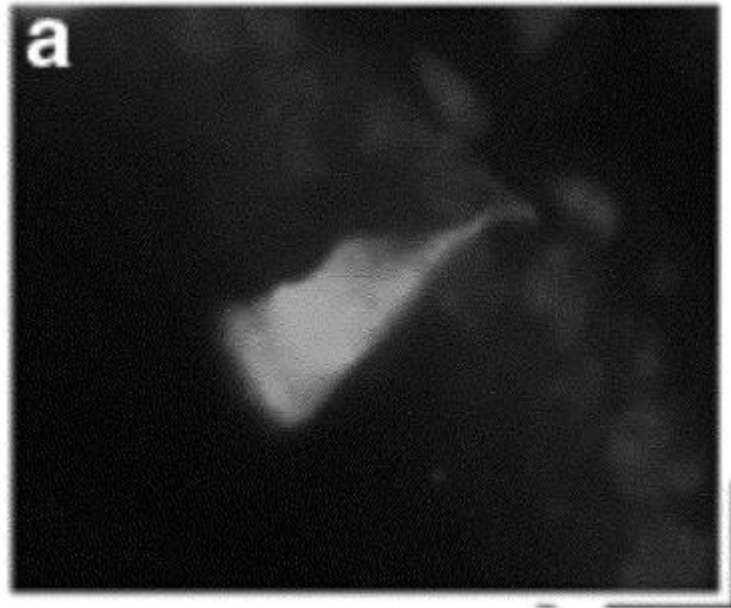


Fig. 5

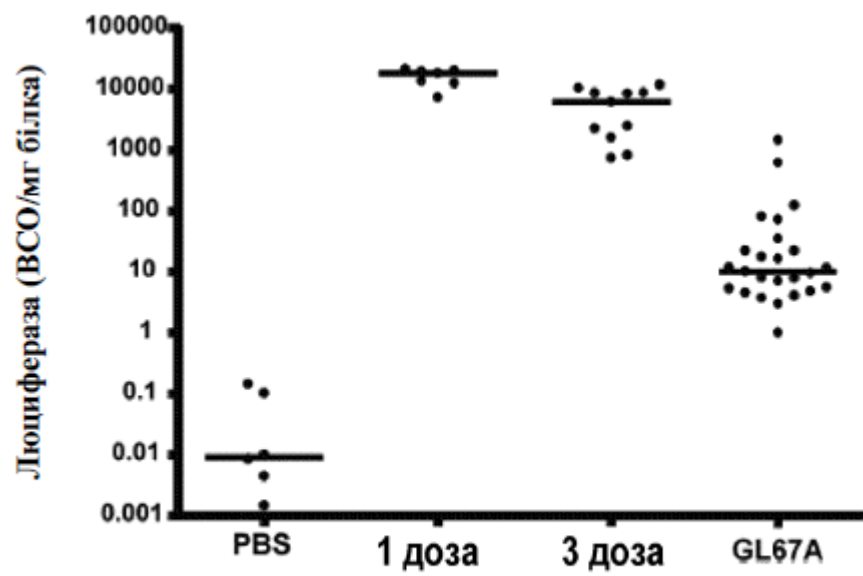
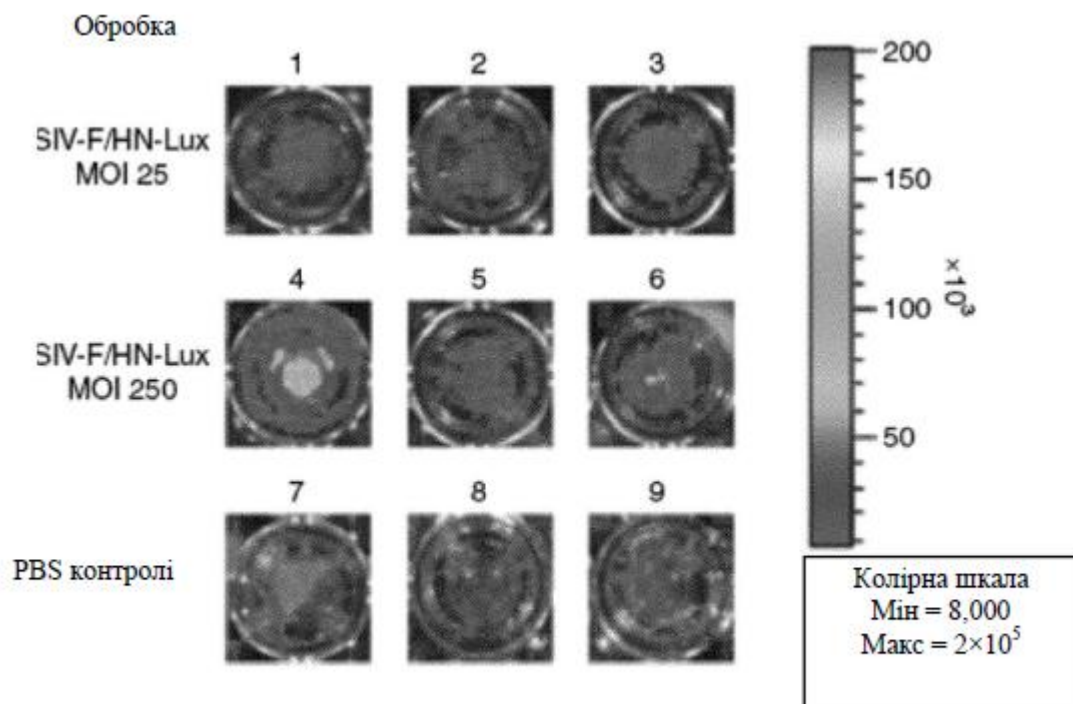
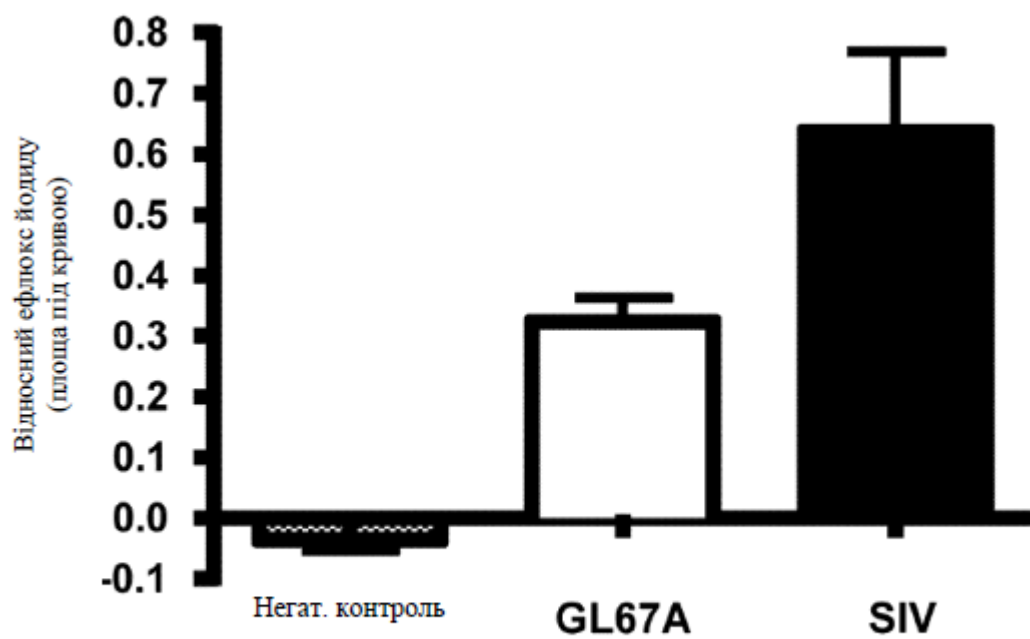


Fig. 6



Фіг. 7



Фіг. 8

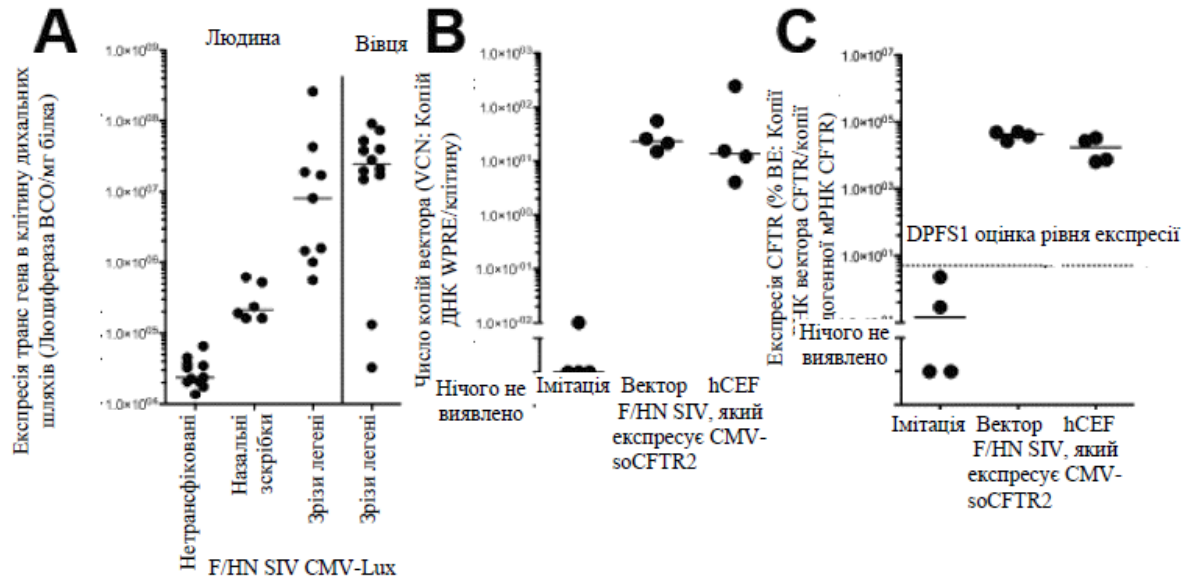


Fig. 9

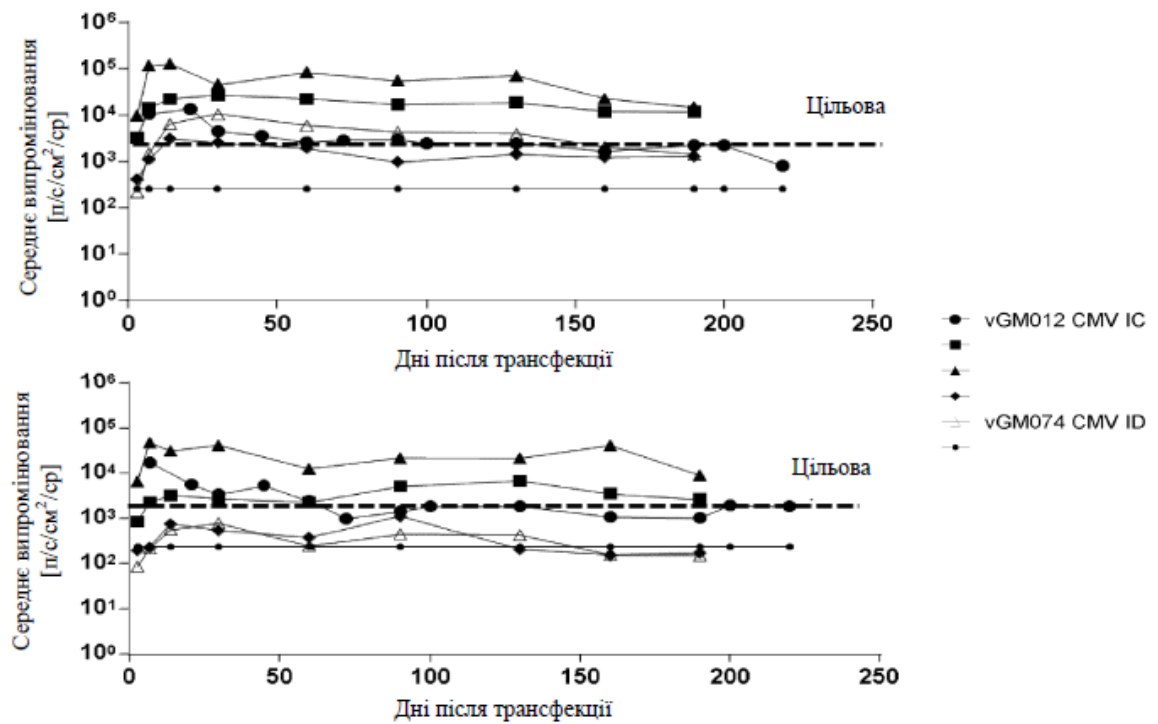
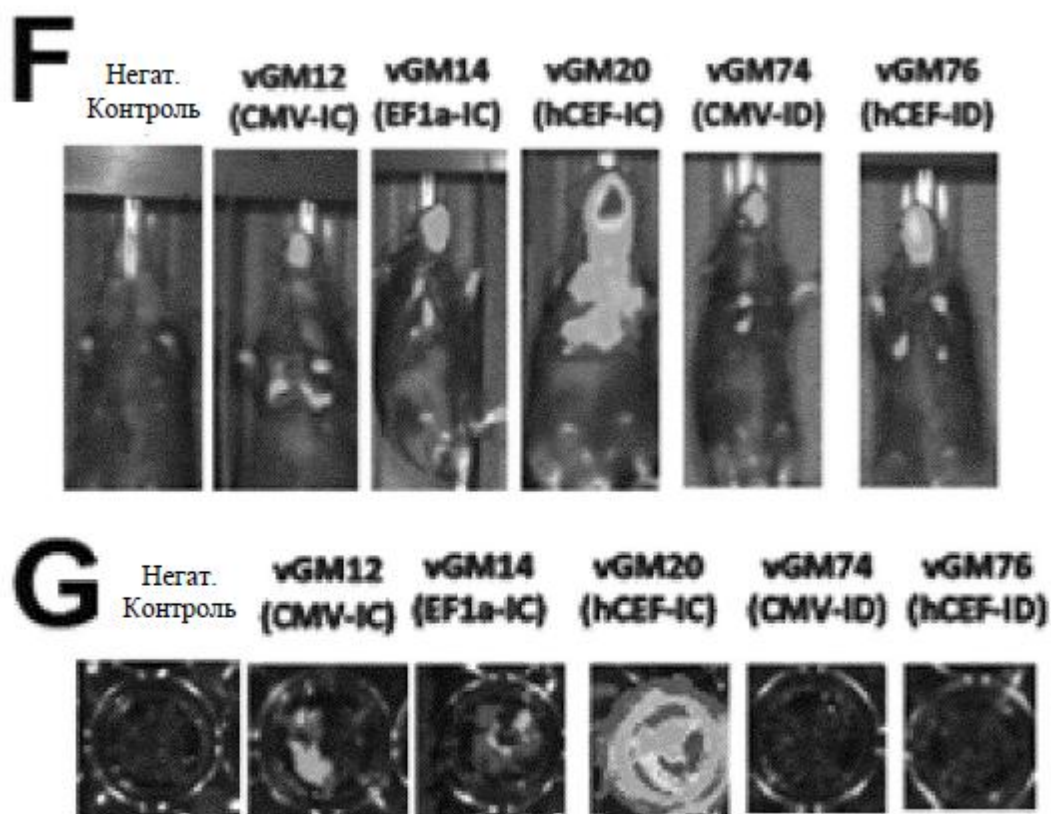
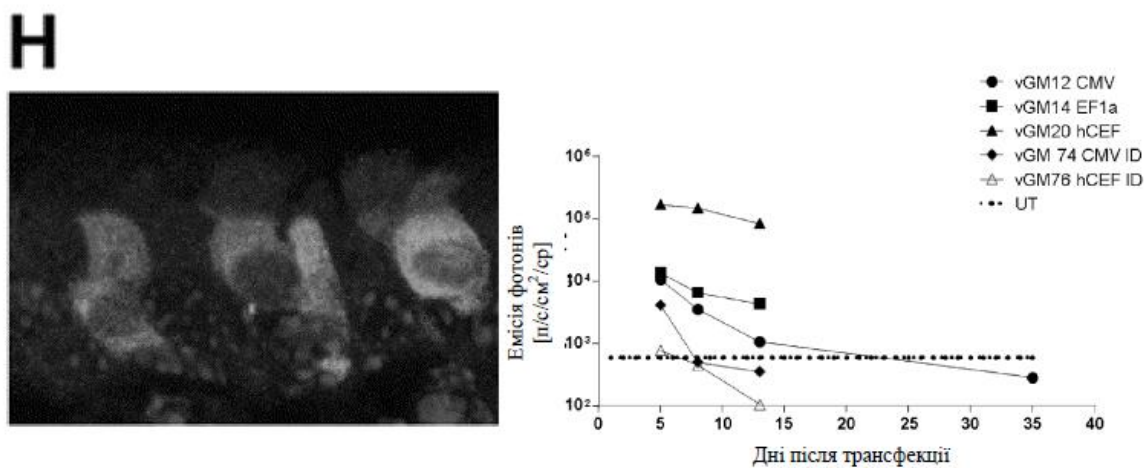


Fig. 9 (Продовження)



Фіг. 9 (Продовження)



Фіг. 9 (Продовження)

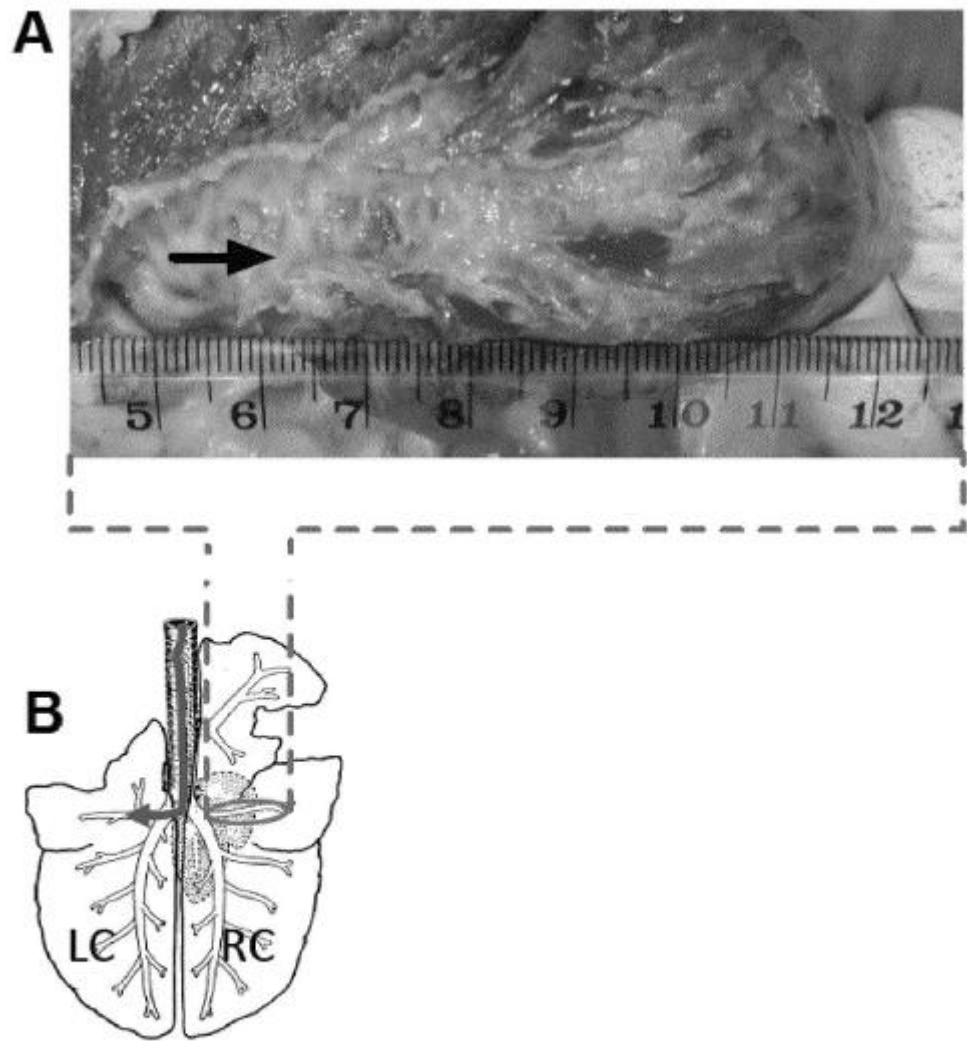
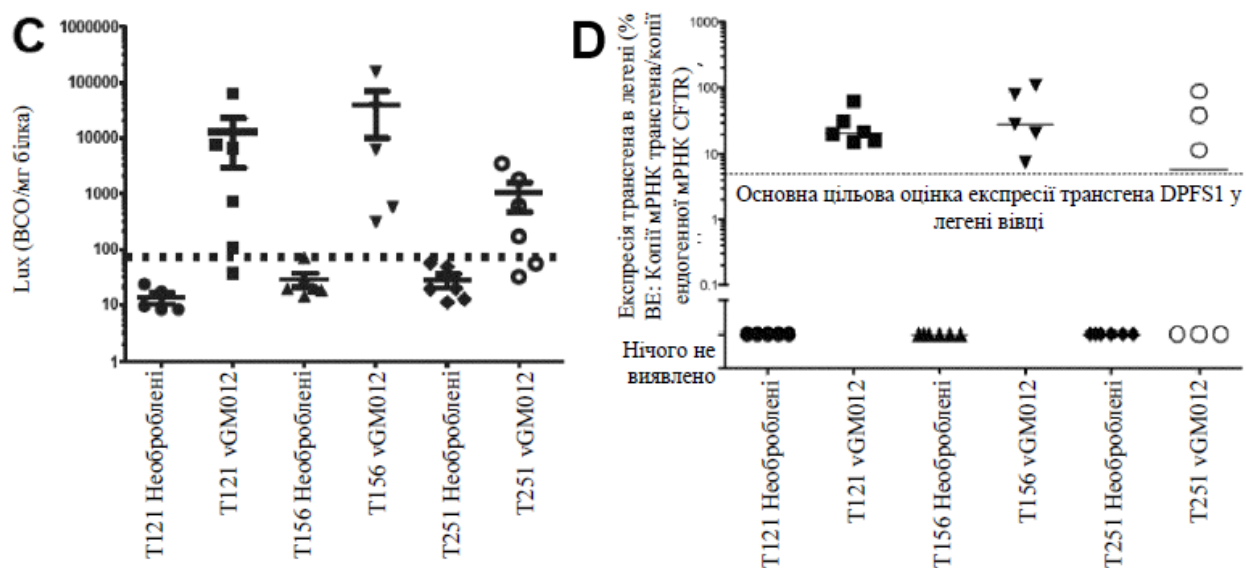
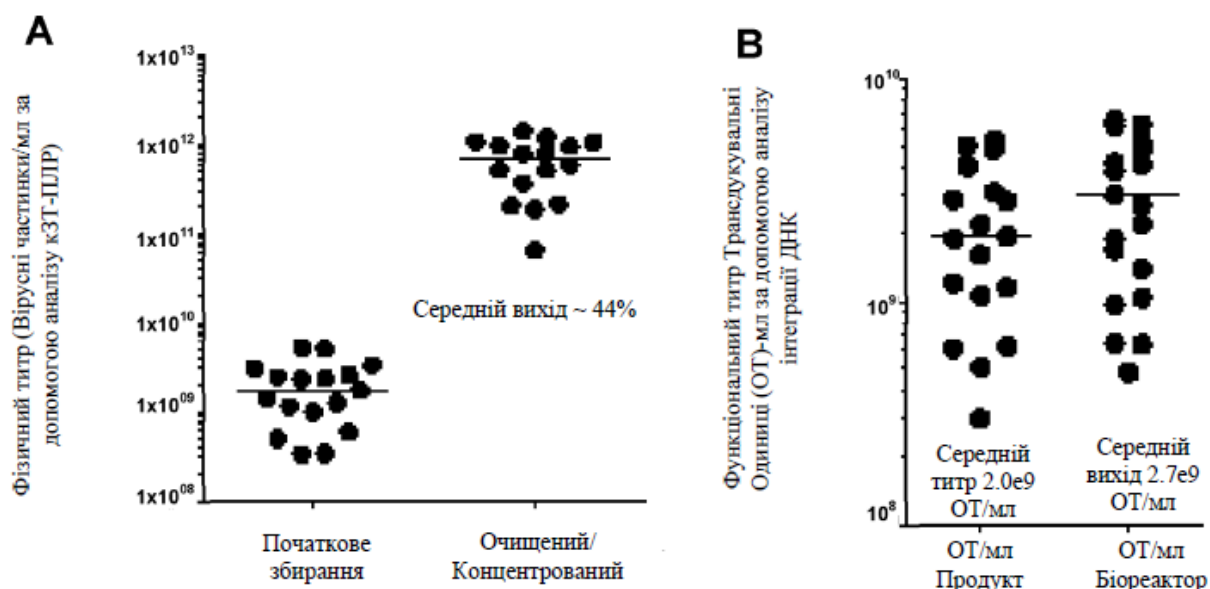


Fig. 10



Фіг. 10 (Продовження)



Фіг. 11

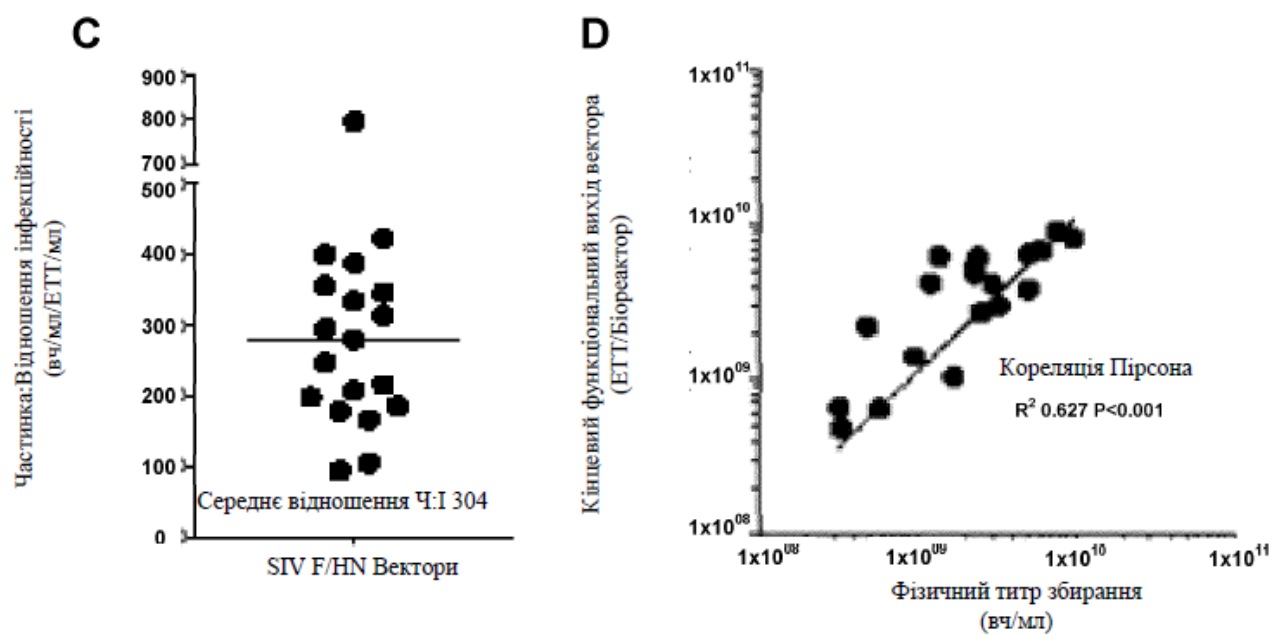


Fig. 11 (Продовження)

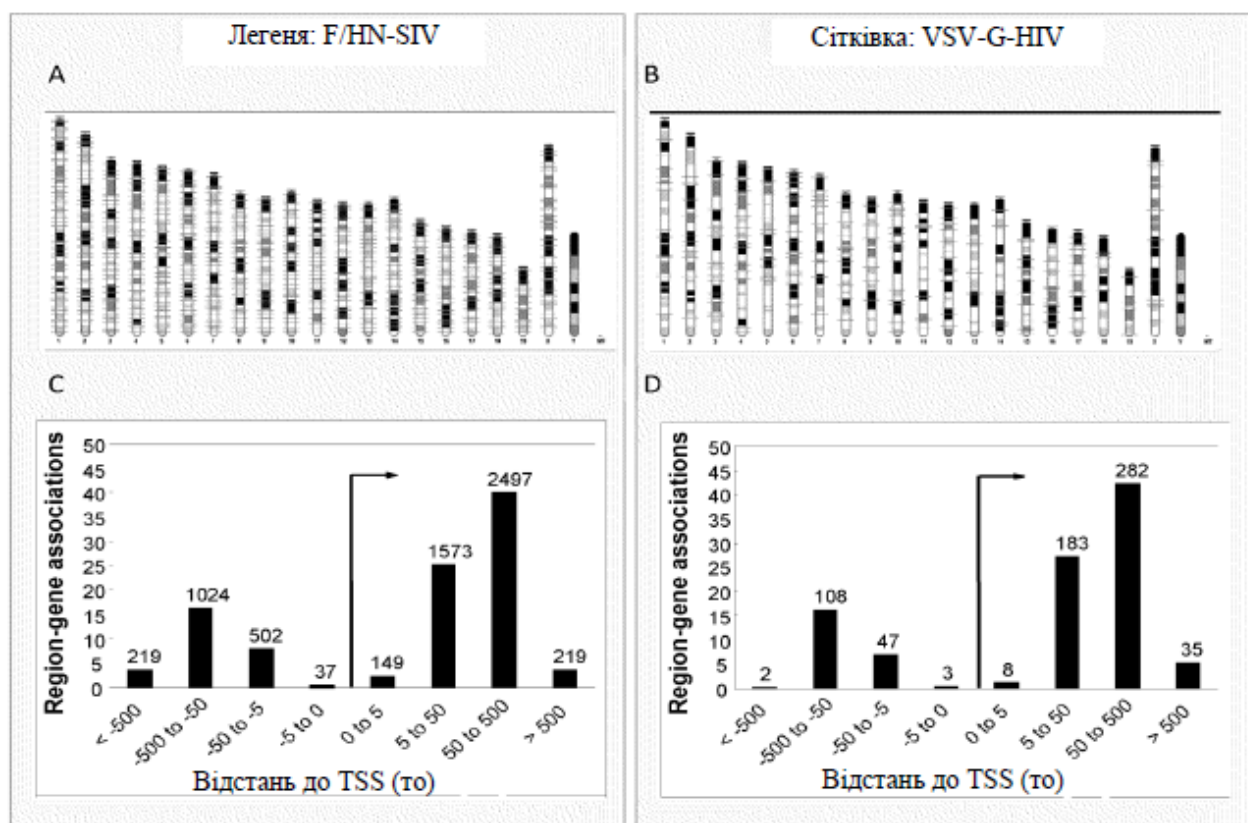
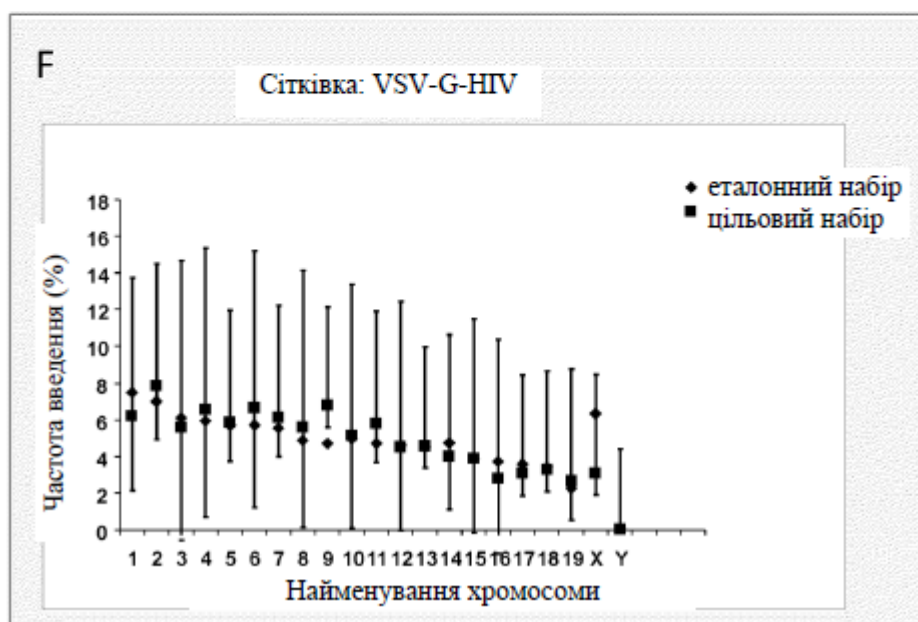
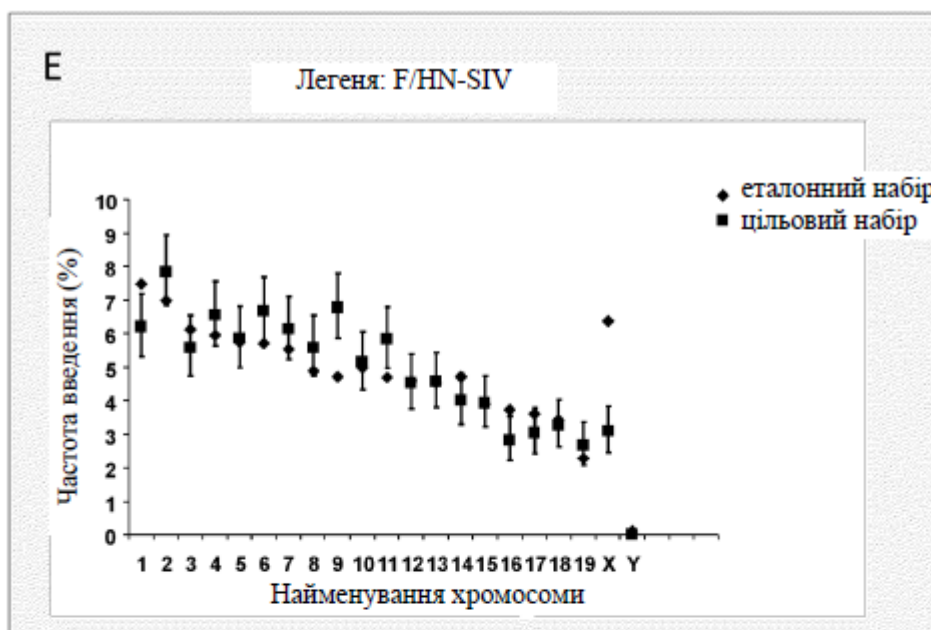
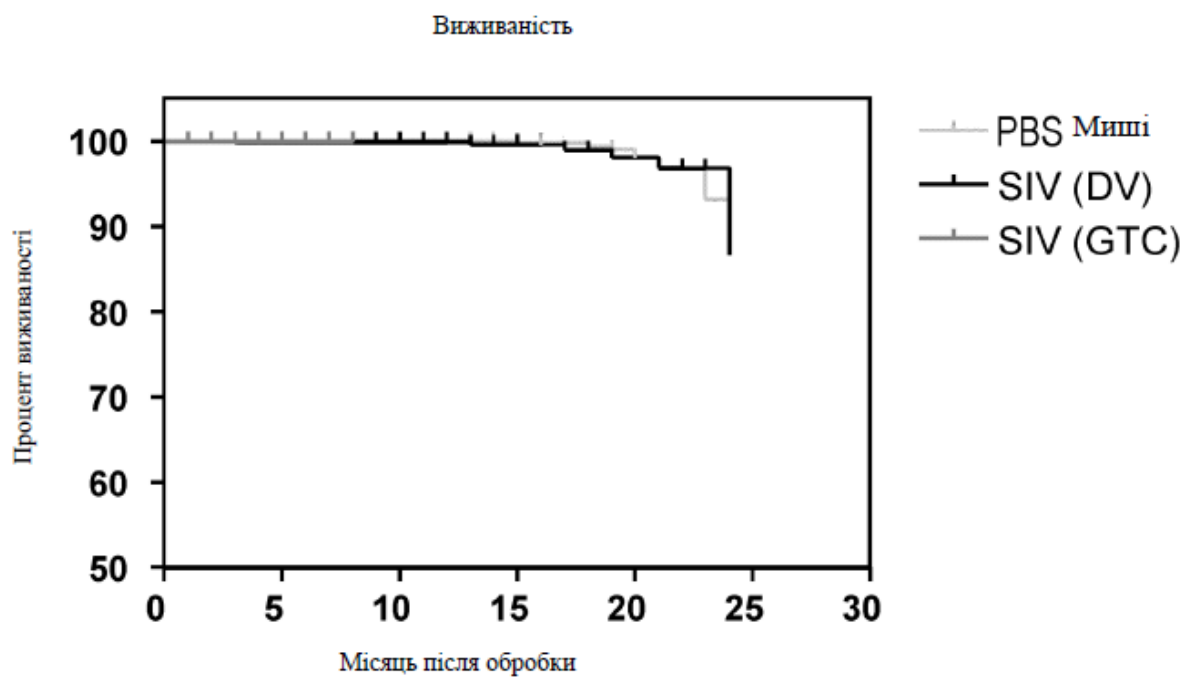


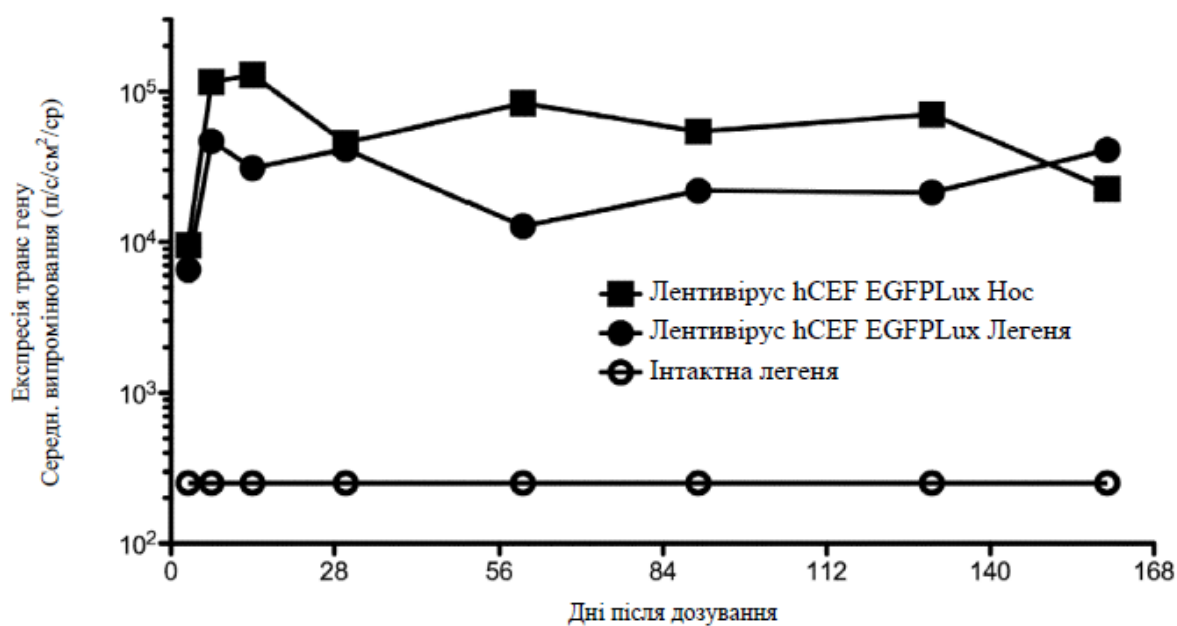
Fig. 12



Фіг. 12 (Продовження)



Фіг. 12 (Продовження)



Фіг. 13

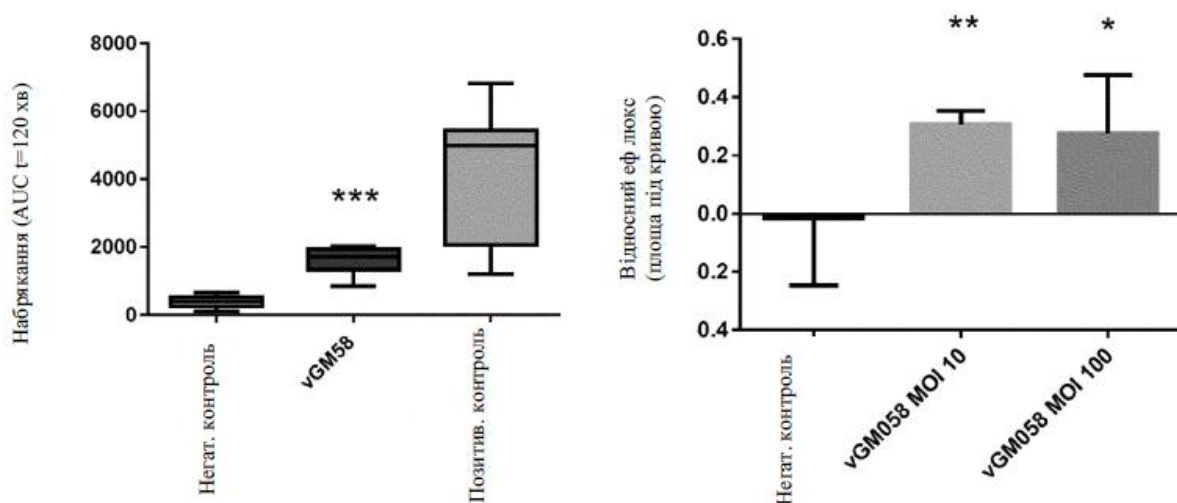


Fig. 14

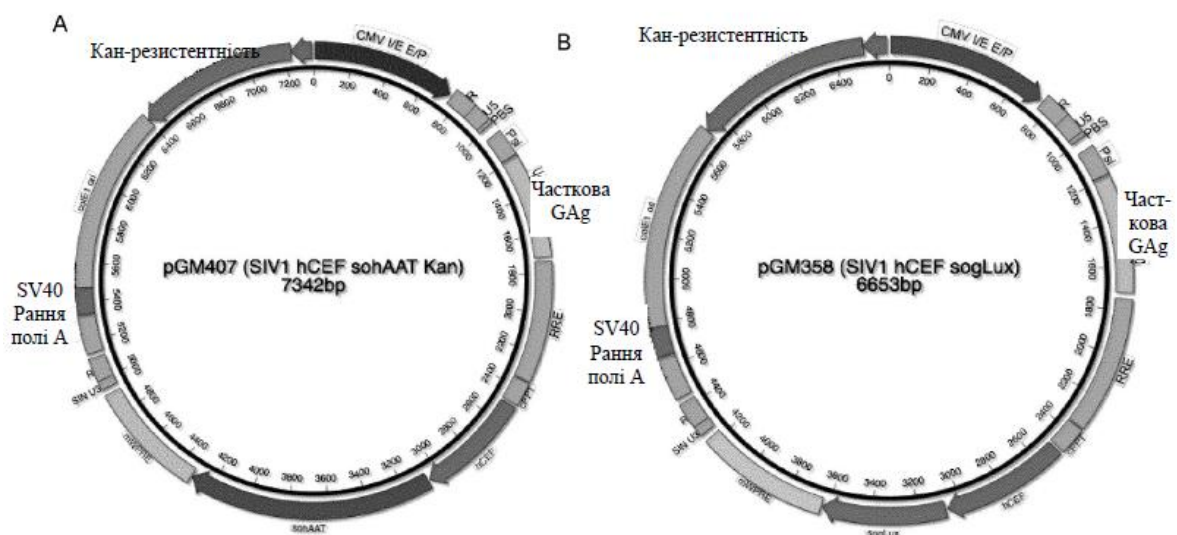


Fig. 15

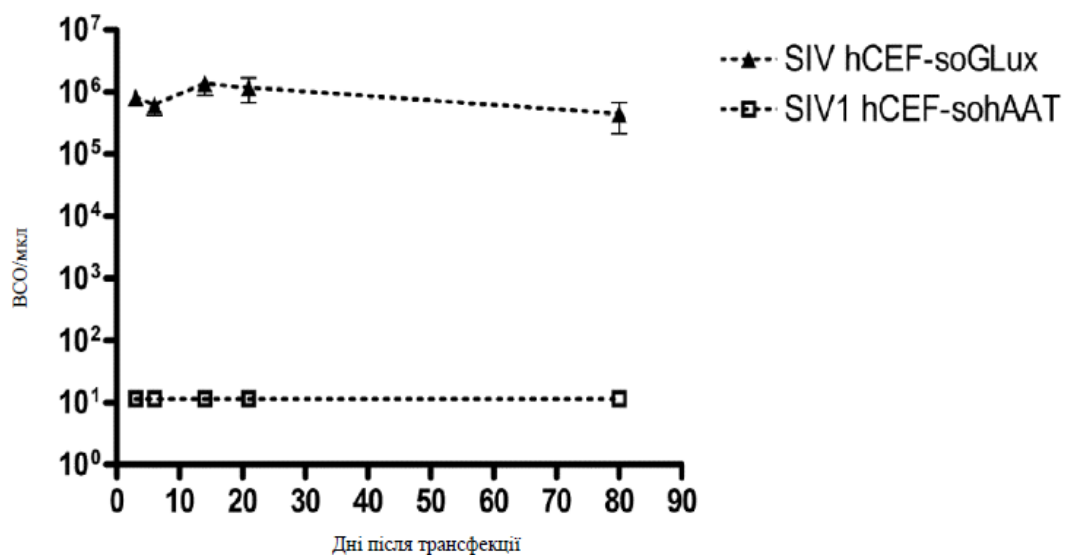
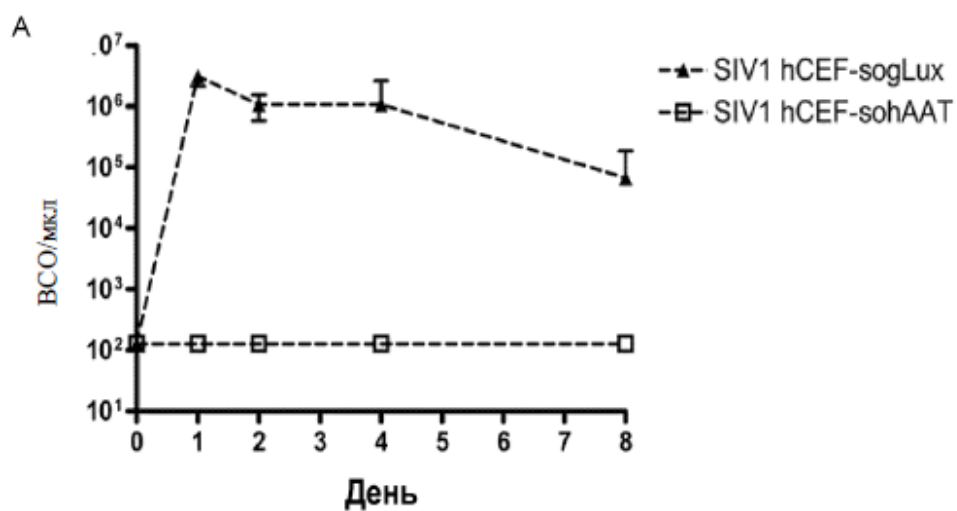


Fig. 16



В

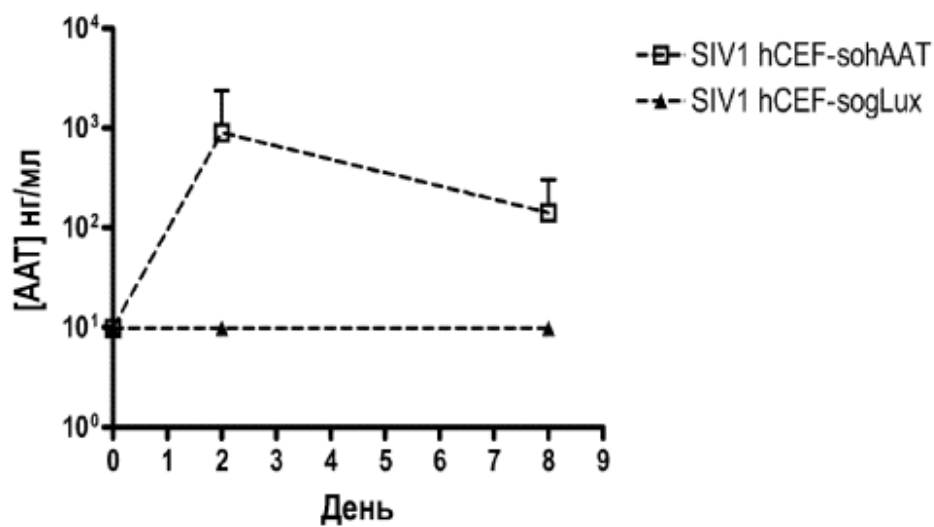
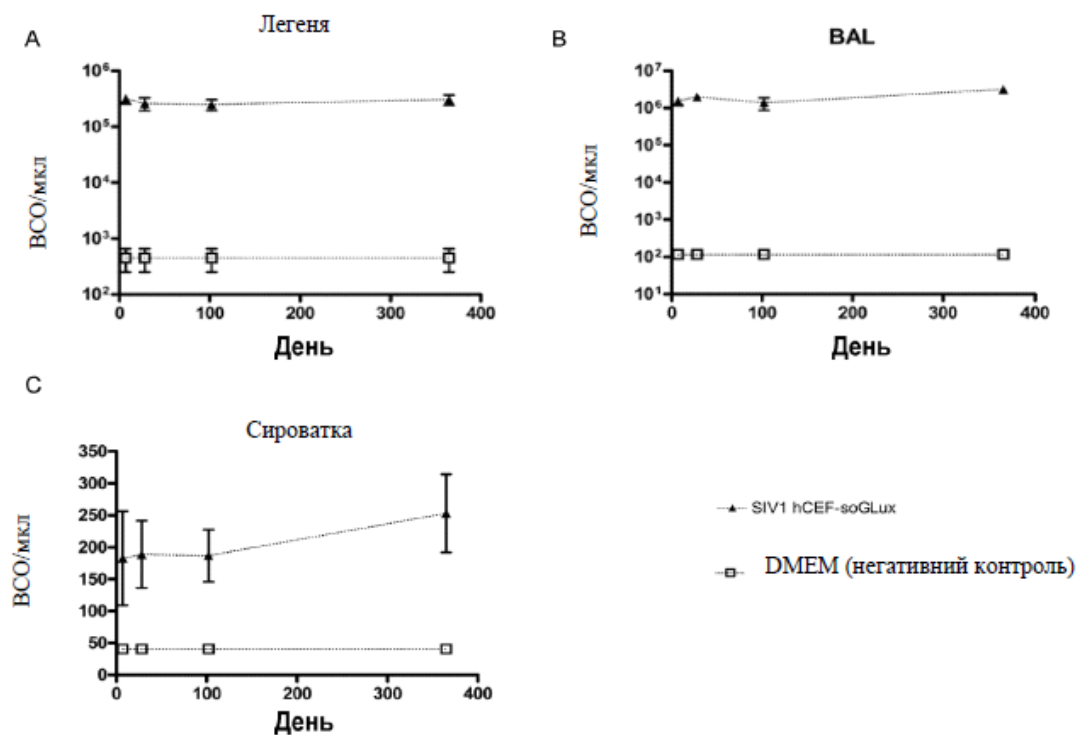
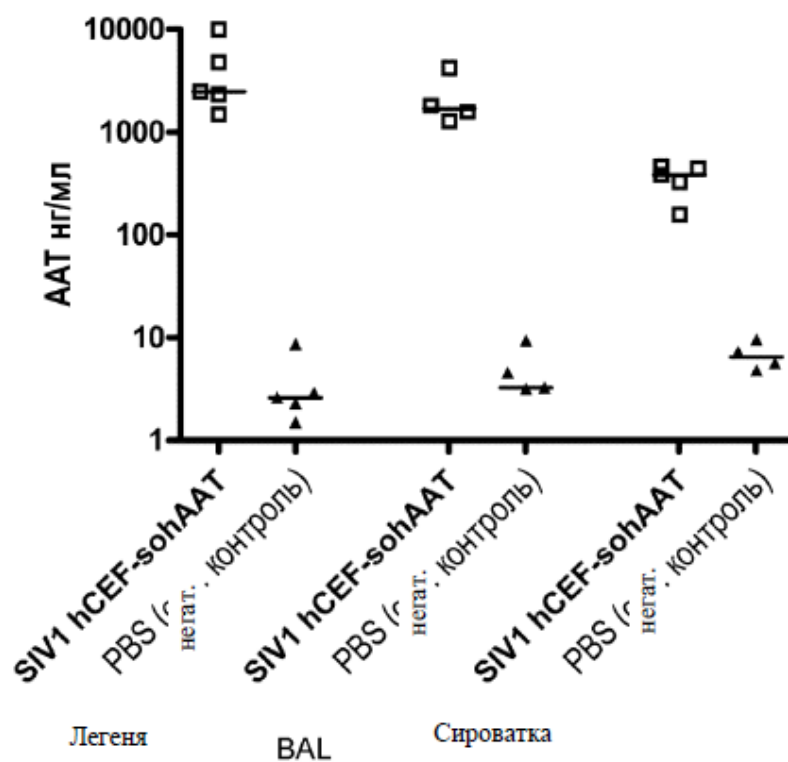


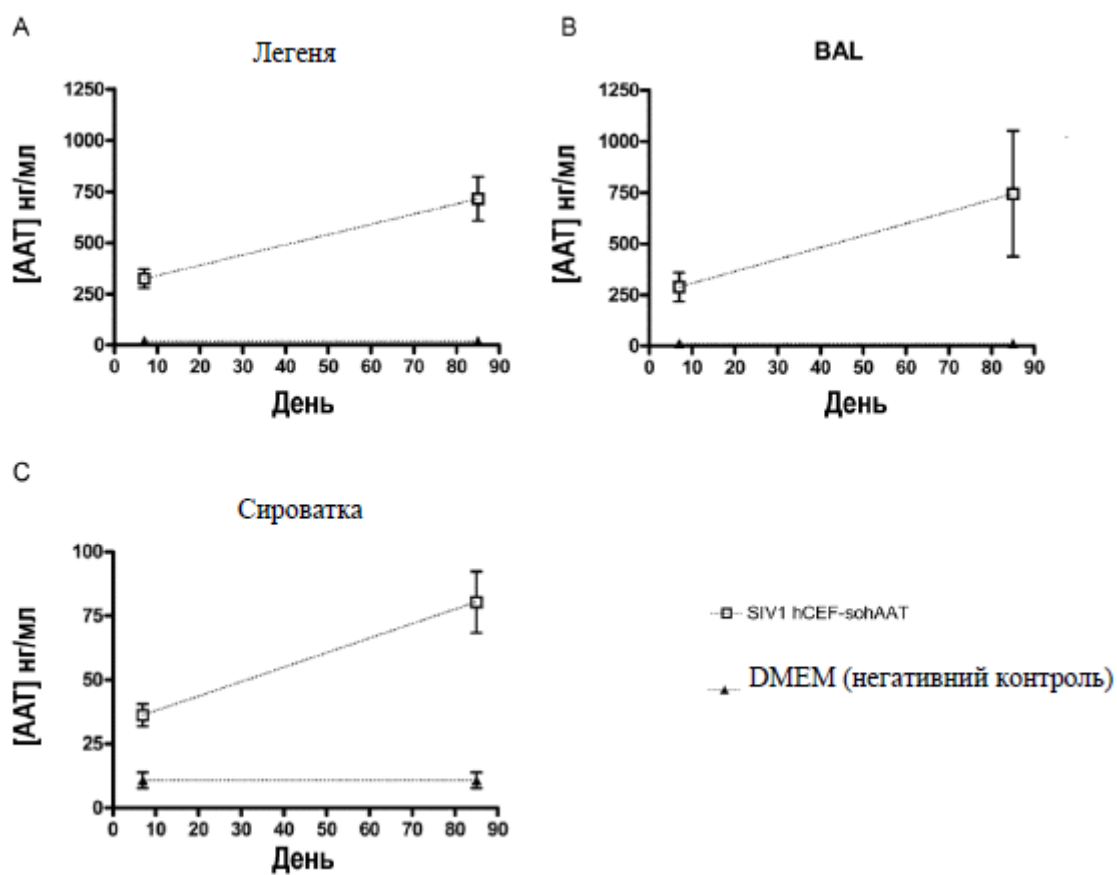
Fig. 17



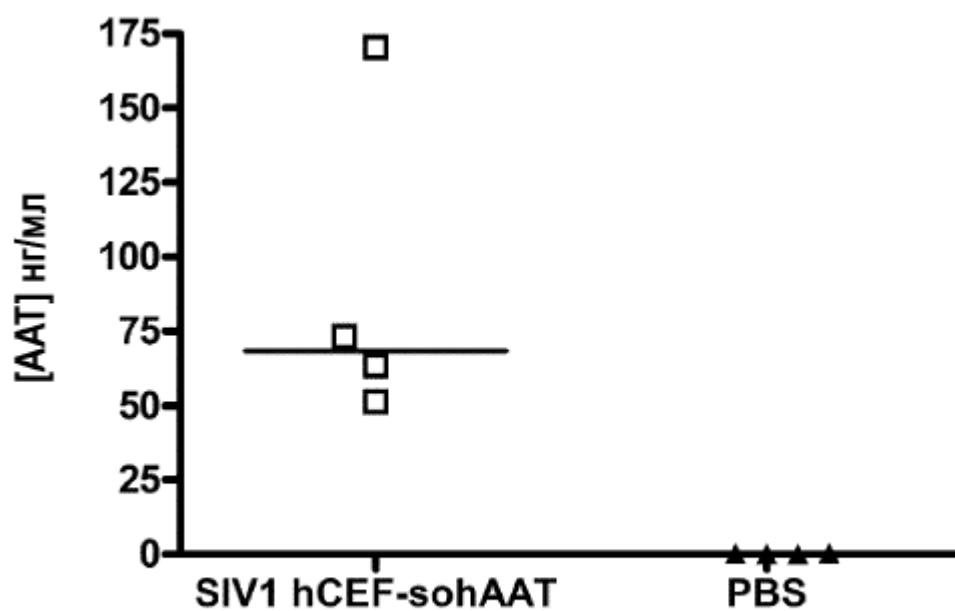
Фиг. 18



Фиг. 19



Фиг. 20



Фиг. 21

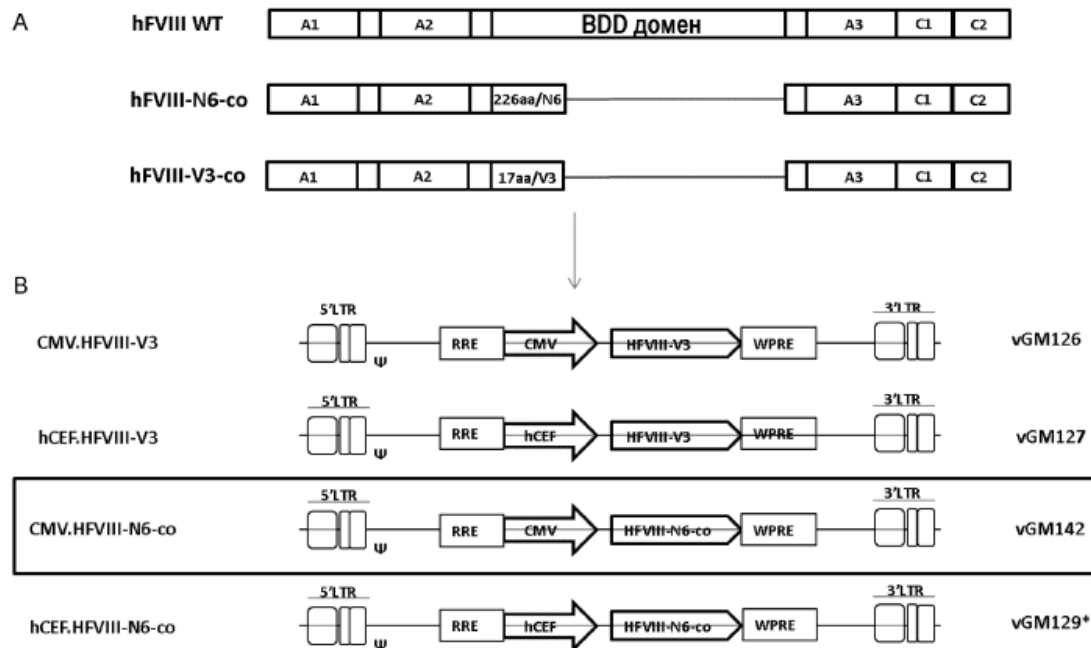
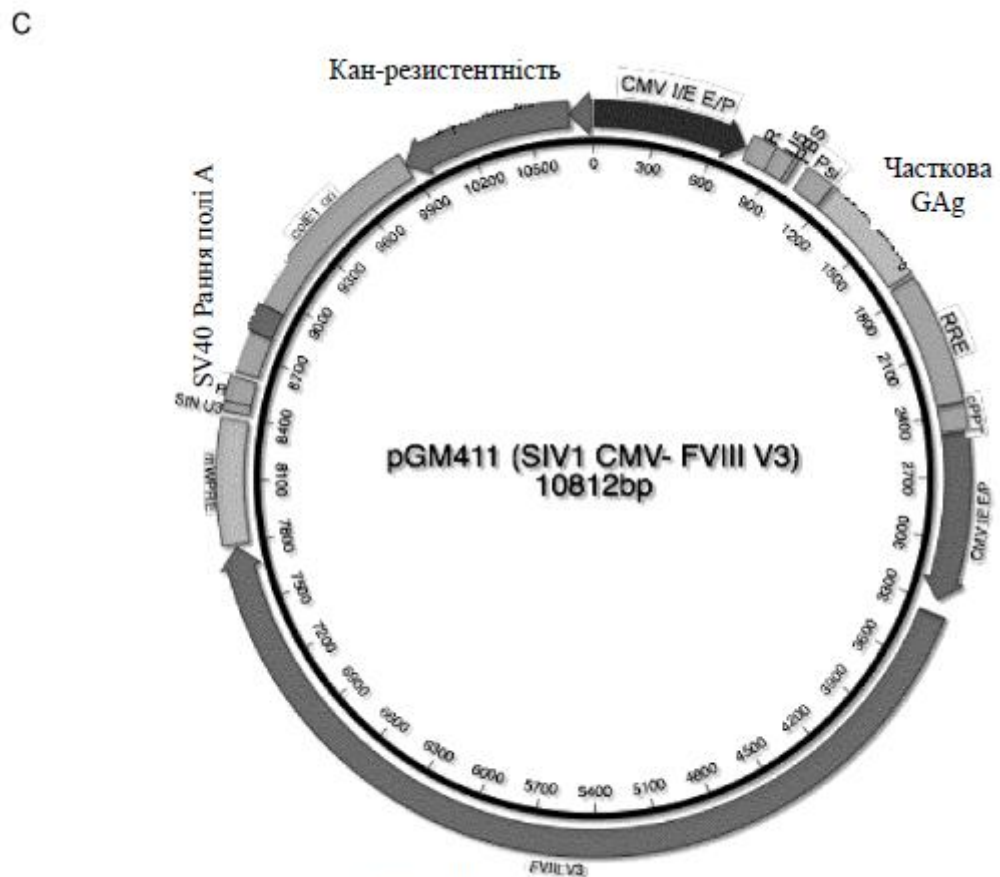
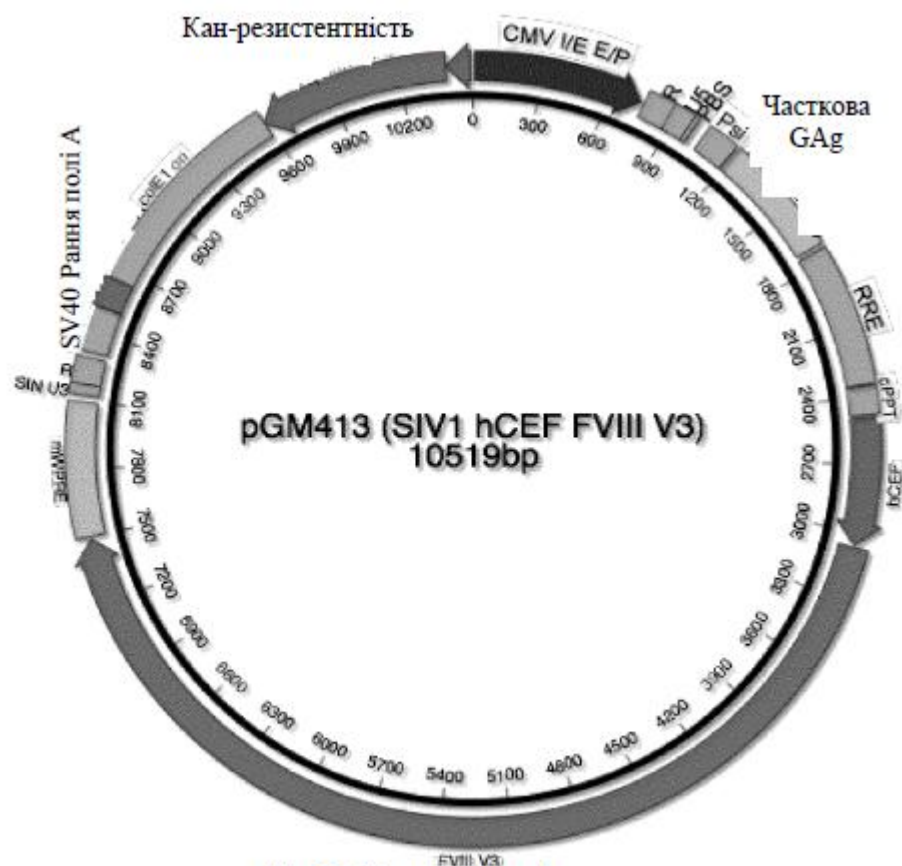


Fig. 22



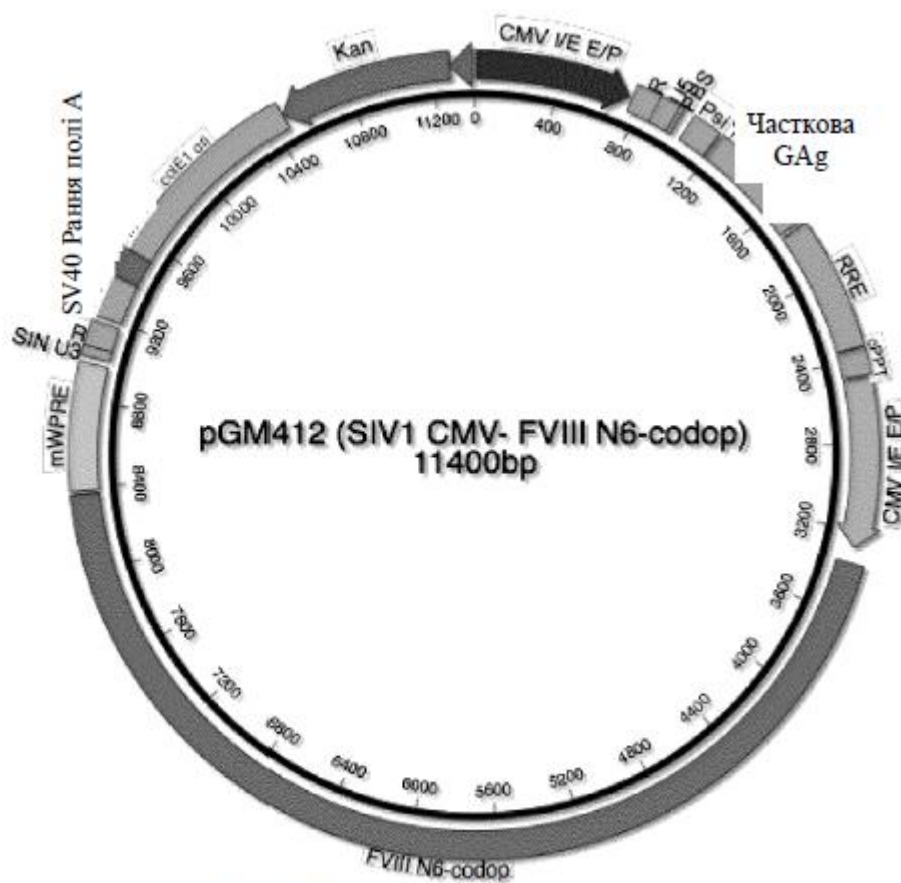
Фіг. 22 (Продовження)

D



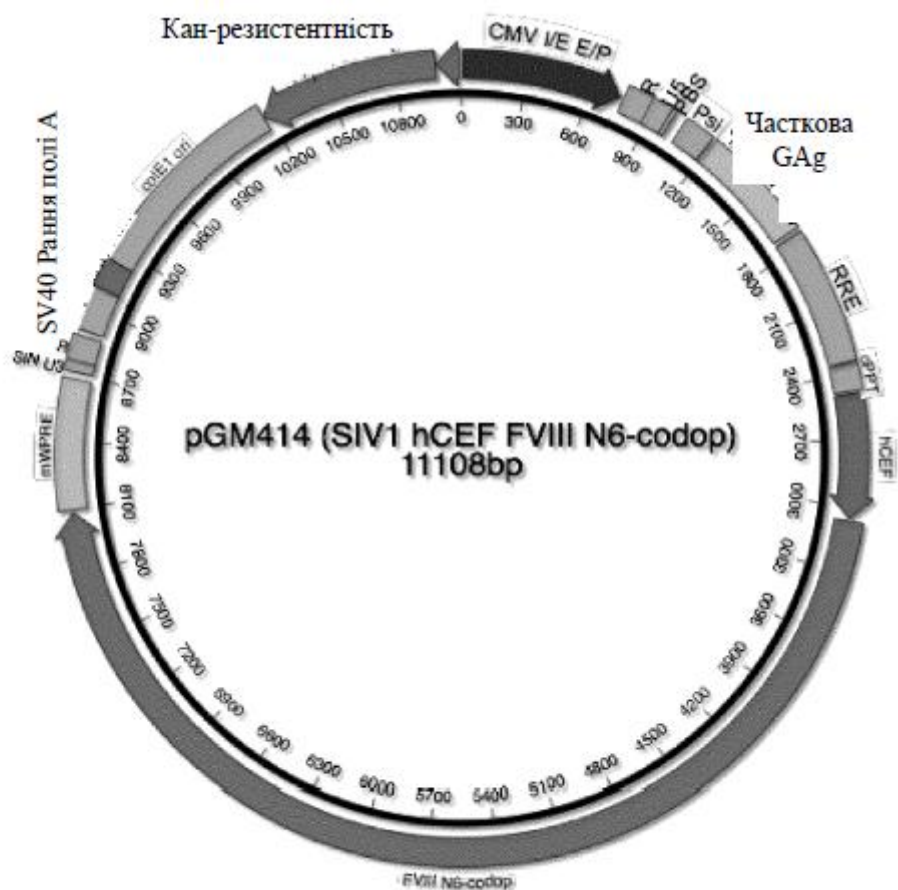
Фіг. 22 (Продовження)

E



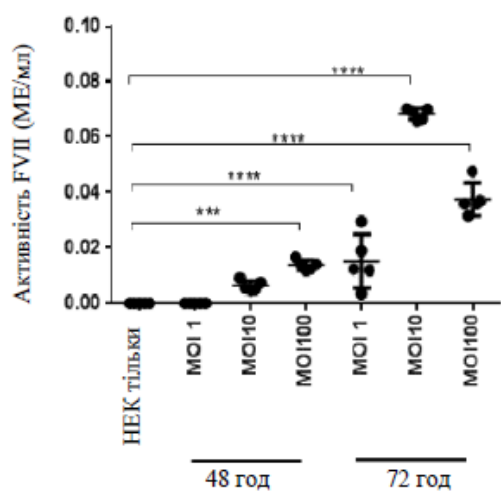
Фіг. 22 (Продовження)

F



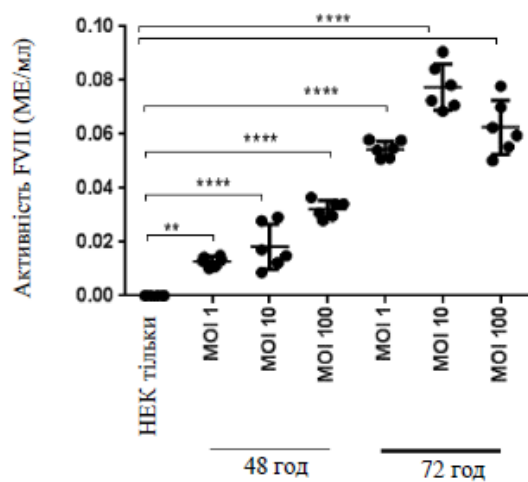
Фіг. 22 (Продовження)

A



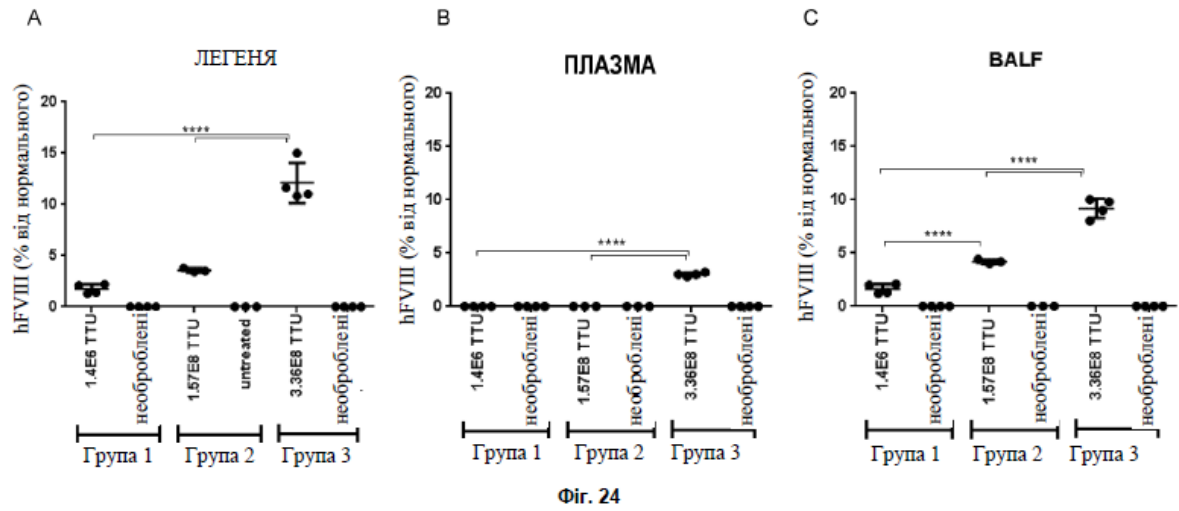
Завантаження $1-5.9 \times 10^8$ ЕТТ/мл

B



Завантаження $2-2.8 \times 10^8$ ЕТТ/мл

Фіг. 23



Комп'ютерна верстка Л. Ціхановська

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

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