



УКРАЇНА

(19) **UA** (11) **128156** (13) **C2**  
(51) МПК

**C07K 14/325** (2006.01)  
**C12N 1/21** (2006.01)  
**C12N 5/04** (2006.01)  
**C12N 15/32** (2006.01)  
**C12N 15/63** (2006.01)  
**A01H 5/10** (2018.01)

НАЦІОНАЛЬНИЙ ОРГАН  
ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ  
ДЕРЖАВНА ОРГАНІЗАЦІЯ  
"УКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ  
ОФІС ІНТЕЛЕКТУАЛЬНОЇ  
ВЛАСНОСТІ ТА ІННОВАЦІЙ"

**(12) ОПИС ДО ПАТЕНТУ НА ВІНАХІД**

(21) Номер заявки: **a 2019 09213**  
(22) Дата подання заявки: **11.01.2018**  
(24) Дата, з якої є чинними права інтелектуальної власності: **25.04.2024**  
(31) Номер попередньої заявки відповідно до Паризької конвенції: **62/445,313**  
(32) Дата подання попередньої заявки відповідно до Паризької конвенції: **12.01.2017**  
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: **US**  
(41) Публікація відомостей про заявку: **26.12.2019, Бюл.№ 24**  
(46) Публікація відомостей про державну реєстрацію: **24.04.2024, Бюл.№ 17**  
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ: **PCT/US2018/013298, 11.01.2018**

(72) Винахідник(и):  
**Баум Джеймс А. (US),  
Боуен Девід Дж. (US),  
Чей Кетрін А. (US),  
Чі Девід Дж. (US),  
Клінтон Уїлльям П. (US),  
Дарт Крістал Л. (US),  
Інгліш Лі (US),  
Фласінскій Станіслав (US),  
Гузов Віктор М. (US),  
Джаррелл Кевін А. (US),  
Кесенейполлі Ума Р. (US),  
Малвар Томас М. (US),  
МакКерролл Роберт М. (US),  
Мілліган Джейсон С. (US),  
Моргенстерн Джей П. (US),  
Рукер Дебора Г. (US),  
Сальвадор Сара А. (US),  
Сміт Темпл Ф. (US),  
Сото Карлос Е. (US),  
Сталц Коллін М. (US),  
Турчік Брайан М. (US),  
Вон Тай Т. (US),  
фон Рехенберг Морітц В. Ф. Ф. (US)**

(73) Володілець (володільці):  
**МОНСАНТО ТЕКНОЛОДЖІ ЛЛС,  
800 North Lindbergh Boulevard, St. Louis, MO  
63167, United States of America (US)**

(74) Представник:  
**Бреус Наталія Володимирівна, реєстр.  
№167**

(56) Перелік документів, взятих до уваги експертизою:  
**WO 2017003811 A1, 05.01.2017  
WO 2013134734 A2, 12.09.2013  
WO 2016061392 A1, 21.04.2016  
EP 0228838 A2, 15.07.1987  
US 2010168387 A1, 01.07.2010  
US 20120035135 A1, 09.02.2012  
Us 20150148288 A1, 28.05.2015**

**UA 128156 C2**

**(54) ПЕСТИЦИДНИЙ БІЛОК ТОКСИНУ, АКТИВНИЙ ПРОТИ ЛУСКОКРИЛИХ КОМАХ**

---

(57) Реферат:

Винахід стосується нуклеотидної послідовності, яка кодує інсектицидний білок, який демонструє інгібуючу активність по відношенню до лускокрилих, химери, які містять токсичний для совки-іпсилон (BCW-black cutworm) білок і інсектицид BCW-токсин, трансгенної рослини, яка експресує вказану химеру або вказаний інсектицид, а також новий інсектицидний білок, позначений як BCW 003, і спосіб виявлення присутності нуклеотидних послідовностей або вказаного інсектициду в біологічному зразку.

## ПОСИЛАННЯ НА СПОРІДНЕНУ ЗАЯВКУ

Ця заявка заявляє пріоритет за попередньою заявкою США № 62/445313, поданою 12 січня 2017 року, яка включена в цей документ з допомогою посилання в повному об'ємі.

## ВКЛЮЧЕННЯ ПЕРЕЛІКУ ПОСЛІДОВНОСТЕЙ

5 Файл, названий "MONS434WO-sequence listing.txt", який містить Перелік послідовностей в машинозчитуваному форматі, був створений 11 січня 2018 року. Цей файл має розмір 166134 байт, як виміряно в MS-Windows®, і повністю включений в цей документ за допомогою посилання. Цей Перелік послідовностей складається з SEQ ID NO:1-SEQ ID NO:12.

## ГАЛУЗЬ ВИНАХОДУ

10 Винахід відноситься, в цілому, до галузі білків, які проявляють інгібуючу активність по відношенню до комах, зокрема до білків, які проявляють інгібуючу активність по відношенню до комах-шкідників сільськогосподарства із ряду Lepidoptera (лускокрилі), таких як совка-іпсилон ("BCW", *Agrotis ipsilon*).

## РІВЕНЬ ТЕХНІКИ

15 Білки, які пригнічують комах, які продукуються видом бактерії *Bacillus thuringiensis* (Bt), відомі в цій галузі техніки. Певні білки Bt можуть бути використані для боротьби з шкідниками сільськогосподарських культур шляхом розприскування прийнятних в сільському господарстві готових форм, які містять один або більшу кількість таких білків, на рослини, шляхом покриття насіння композицією, приготованою так, що вона містить інсектицидно-ефективну кількість таких білків, або шляхом результативно-ефективної експресії вказаних одного або більшої кількості білків в рослинах/насінні.

20 Були розроблені лише кілька білків Bt для використання в якості трансгенних ознак для комерційного використання фермерами для боротьби з комахами-шкідниками. Фермери покладаються на ці білки, щоб отримати вказаний спектр стримування шкідників, і можуть продовжувати покладатися на хімічні речовини широкого спектра дії при застосуванні до листя і 25 ґрунту для боротьби зі шкідниками. Доведено, що особливо важко боротися з деякими лускокрилими комахами, такими як види *Agrotis* і види *Striacosta*, з використанням трансгенних інсектицидних ознак, що застосовуються в даний час, включаючи Cry1Ab, Cry1Ac, Cry1Fa, Cry2Ab, Cry2Ae, VIP3Aa і різні інші токсини Bt, які використовуються менш часто. Отже, існує 30 потреба в білках, які пригнічують комах, які проявляють активність проти більш широкого спектра видів комах-шкідників, і у використанні токсинів з метою подолання розвитку стійкості шкідників до існуючих пестицидів, включаючи токсини, які застосовуються в даний час в системах контролю шкідників.

35 Ця заявка описує нову родину білків, варіанти і химерні токсичні білкові конструкції, кожна з яких проявляє дивовижно ефективну інсектицидну активність по відношенню до лускокрилих, особливо проти шкідників видів *Agrotis*, таких як совка-іпсилон.

## КОРОТКИЙ ОПИС СУТІ ВИНАХОДУ

40 Продемонстровано, що нова група поліпептидів, які пригнічують комах (токсичні білки BCW 001, BCW 002 і BCW 003 та їх пестицидні фрагменти), володіє інгібуючою активністю по відношенню до кількох лускокрилих-шкідників сільськогосподарських рослин, зокрема по відношенню до виду совки-іпсилон (вид *Agrotis*). Кожний білок може бути використаний окремо або в комбінації один з одним і з іншими білками Bt і агентами, які інгібують комах, в готових формах та *in planta*, тим самим пропонуючи альтернативи Bt білкам та хімічним інсектицидам, які в даний час застосовуються в сільськогосподарських системах.

45 Згідно цього винаходу запропоновані поліпептидні конструкції, які містять у функціональному зв'язку сегмент гетерологічного промотора, зв'язаний з нуклеотидною послідовністю, яка кодує інсектицидний білок, який володіє характеристиками Cry1A, який менший ніж повнорозмірний токсичний білок класу Cry1A, і який має амінокислотну послідовність від близько позиції 1 до позиції 607, як вказано в SEQ ID NO: 2, SEQ ID NO: 4, 50 SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10 і SEQ ID NO: 12, або його інсектицидно-активний фрагмент. Менший ніж повнорозмірний поліпептид, який демонструє таку інсектицидну активність, має демонструвати ідентичність, щонайменше, близько 100 %, 99 %, 98 %, 97 %, 96 %, 95 %, 94 %, 94 %, 92 %, 91 % або 90 % з амінокислотною послідовністю BCW 001, як вказано в SEQ ID NO: 2, від близько позиції 1 до близько позиції 606, або від близько позиції 5 55 до близько позиції 600. Якщо необхідно використовувати повнорозмірний або значно більшого розміру фрагменти токсину, відсоток ідентичності має бути менш обмежуючим, і являти собою відсоток ідентичності від близько 100, до близько 95, до близько 90, до близько 85 або навіть до 80 % ідентичності по відношенню до послідовностей повнорозмірних білкових токсинів, як вказано в SEQ ID NO: 2, 4 і 6, так як ці токсичні білки демонструють комерційно корисні рівні 60 біологічної активності при тестуванні проти личинок совки-іпсилон в харчових біоаналізах і при

тестуванні in planta на трансгенній кукурудзі, бавовні та сої, які експресують ці білки.

Згідно цього винаходу запропоновані білки, токсичні для видів лускокрилих, включаючи білки, які мають амінокислотну послідовність, як вказано в SEQ ID NO: 2 від позиції 256 до 606 (білок BCW 001), і білки, які мають амінокислотну послідовність, як вказано в будь-якій з SEQ ID NO: 4 і SEQ ID NO: 6 від амінокислотної позиції 257 до 607 (в цьому документі відповідно позначаються як токсичний білок BCW 002 і токсичний білок BCW 003).

Такі інсектицидні білки також проявляють активність проти видів лускокрилих, вибраних з групи, що складається з: *Spodoptera frugiperda*, *Spodoptera exigua*, *Spodoptera litura*, *Mamestra configurata*, *Striacosta albicosta*, *Trichoplusia ni*, *Pseudoplusia includens*, *Anticarsia gemmatalis*, *Hypena scabra*, *Heliothis virescens*, *Agrotis subterranea*, *Pseudaletia unipuncta*, *Agrotis ipsilon*, *Agrotis orthogonia*, *Ostrinia nubilalis*, *Amyeloides transitella*, *Crambus caliginosellus*, *Herpetogramma licarsisalis*, *Homoeosoma electellum*, *Elasmopalpus lignosellus*, *Cydia pomonella*, *Endopiza viteana*, *Grapholita molesta*, *Suleima helianthana*, *Plutella xylostella*, *Pectinophora gossypiella*, *Lymantria dispar*, *Alabama argillacea*, *Archips argyrospila*, *Archips rosana*, *Chilo suppressalis*, *Cnaphalocrocis medinalis*, *Crambus caliginosellus*, *Crambus teterrellus*, *Diatraea grandiosella*, *Diatraea saccharalis*, *Earias insulana*, *Egrias vittella*, *Helicoverpa armigera*, *Helicoverpa zea*, *Heliothis virescens*, *Herpetogramma licarsisalis*, *Lobesia botrana*, *Pectinophora gossypiella*, *Phyllocnistis citrella*, *Pieris brassicae*, *Pieris rapae*, *Plutella xylostella*, і *Tuta absoluta*.

Білки згідно цього винаходу також можуть проявляти біологічну активність по відношенню до видів лускокрилих, вибраних з групи, що складається з: *Spodoptera frugiperda*, *Spodoptera exigua*, *Spodoptera litura*, *Mamestra configurata*, *Striacosta albicosta*, *Trichoplusia ni*, *Pseudoplusia includens*, *Anticarsia gemmatalis*, *Hypena scabra*, *Heliothis virescens*, *Agrotis subterranea*, *Pseudaletia unipuncta*, *Agrotis ipsilon*, *Agrotis orthogonia*, *Ostrinia nubilalis*, *Amyeloides transitella*, *Crambus caliginosellus*, *Herpetogramma licarsisalis*, *Homoeosoma electellum*, *Elasmopalpus lignosellus*, *Cydia pomonella*, *Endopiza viteana*, *Grapholita molesta*, *Suleima helianthana*, *Plutella xylostella*, *Pectinophora gossypiella*, *Lymantria dispar*, *Alabama argillacea*, *Archips argyrospila*, *Archips rosana*, *Chilo suppressalis*, *Cnaphalocrocis medinalis*, *Crambus caliginosellus*, *Crambus teterrellus*, *Diatraea grandiosella*, *Diatraea saccharalis*, *Earias insulana*, *Egrias vittella*, *Helicoverpa armigera*, *Helicoverpa zea*, *Heliothis virescens*, *Herpetogramma licarsisalis*, *Lobesia botrana*, *Pectinophora gossypiella*, *Phyllocnistis citrella*, *Pieris brassicae*, *Pieris rapae*, *Plutella xylostella*, і *Tuta absoluta*.

Білки згідно цього винаходу і конструкції, які розглядаються в цьому документі, можуть бути включені в будь-який вектор, включаючи плазміди, косміди, фагміди і т. п.

Такі вектори можуть бути використані для введення конструкцій згідно цього винаходу в будь-яку кількість клітин-господарів, включаючи бактеріальні клітини, дріжджові клітини і клітини рослин.

Клітинами-господарями, які є дріжджовими клітинами, можуть бути *Saccharomyces cerevisiae* або *Saccharomyces rombe*, і т. п. Бактеріальні клітини-господарі можуть являти собою будь-яку кількість відомих таких клітин-господарів, включаючи, але не обмежуючись *E. coli*, *B. thuringiensis*, та інші споріднені бацили. Рослинні клітини-господарі можуть бути отримані з будь-якої кількості видів рослин, включаючи, але не обмежуючись рослинними клітинами з: люцерни, банана, ячменю, квасолі, брокколи, капусти, капусти декоративної, моркви, маніюки, рицини, цвітної капусти, селери, нуту, китайської капусти, цитрусових, кокосової пальми, кави, кукурудзи, конюшини, бавовни, гарбузових, огірка, псевдотсуги Мензиса, баклажана, евкаліпта, льону, часнику, винограду, хмелю, цибулі-порею, салату-латуку, сосни ладанної, проса, дині, горіха, вівса, оливкового дерева, ріпчастої цибулі, декоративних рослин, пальмових, пасовищних трав, гороху, арахісу, перцю, голубиноного гороху, соснових, картоплі, тополі, гарбуза, сосни променистої, редьки, ріпаку, рису, кореневищ, жита, дикого шафрану, чагарникових, сорго, сосни південної, сої, шпинату, гарбузових, полуниці, цукрового буряка, цукрової тростини, соняшнику, кукурудзи цукрової, амбрового дерева, батату, проса прутувидного, чаю, тютюну, помідора, тритикале, дернової трави, кавуна і пшениці.

Трансгенні рослини, зокрема трансгенні рослини кукурудзи, бавовни та сої, можуть бути отримані шляхом введення конструкцій згідно цього винаходу, які містять відповідним чином модифіковані полінуклеотидні послідовності, такі як вказані в SEQ ID NO: 7, SEQ ID NO: 9 і SEQ ID NO: 11, наприклад, в геном таких рослинних клітин і шляхом відновлення фертильної трансгенної рослини кукурудзи, сої або бавовни, які містять в своєму геномі генетичну конструкцію для експресії щонайменше білкового токсину згідно цього винаходу, тобто BCW 001, BCW 002 або білкового токсину BCW 003. Такі трансгенні рослини будуть мати розміщену в їх рослинному геномі полінуклеотидну конструкцію, яка містить щонайменше гетерологічний промоторний сегмент, функціонально зв'язаний з нуклеотидною послідовністю, яка кодує

інсектицидний білок BCW 001, BCW 002 або BCW 003, який має амінокислотну послідовність, вибрану з групи, що складається з SEQ ID NO: 8, SEQ ID NO: 10 і SEQ ID NO: 12, або їх інсектицидний білковий фрагмент.

5 Насіння також розглядається як ознака винаходу, в якому насіння отримують з таких трансгенних рослин, і таке насіння містить кількість полінуклеотидної конструкції, введеної в геном рослини, яка може бути визначена. Пилок, насіння, рослинні клітини потомства, рослинна тканина і товарні продукти, отримані з кожної такої трансгенної рослини, будуть містити кількість полінуклеотидної конструкції, яка може бути детектована.

10 Будь-який біологічний зразок, який містить, щонайменше, кількість полінуклеотидної конструкції, яка може бути детектована, яка кодує такий білок BCW 001, BCW 002 або BCW 003, буде знаходитися в межах об'єму винаходу.

15 Запропоновано композиції, які забезпечують інсектицидно-ефективну кількість білка BCW 001, 002 або 003 згідно цього винаходу і призначені для боротьби з видами лускокрилих-шкідників. Такі композиції можуть містити додатковий агент, який відрізняється від токсичного білка BCW. Такий агент також буде токсичним для тих самих видів лускокрилих, що і токсичний білок BCW. Додатковий агент має бути вибраний з групи агентів, яка складається з білків або поліпептидів, амінокислотна послідовність яких відрізняється від білка BCW, і також може бути агентом, який являє собою молекулу РНК, яка проявляє токсичну дію по відношенню до цільової комахи-шкідника (таку як дцРНК, мікроРНК або мала інтерферуюча РНК (міРНК)), а також може бути інсектицидною хімічною сполукою, такою як пиретрин, фосфорорганічний пестицид і т. п. В альтернативному варіанті, додатковий агент може бути будь-яким сумісним Cry або спорідненим токсичним білком, таким як інший Cry1A, Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae (але вони небажані, оскільки вони можуть не забезпечувати відповідні властивості для контролю стійкості), або Cry1B, Cry1C, Cry1D, Cry1E, Cry1F, Cry1G, Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab, Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, ET35, ET66, TIC400, TIC800, TIC807, TIC834, TIC853, TIC1415, VIP3A, VIP3Ab, інсектицидними білками Aхmі, інсектицидними білками DIG, eHІPs і білками VIP і будь-яким токсичним білком, відомим в цій галузі техніки, для забезпечення токсичних властивостей по відношенню до личинок виду совки-іпсилон або які можуть бути застосовані до інших видів-мішеней лускокрилих шкідників.

30 Такі композиції можуть також містити додаткові пестицидні агенти, які необов'язково є токсичними для одного і того ж шкідника-мішені, такі як додаткові агенти, вибрані з групи, що складається з Cry1C, Cry3A, Cry3B, Cry34, Cry35, Cry51Aa1, ET29, ET33, ET34, ET70, TIC407, TIC417, TIC431, TIC901, TIC1201, TIC3131, 5307, DIG-10, Aхmі184, Aхmі205 і AхmіR1.

35 Також передбачаються способи отримання насіння, в яких використовуються корисні пестицидні властивості білків згідно цього винаходу. Такі способи включають в себе полінуклеотидну конструкцію, спроектовану для експресії білка BCW 001, BCW 002 або BCW 003, або білка, який демонструє щонайменше близько 90 % ідентичності із вказаним білком, причому вказаний спосіб включає в себе посадку одного або більшої кількості насіння, яке містить полінуклеотид, який експресує один або більшу кількість білкових токсинів BCW згідно цього винаходу, вирощування рослин з такого насіння і потім збір врожаю такого насіння з рослин. Зібране насіння буде містити полінуклеотидну конструкцію і дасть рослини, які також будуть стійкі до зараження совкою-іпсилон.

45 Такими рослинами можуть бути кукурудза, бавовна, соя або будь-які інші рослини, сприйнятливі до видів лускокрилих-шкідників, які, як продемонстровано, контролюються білками згідно цього винаходу. Передбачається, що такі рослини є попередньо отриманими трансгенними рослинами, які отримують користь від ефектів токсичних властивостей білків згідно цього винаходу. Рослини кукурудзи, які потрапляють в цю категорію, включають в себе, але не обмежуються лише трансформантами, вибраними з групи, яка складається з: DKB89614-9, MON801, MON802, MON809, MON810, MON863, MON88017, MON89034, трансформанта 4114-3, трансформанта 5307, DAS59122-7, Bt10, Bt11, Bt176, CBH-351, DKB-83614-9, MIR162, MIR604, TC1507, TC6275, трансформанта 676, трансформанта 678, трансформанта 680, трансформанта 98140, DAS40278-9, DKB89790-5, MON21-9, HCEM485, MON832, MON87427, NK603, T14, T25 і VCO01981-5. Рослини сої, які потрапляють в цю категорію трансгенних рослин, вибирають з групи, яка складається з: MON87751, DAS81419-2, MON87701, A2704-12, A2704-21, A5547-127, A5547-35, CV127, DAS44406-6, DAS68416-4, DP356043, FG72, MON4032, ACS-GM003-1, MON87705, MON87708, MON89788, W62, W98 і GFM Cry1A. Трансгенні рослини бавовни, які потрапляють в цю категорію, вибирають з групи, яка складається з: DAS24236-5, DAS21023-5, трансформанта 31707, трансформанта 31803, трансформанта 31807, трансформанта 31808, трансформанта 42317, BNLA-601, COT102, COT67B, трансформанта 1, GHB119, GK12, MON15985, MLS9124, MON1076, MON531, MON757, T303-3, T304-40, SGK321,

трансформанта 19-51a, GHB614, LLCotton25, MON88701, MON88702, MON1445, MON1698 і MON88913. Трансгенні рослини цукрової тростини, які потрапляють в цю категорію, включають в себе трансформант NX1-1T цукрової тростини. Відомі в цій галузі техніки рослини рису, яким було б корисно мати конструкцію, яка кодує такі білкові токсини BCW, включають в себе трансформанти рису, вибрані з групи, яка складається з: LLRICE06, LLRICE601, LLRICE62, GM-A17054 і GM-A17054.

Додатково, запропоновані оброблені рослинні продукти, які містять кількість розкритих рекомбінантних полінуклеотидів яка може бути детектована. Такі оброблені продукти включають в себе, але не обмежуються лише цими: рослинну біомасу, масло, шрот, корм для тварин, борошно, пластівці, висівки, пух, лушпиння і оброблене насіння.

Також пропонуються способи отримання трансгенних рослин. Такі способи включають в себе введення рекомбінантного полінуклеотида в клітину рослини і відбір трансгенної рослини, яка експресує кількість рекомбінантного поліпептиду, який кодується рекомбінантним полінуклеотидом, яка інгібує комах.

Інші варіанти здійснення, ознаки та переваги винаходу будуть очевидні з наступного докладного опису, прикладів і формули винаходу.

#### КОРОТКИЙ ОПИС ГРАФІЧНИХ МАТЕРІАЛІВ

На Фіг. 1 зображений список видів лускокрилих комах-шкідників, які були протестовані в біоаналізі з токсичними білками BCW 001, BCW 002 і BCW 003. «+» вказує на смертність по відношенню до контролю у вигляді буфера; «-» вказує на відсутність статистично значущої спостережуваної смертності вище рівня контролю у вигляді буфера; "НД" означає, що не тестувалось з використанням відповідного токсичного білка. BCW 001 продемонстрував дію у вигляді смерті проти *Agrotis ipsilon* (совка-іпсилон), *Striacosta albicostsa*, *Helicoverpa zea*, *Ostrinia nubilalis* (вогнівка кукурудзяна), *Diatraea grandiosella*, *Trichoplusia ni*, *Pseudoplusia includens*, і не продемонстрував дії у вигляді смерті або затримку росту при тестуванні проти *Spodoptera frugiperda* (кукурудзяна листяна совка). *Diatraea saccharalis* не тестували з BCW 001. BCW 002 продемонстрував дію у вигляді смерті проти *Agrotis ipsilon* (совка-іпсилон), *Striacosta albicostsa*, *Helicoverpa zea*, і *Ostrinia nubilalis* (вогнівка кукурудзяна), і не продемонстрував дії у вигляді смерті або затримку росту при тестуванні проти *Spodoptera frugiperda* (кукурудзяна листяна совка). *Diatraea saccharalis*, *Diatraea grandiosella*, *Trichoplusia ni*, і *Pseudoplusia includens* не тестували з BCW 002. BCW 003 продемонстрував дію у вигляді смерті проти *Agrotis ipsilon* (совка-іпсилон), *Striacosta albicostsa*, *Helicoverpa zea*, *Ostrinia nubilalis* (вогнівка кукурудзяна), *Diatraea saccharalis*, *Diatraea grandiosella*, *Trichoplusia ni*, і *Pseudoplusia includens*, і не продемонстрував дії у вигляді смерті або затримку росту при тестуванні проти *Spodoptera frugiperda* (кукурудзяна листяна совка).

На Фіг. 2 зображено вирівнювання амінокислотної послідовності BCW 001 (SEQ ID NO: 2, верхня) проти BCW 002 (SEQ ID NO: 4, середня), і проти BCW 003 (SEQ ID NO: 6, нижня); зірочки під трьома послідовностями представляють відмінності по амінокислоті в обраній позиції щонайменше в одній з трьох різних послідовностей.

#### КОРОТКИЙ ОПИС ПОСЛІДОВНОСТЕЙ

SEQ ID NO: 1 являє собою нуклеотидну послідовність природного штаму EG4384 *B. thuringiensis*, яка кодує токсичний для лускокрилих білок BCW 001.

SEQ ID NO: 2 є встановленою амінокислотною послідовністю BCW 001 з відкритою рамкою зчитування, як вказано в SEQ ID NO: 1.

SEQ ID NO: 3 являє собою штучну послідовність, яка кодує химерний токсичний для лускокрилих білок BCW 002.

SEQ ID NO: 4 являє собою встановлену амінокислотну послідовність BCW 002 з відкритою рамкою зчитування, як вказано в SEQ ID NO: 3, в якій такий білок BCW 002 складається з домену I Cru1Ac, функціонально зв'язаного з доменами II і III BCW 001 (з амінокислотної позиції 258 до амінокислотної позиції 606, як вказано в SEQ ID NO: 2) і функціонально зв'язаного з доменом протоксину Cru1Ac з амінокислотної позиції 608 до 1177, як вказано в SEQ ID NO: 4 і SEQ ID NO: 6. SEQ ID NO: 4 (BCW 002) відрізняється від SEQ ID NO: 6 (BCW 003) лише в позиції 259, BCW 002 містить ізолейцин в цій позиції, BCW 003 містить треонін, як в BCW 001.

SEQ ID NO: 5 являє собою штучну послідовність, яка кодує химерний токсичний для лускокрилих білок BCW 003.

SEQ ID NO: 6 являє собою встановлену амінокислотну послідовність BCW 003 з відкритою рамкою зчитування, як вказано в SEQ ID NO: 3, в якій такий білок BCW 003 складається з домену I Cru1Ac, функціонально зв'язаного з доменами II і III BCW 001 (з амінокислотної позиції 258 до амінокислотної позиції 606, як вказано в SEQ ID NO: 2) і функціонально зв'язаного з доменом протоксину Cru1Ac з амінокислотної позиції 608 до 1177, як вказано в SEQ ID NO: 4 і

SEQ ID NO: 6. SEQ ID NO: 6 (BCW 003) відрізняється від SEQ ID NO: 4 (BCW 002) лише в позиції 259, BCW 002 містить ізолеїцин в цій позиції, BCW 003 містить треонін, як в BCW 001.

SEQ ID NO: 7 являє собою штучну послідовність, яка кодує білок BCW 001 для експресії в рослинах.

5 SEQ ID NO: 8 являє собою встановлену амінокислотну послідовність BCW 001, отриману з SEQ ID NO: 7.

SEQ ID NO: 9 являє собою штучну послідовність, яка кодує білок BCW 002 для експресії в рослинах.

10 SEQ ID NO: 10 являє собою встановлену амінокислотну послідовність BCW 002, отриману з SEQ ID NO: 9.

SEQ ID NO: 11 являє собою штучну послідовність, яка кодує білок BCW 003 для експресії в рослинах.

SEQ ID NO: 12 являє собою встановлену амінокислотну послідовність BCW 003, отриману з SEQ ID NO: 11.

#### 15 ДЕТАЛЬНИЙ ОПИС СУТІ ВИНАХОДУ

Альтернативою боротьбі з сільськогосподарськими шкідниками культурних рослин шляхом обприскування готовими формами, які містять інсектицидні білки, рослин/насіння, є введення полінуклеотидів, які кодують ці білки, в геном рослини для експресії в рослині або частинах рослини. Рослини, трансформовані такими полінуклеотидами, володіють стійкістю до комах, яку ці експресовані білки забезпечують у вигляді трансгенних ознак.

20 Щоб уникнути розвитку або обійти стійкість комах до використовуваних в даний час білків, для боротьби з лускокрилими необхідні нові білки з іншим способом дії (СД), а також широким спектром і ефективністю. Одним із способів вирішити цю проблему є сіквенування геномів *Bt* сподіваючись відкрити нові інсектицидні білки. Інший підхід полягає в обміні сегментів з різних білків *Bt* для створення нових химерних білків *Bt*, які мають пригнічуючі властивості по відношенню до комах. Ймовірність випадкового створення химерного білка з покращеними властивостями шляхом рекомбінування доменних структур численних природних інсектицидних кристалічних білків, відомих в цій галузі техніки, вкрай мала (дивіться, наприклад, A Strategy for Shuffling Numerous *Bacillus thuringiensis* Crystal Protein Domains; J. Economic Entomology, 97 (6): 1805-1813. 2004).

В цьому документі розкриті нуклеотидні послідовності, які кодують інсектицидні білки, позначені в цьому документі як білки BCW, які задовольняють потребу в альтернативному СД, забезпечують активність проти більш широкого спектру комах-шкідників і діють, затримуючи або запобігаючи розвитку стійкості, особливо при використанні в боротьбі з шкідниками совки-іспилон (BCW).

BCW 001 був відкритий як відкрита рамка зчитування, яка передбачає амінокислотну послідовність, яка має властивості білка типу *Cry1A*, після сіквенування генома штаму EG4384 *Bacillus thuringiensis*. Відкрита рамка зчитування BCW 001 (ORF) кодувала білок з 1180 амінокислот, і було передбачено, що білок володіє багатьма характеристиками білкових токсинів *Cry1*, включаючи ідентифіковану структуру доменів I, II і III і характерний домен протоксину типу *Cry1A* на карбоксикінцевій половині передбаченого білка. Передбачена амінокислотна послідовність домену I (залишки з 1 по близько 258 SEQ ID NO: 2) демонструє близько 67 % ідентичності з доменом I білкового токсину *Cry1Ac*. Передбачена амінокислотна послідовність домену II (залишки від близько 259 до близько залишку 459, як вказано в SEQ ID NO: 2) демонструє повну (100 %) ідентичність з доменом II *Cry1Ai2*. Передбачена амінокислотна послідовність домену III (залишки від близько 260 до близько 606, як вказано в SEQ ID NO: 2) демонструє близько 63 % ідентичності з відповідними залишками домену III в *Cry1Ah2*. Структура домену протоксину передбаченого білка BCW 001 (від близько залишку 607 до 1180, як вказано в SEQ ID NO: 2) демонструє близько 96 % ідентичності з відповідними залишками в *Cry1Aa9*. В цілому, цей передбачений повнорозмірний білок демонструє близько 83 % ідентичності амінокислотної послідовності з *Cry1Ai*, а передбачена область токсину від амінокислотної позиції 1 до близько залишку 607, як вказано в SEQ ID NO: 2, демонструє 76 % ідентичність з *Cry1Ai1*. Важко віднести цей новий токсичний білок до певного класу *Cry1A*, і тому комітету по номенклатурі *Bacillus thuringiensis* буде надана ця послідовність і він визначить, чи заслуговує цей білок свій власний окремий і новий клас.

55 Білок BCW 001 був отриманий з плазмідного вектора в безкрystalоносному штамі *Bacillus thuringiensis*, а препарати кристалів спор були протестовані проти різноманітних лускокрилих. Дивіться Фіг. 1, стовпчик 2 стосовно даних. Факти вказують на те, що цей білок не схожий на будь-який інший білок, з якими він розділяє джерело походження, тобто токсичні білки *Cry1Aa*, *Cry1Ah* або *Cry1Ai*, відомі в цій галузі техніки. Жоден з білків попереднього рівня техніки не

проявляє хоча б якусь помітну активність при тестуванні проти совки-іпсилон, проте цей новий білок BCW 001 був токсичним в біоаналізі при тестуванні проти совки-іпсилон і несподівано проявляв токсичні властивості при тестуванні проти групи інших лускокрилих, як викладено в Фіг. 1.

5 Зокрема, в біоаналізі в порівнянні з необробленим контролем-раціоном комах білок BCW 001 проявляв активність проти *Striacosta albicosta* ("WBC"), *Helicoverpa zea* ("CEW"), вогнівки кукурудзяної ("ECB", *Ostrinia nubilalis*), *Diatraea grandiosella* ("SWC"), *Pseudoplusia includes* ("SL"), *Trichoplusia ni* ("CLW"), і совки іпсилон 1-3 личинкової стадії ("BCW", *Agrotis ipsilon*).

10 Як описано додатково нижче, химерні токсичні білки отримували з використанням домену I Cру1Ac, і доменів II і III BCW 001 (тобто BCW 002 і BCW 003, як вказано в SEQ ID NO: 4 і SEQ ID NO: 6, відповідно), і ці химерні токсичні білки були введені в рослини кукурудзи і цукрової тростини. Обидва химерних білка проявляли активність в кукурудзі проти BCW, WBC, CEW і SWC. BCW 003 проявляв активність по відношенню до SCB в цукровій тростині.

15 Термін "білок BCW", як використовується в цьому документі, відноситься до будь-якого нового інгібуючого комах білка, який містить, який складається з, який є по суті гомологічним, або який отриманий з будь-якої інгібуючої комах поліпептидної послідовності: BCW 001 (SEQ ID NO: 2), BCW 002 (SEQ ID NO: 4) і BCW 003 (SEQ ID NO: 6) та їх сегментів, які пригнічують комах, або їх комбінацій, які забезпечують активність проти лускокрилих, зокрема, але не обмежуючись, BCW, WBC і/або SCB. Полінуклеотид, який кодує BCW 001, був отриманий з  
20 штаму EG4384. Основна токсична амінокислотна послідовність для BCW 001 відповідає амінокислотам від близько позиції 28 до близько позиції 606 і до позиції 618, як вказано в SEQ ID NO: 2, а основна токсична амінокислотна послідовність для BCW 002 і 003 відповідає амінокислотам від близько позиції 29 до близько позиції 607 і до близько позиції 619, як вказано в SEQ ID NO: 4 і SEQ ID NO: 6, відповідно.

25 В одному варіанті здійснення, білки, розкриті в цьому документі, є спорідненими по первинній структурі дельта-ендотоксину по довжині (близько 1176-1180 амінокислот), по довжині білка без протоксину (від близько 600 до близько 619 амінокислот), по довжині основної частини токсину (близько 591 амінокислот) або за наявністю щонайменше одного специфічного для BCW сегмента.

30 Ілюстративні білки були вирівняні один з одним з використанням алгоритму Clustal W, даючи в результаті число амінокислотних збігів для пар вирівнювання і відсоток амінокислотної ідентичності для пар вирівнювання з використанням таких параметрів за замовчуванням: Матриця вагів: Blosum; Штраф за відкриття пропуску: 10,0; Штраф за продовження пропуску: 0,05; Гідрофільні пропуски: Увімк.; Гідрофільні залишки: GPSNDQERK; Штрафи за пропуск конкретних залишків: Увімк. Алгоритм Clustal W описаний в Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice." *Nucleic Acids Research*, 22:4673-4680.

40 Інші алгоритми вирівнювання також доступні в цій галузі техніки і забезпечують результати, аналогічні тим, які отримані з використанням вирівнювання Clustal W.

Термін "близько" використовується в цьому документі для опису того, що такі межі сегментів можуть відрізнятись на 1, 2, 3, 4, 5, 10, 20, 25, 30 або 35 залишків, в залежності від послідовності вихідних білків і їх вирівнювання один з одним. Щоб додатково описати мінливість і конфігурацію різних сегментів, в Таблицях 2 і 3 наведені межі сегментів BCW 002 і 003, та  
45 інших активних проти совки-іпсилон химерних білків.

Термін "фрагмент" використовується в цьому документі для опису безперервних амінокислотних або нуклеотидних послідовностей, які коротші, ніж повна амінокислотна або нуклеотидна послідовність, яка описує токсичний білок BCW.

50 Словосполучення "інгібітор комах" та "інсектицид" використовуються в цьому документі як взаємозамінні і відносяться до білка, фрагменту білка, сегменту білка або полінуклеотиду, який призводить до будь-якого вимірюваного пригнічення життєздатності, росту, розвитку комах, розмноження комах, поведінки комах при годуванні, шлюбної поведінки комах і/або будь-якого вимірюваного зменшення несприятливих ефектів, викликаних поглинанням комахою цього білка, фрагмента білка, сегмента білка або полінуклеотида.

55 Терміни "біологічна активність", "активний", "активність", "ефективний", "ефективність" або їх варіації використовуються в цьому документі як взаємозамінні для опису ефектів білків згідно цього винаходу по відношенню до цільових комах-шкідників.

60 Культура - це самопоширювана рослина або рослина яку культивують, продукт якої збирають в певний момент її стадії розвитку. Необмежуваними прикладами таких продуктів є насінина, корбочка, лист, квітка, стебло, корінь або будь-яка їх частина.

Біологічний зразок, отриманий з будь-якої тканини рослини, бактерії, вірусу або вектора, які містять полінуклеотид або які експресують білок, як проілюстровано в цьому документі, такий як, але не обмежуючись лише цими: насіння, коробочка, лист, квітка, стебло, корінь або будь-яка його частина, і містять визначувану кількість полінуклеотиду, білка або обох.

5 Словосполучення "визначувана кількість" використовується в цьому документі для опису мінімальної кількості білка або полінуклеотиду, розкритого в цьому документі, яка може бути виявлена стандартними аналітичними способами, такими як, але не обмежуючись лише методами полімеразної ланцюгової реакції (ПЛР) та імуноферментного аналізу (ІФА), і т. п.

10 В одному варіанті здійснення, токсини, описані в цьому документі, споріднені по загальній функції і проявляють інсектицидну активність по відношенню до видів лускокрилих комах.

Сегменти BCW 001 забезпечують активність проти совки-іпсилон, коли містяться в химерах Cry1. Приклади сегментів BCW 001, які проявляють активність проти совки-іпсилон, наведені в SEQ ID NO: 2 від близько амінокислотної позиції 250 до близько 606, і більш конкретно від близько амінокислотної позиції 255 до близько 606. Химери Cry1, які містять цей амінокислотний сегмент, який відповідає доменам II і III токсичного білка BCW 001, часто також матимуть крім властивостей химерного білка токсичні властивості, пов'язані з контролем совки-іпсилон, і це було перевірено в каркасах Cry1A, Cry1B, Cry1C, Cry1D, Cry1E і Cry1F, в яких відповідна токсична конструкція мала компоненти домену II і III, заміщені цією ділянкою амінокислот з BCW 001, і в багатьох випадках активність проти совки-іпсилон несподівано зберігалася в химерній конструкції (дані не показані).

20 В одному аспекті винаходу, шкідник, якого контролюють застосуванням токсичним білком BCW, є, зокрема, комахою-шкідником ряду Lepidoptera, включаючи дорослих особин, лялечок, личинок і личинок які щойно з'явилися.

25 Комахи ряду лускокрилих (Lepidoptera) включають в себе, але не обмежуються лише цими: комах з родини совок (Noctuidae) (арміvormз, катvormз, луперз, геліотінез), наприклад кукурудзяну листяну совку (*Spodoptera frugiperda*), совку помідорну (*Spodoptera exigua*), *Mamestra configurata*, совку-іпсилон (*Agrotis ipsilon*), *Trichoplusia ni*, *Pseudoplusia includens*, *Anticarsia gemmatalis*, *Hypona scabra*, *Heliothis virescens*, *Agrotis subterranea*, *Agrotis orthogonia*, *Pseudaletia unipuncta*; точильників, чехлоносиків та інших комах з родини вогнівок (*Pyralidae*), наприклад вогнівку кукурудзяну (*Ostrinia nubilalis*), *Amyelois transitella*, *Crambus caliginosellus*, *Herpetogramma licarsisalis*, *Homoeosoma electellum*, *Elasmopalpus lignosellus*; комах з родини листокруток (*Tortricidae*), наприклад яблуневу плодожерку (*Cydia pomonella*), *Endopiza viteana*, східну плодожерку (*Grapholita molesta*), *Suleima helianthana*; і багатьох інших економічно важливих лускокрилих (наприклад, капустяну міль (*Plutella xylostella*), *Pectinophora gossypiella*, і непарного шовкопряда (*Lymantria dispar*)). Інші комахи-шкідники ряду лускокрилих (Lepidoptera) 35 включають в себе, наприклад, *Alabama argillacea*, *Archips argyrospila*, *Archips rosana* та інші види *Archips*, *Chilo suppressalis* (вогнівку жовту рисову), *Snaphalocrocis medinalis*, *Crambus caliginosellus*, *Crambus teterrellus*, *Diatraea grandiosella*, *Diatraea saccharalis*, *Earias insulana*, *Earias vittella*, *Helicoverpa armigera* (совку бавовняну), *Helicoverpa zea*, *Heliothis virescens*, *Herpetogramma licarsisalis*, *Lobesia botrana* (гронову листокрутку), *Pectinophora gossypiella*, *Phyllocnistis citrella*, *Pieris brassicae* (капустяну міль), *Pieris rapae* (білана ріпакового), *Plutella xylostella* (капустяну міль), *Spodoptera exigua*, *Spodoptera litura*, *S. frugiperda*, и *Tuta absoluta* (томатну мінуючу міль).

45 Розкриті в цьому документі білки також можуть бути використані для отримання антитіл, які специфічно зв'язуються з BCW-специфічними токсичними білками, і можуть використовуватися для скринінгу і пошуку інших представників родини токсинів BCW.

В одному варіанті здійснення, ілюстративні полінуклеотиди, які кодують білки, які пригнічують комах, споріднені з BCW 001, вказані в SEQ ID NO: 1, 3, 5, 7, 9 і 11. Нуклеотидні послідовності, які кодують ці білки, можна використовувати в якості зондів і праймерів для скринінгу для ідентифікації інших представників родини з використанням способів термічної або ізотермічної ампліфікації і/або гібридизації та інших способів ідентифікації, відомих фахівцям в цій галузі техніки.

55 Згідно аспекту винаходу, запропоновано способи виявлення споріднених білків, і такі способи включають в себе сіквенування геномів Bt, складання таких сіквенувань, ідентифікацію та клонування генів Bt, які кодують такі пестицидні білки, а також експресію і тестування нових білків Bt для аналізу пестицидної активності. В іншому аспекті винаходу використовуються молекулярні способи для конструювання і клонування комерційно корисних білків, які включають в себе химери білків з родини пестицидних білків, наприклад, химери можуть бути зібрані з сегментів токсичних білків BCW для отримання додаткових варіантів здійснення. 60 Розкриті білки можуть бути вирівняні один з одним та з іншими пестицидними білками Bt, і

можуть бути ідентифіковані сегменти кожного такого білка, які корисні для заміщення між вирівняними білками, що призводить до конструювання химерних білків. Такі химерні білки можуть бути проаналізовані - з допомогою біоаналізу з шкідником, і охарактеризовані на наявність підвищеної біологічної активності або розширення спектра шкідників-мішеней, у порівнянні з вихідними білками, з яких було отримано кожен такий сегмент химери. Пестицидну активність поліпептидів можна додатково сконструювати у вигляді активності проти конкретного шкідника або більш широкого спектра шкідників шляхом заміни доменів або сегментів сегментами з інших білків.

В одному варіанті здійснення, білки, розкриті в цьому документі, включають в себе функціонально еквівалентні фрагменти (N- або C-кінцеві делеції) білків, розкритих в цьому документі.

BCW токсичні білки запропоновані в цьому документі. У деяких варіантах здійснення, можуть бути виділені, можуть бути представлені в композиції, в трансгенному мікроорганізмі або в трансгенній рослині токсини, споріднені з BCW 001. В цьому варіанті здійснення, білки BCW 002 і, зокрема, білки BCW 003 забезпечують інгібуючу активність по відношенню до лускокрилих, особливо інгібуючу активність по відношенню до совки-іпсилон і/або *Diatraea saccharalis*. Посилання в цій заявці на "виділений білок", або еквівалентний термін чи словосполучення, призначене для позначення того, що білок являє собою молекулу, яка присутня окремо або в комбінації з іншими сполуками, але не в її природному середовищі. Наприклад, токсичні білки згідно цього винаходу і т. п., які природним чином виявляються в організмі, не вважаються "виділеними", якщо вони знаходяться в організмі, в якому вони знаходяться в природному стані. Проте, кожен з них буде "виділений" в рамках цього розкриття винаходу, якщо білок не знаходиться в організмі, в якому він знаходиться в природному стані.

Термін "функціонально зв'язаний", як використовується в цьому документі, відноситься до з'єднання послідовностей нуклеїнових кислот або амінокислотних послідовностей, так що одна послідовність може забезпечувати потрібну функцію або сумісну або корисну властивість для зв'язаної послідовності.

Пептиди, поліпептиди і білки, біологічно функціонально еквівалентні BCW 001, BCW 002 і BCW 003, включають в себе, але не обмежуються лише цими: амінокислотні послідовності, які містять консервативні амінокислотні заміни в цих білкових послідовностях токсичного білка BCW. В таких амінокислотних послідовностях одна або більша кількість амінокислот в послідовності заміщені іншою амінокислотою (амінокислотами), що призводить до безмовної або консервативної заміни в амінокислотній послідовності.

Хоча розкриті поліпептиди, які інгібують комах, бажано містять білкову послідовність BCW 001, BCW 002 або BCW 003, фрагменти і варіанти цієї послідовності, які мають таку ж або подібну інгібуючу активність по відношенню до комах, як і активність такого білка, який інгібує комах, також розкриті в цьому документі. Наприклад, суміжні послідовності з щонайменше 30, 35, 38, 40, 45, 50, 55, 60, 65, 70, 75, 100, 150, 200, 500, 550 або більшої кількості амінокислот в спорідненому BCW токсичному білку з інгібуючою активністю по відношенню до комах. В іншому варіанті здійснення, фрагменти спорідненого BCW токсичного білка з інгібуючою активністю по відношенню до комах можуть містити амінокислотні заміни, делеції, вставки або додавання в послідовності токсичного білка BCW.

В одному варіанті здійснення, поліпептид, який інгібує комах, містить сегмент, який інгібує комах, від близько залишку 28 до близько залишку 618 білкової послідовності BCW 001, як вказано в SEQ ID NO: 2. Необмежуючі приклади включають в себе будь-яку з SEQ ID NO: 2, 4 або 6, або більш короткі фрагменти, або варіанти, які мають таку ж або подібну інгібуючу активність по відношенню до комах, що і активність такого конкретного білка BCW 001, самі по собі або в функціональному зв'язку в химерним білком. В іншому варіанті здійснення, сегменти, які мають суміжні амінокислотні послідовності, щонайменше, близько 38 або більшої кількості амінокислот в будь-якій з SEQ ID NO: 2, 4 або 6, які володіють інгібуючою активністю по відношенню до комах, також дають функціональний інсектицидний білок. Токсичні фрагменти BCW 001, які інгібують комах, також можуть містити сегменти, щонайменше, з 30, 35, 38, 40, 45, 50, 100, 150, 200, 500, 550, 555, 560, 565, 570, 572, 574, 580 або 585 амінокислотних залишків з області 591 амінокислотного залишку, яка відповідає залишкам від близько 28 до близько 618 послідовностей, будь-якої з SEQ ID NO: 2, 4 або 6.

У деяких варіантах здійснення, фрагменти зрілого білка BCW 001 (зрілий позначає протоксинову форму білка, яка складається з 1180 амінокислот, розщеплюється шляхом протеолізу в кишечнику комах-шкідника, щоб вивільнити лише N-кінцеве ядро токсину до залишку 607 включно близько залишку 618, вивільняючи активний сегмент токсину, який містить, більше або менше, залишки з 1 по 606 або будь-яку кількість залишків від близько 5 до

близько 618, як вказано в SEQ ID NO: 2, за умови, що сегмент який вивільнився демонструє токсичні властивості при дії на личинки совки-іпсилон) можуть бути вкороченими формами, в яких одна або більша кількість амінокислот видалені з N-кінця, C-кінця, середини білка або їх комбінацій, з активністю у вигляді пригнічення комах. Ці фрагменти можуть бути природними або синтетичними варіантами BCW 001 і зберігати інгібуючу активність по відношенню до комах токсичного білка BCW. У деяких варіантах здійснення, фрагменти зрілих білків BCW 001, BCW 002 або BCW 003 проявляють пестицидну активність, яку мають вихідні молекули білка, з яких вони отримані. Описаний в цьому документі фрагмент або варіант може додатково містити позначений в цьому документі домен, який відповідає за пестицидну активність білка.

Вкорочене похідне, яке володіє інгібуючою активністю по відношенню до комах, є токсичним білком BCW, який відповідає залишкам від близько 28 до близько 606, або до близько 618, послідовності токсичного білка BCW 001, як вказано в SEQ ID NO: 2, або залишків від близько 29 до близько 607 і до залишку 619 токсичного білка BCW, як вказано в SEQ ID NO: 4 або SEQ ID NO: 6.

Ще в одному варіанті здійснення, вкорочені N-кінцеві делеційні мутації включають в себе, але не обмежуються токсичними білками BCW 001, в яких відсутні амінокислотні залишки або з N-кінця, і/або з C-кінця токсичної частини без протоксину, або токсичне ядро токсичних білків BCW. Наприклад, від 1 до 6 N-кінцевих амінокислотних залишків токсичного ядра білка BCW 001, яке відповідає залишкам з 28 по 618 SEQ ID NO: 2 або залишкам з 29 по 619 SEQ ID NO: 4 або 6, можуть бути видалені. Вкорочені C-кінцеві делеційні мутанти токсичного білка BCW, який відповідає залишкам з 28 по 618 послідовності SEQ ID NO: 2 або залишкам з 29 по 619 послідовності SEQ ID NO: 4 або 6, включають в себе, але не обмежуються токсичними білками BCW, в яких відсутні від 1 до 6 C-кінцевих амінокислотних залишків. В інших варіантах здійснення, токсичний білок BCW з відповідними залишками 28-618 SEQ ID NO: 2 або відповідними залишками 29-619 SEQ ID NO: 4 або 6, може мати як N-кінцеве вкорочення від 1 до 6 амінокінцевих залишків, так і C-кінцеве вкорочення від 1 до 6 карбоксикінцевих залишків.

У деяких варіантах здійснення, окремі сегменти з 1 по 6 білка CPR24719 або комбінація з сегментів з 1 по 6, які надають активність проти совки-іпсилон білку, який відрізняється від CPR24719-1, також можуть демонструвати таку ж або подібну функцію.

Фрагменти і варіанти токсичного білка BCW, розкриті в цьому документі, можуть мати близько 62 % або більше ідентичності послідовності, близько 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 92 %, 95 %, 96 %, 97 %, 98 % або більше ідентичності послідовності і близько 99 %, 99,5 %, 100 % ідентичності амінокислотної послідовності з відповідними сегментами зрілого токсичного білка BCW, який має відповідні амінокислотні послідовності, показані як залишки 28-618 SEQ ID NO: 2 або залишки 29-619 SEQ ID NO: 4 або 6.

Варіант здійснення винаходу включає в себе рекомбінантні полінуклеотидні композиції, які кодують токсичні білки BCW. Наприклад, токсичні білки BCW можуть експресуватися з допомогою рекомбінантних конструкцій ДНК, в яких виділена полінуклеотидна молекула з відкритою рамкою зчитування, яка кодує білок, функціонально зв'язана з такими елементами, як промотор і будь-яким іншим регуляторним елементом, який функціонує для експресії в системі, для якої призначена конструкція. Посилання в цій заявці на "виділену молекулу ДНК", або еквівалентний термін чи фразу, призначене для позначення того, що молекула ДНК являє собою молекулу, яка присутня окремо або в комбінації з іншими сполуками, але не в її природному середовищі. Наприклад, елементи нуклеїнової кислоти, такі як кодуюча послідовність, послідовність інтрона, лідерна послідовність яка не транскрибується, послідовність промотора, послідовність термінації транскрипції і т. п., які природно виявляють в ДНК генома організму, не рахуються "виділеними" до тих пір, поки елемент знаходиться в геномі організму і в тому місці в геномі, в якому його природно виявляють. Проте, кожен з цих елементів, і частин таких елементів, будуть "виділені" в рамках цього розкриття винаходу, якщо елемент знаходиться не в геномі організму, і не місці в геномі, в якому його природно виявляють. Подібним чином, нуклеотидна послідовність, яка кодує інсектицидний білок або будь-який інсектицидний варіант такого білка який зустрічається в природі, буде являти собою виділену нуклеотидну послідовність, до тих пір поки нуклеотидна послідовність не знаходиться в ДНК бактерії, в якій природно виявляють послідовність, яка кодує білок. Синтетична нуклеотидна послідовність, яка кодує амінокислотну послідовність інсектицидного білка який зустрічається в природі, буде вважатися виділеною для цілей цього розкриття винаходу. Для цілей цього розкриття винаходу, будь-яка трансгенна нуклеотидна послідовність, тобто нуклеотидна послідовність ДНК, вставлена в геном клітин рослини чи бактерії, або присутня у позахромосомному векторі, буде розглядатися як виділена нуклеотидна послідовність, незалежно від того, чи є вона присутньою в плазміді або аналогічній структурі, яка

використовується для трансформації клітин, в геномі рослини або бактерії, або присутня в кількостях, які виявляються, в тканинах, потомстві, біологічних зразках або товарних продуктах, отриманих з рослини або бактерії. Необмежуючі приклади включають в себе функціональні в рослинах промотори, функціонально зв'язані з послідовностями, які кодують токсичний білок BCW, для експресії білка в рослинах, або Vt-функціональні промотори, функціонально зв'язані з послідовностями, які кодують токсичний білок BCW, для експресії білка в Vt. Інші елементи, які можуть бути функціонально зв'язані з послідовностями, які кодують токсичний білок BCW, включають в себе, але не обмежуються лише цими: енхансери, інтрони, лідерні послідовності, кодовані теги іммобілізації білка (HIS-ter), кодовані пептиди субклітинної транслокації (наприклад, пластидні транзитні пептиди, сигнальні пептиди), кодовані поліпептидні сайти ферментів посттрансляційних модифікацій, рибосомні сайти зв'язування і сайти-мішені РНКі.

Як застосовується в цьому документі, термін "рекомбінантна молекула ДНК" є молекулою ДНК, яка містить комбінацію молекул ДНК, які б не зустрічалися в природі разом без втручання людини. Наприклад, рекомбінантна молекула ДНК може являти собою молекулу ДНК, яка складається щонайменше з двох молекул ДНК, гетерологічних по відношенню одна до одної, молекулу ДНК, яка містить послідовність ДНК, яка відрізняється від послідовностей ДНК, які існують в природі, або молекулу ДНК, яка була включена в ДНК клітини-господаря шляхом генетичної трансформації або редагування генів. Аналогічним чином, "рекомбінантна молекула білка" являє собою молекулу білка, яка містить комбінацію амінокислот, які б не зустрічалися в природі разом без втручання людини. Наприклад, рекомбінантна молекула білка може бути молекулою білка, яка складається щонайменше з двох амінокислотних молекул, гетерологічних по відношенню одна до одної, молекулою білка, яка містить амінокислотну послідовність, яка відрізняється від амінокислотних послідовностей, які існують в природі, або молекулою білка, яка експресується в клітині-господарі в результаті генетичної трансформації клітини-господаря або через редагування генів геному клітини-господаря.

Ілюстративні рекомбінантні полінуклеотидні молекули, запропоновані в цьому документі, включають в себе, але не обмежуються лише цими: SEQ ID NO: 7, SEQ ID NO: 9 і SEQ ID NO: 11, а також кожен з нуклеотидних сегментів, вказаних в SEQ ID NO: 3 і SEQ ID NO: 5 і що кодують відповідні поліпептиди або білки, які мають амінокислотну послідовність, як вказано в SEQ ID NO: 2 (BCW 001), SEQ ID NO: 4 (BCW 002), SEQ ID NO: 6 (BCW 003) і SEQ ID NO: 8, SEQ ID NO: 10 і SEQ ID NO: 12. Розкриті в цьому документі кодони рекомбінантної полінуклеотидної молекули, яка кодує білки, можуть бути замінені синонімічними кодонами (які також називають безмовною заміною). Також запропоновані рекомбінантні полінуклеотиди, які кодують будь-який з білків-варіантів токсину BCW, описаних в цьому документі.

Конструкція рекомбінантної ДНК, яка містить послідовності, які кодують токсичний білок BCW, може також додатково містити область ДНК, яка кодує один або більшу кількість агентів, які пригнічують комах, які можуть бути сконфігуровані для одночасної експресії або ко-експресії з послідовністю ДНК, яка кодує токсичний білок BCW, білка, який відрізняється від токсичного білка BCW, молекули дцРНК яка інгібує комах або інсектицидною хімічною сполукою. Необмежуючими прикладами інсектицидних хімічних сполук є органохлориди, органофосфати і карбамати, піретроїди, неонікотиноїди і ріаноїди.

Конструкція рекомбінантної ДНК може бути зібрана таким чином, що всі білки або молекули дцРНК експресуються одним промотором, або кожний білок або молекули дцРНК знаходиться під контролем окремого промотора або будь-якої їх комбінації. Розкриті в цьому документі білки можуть бути експресовані з мультигенної системи експресії, в якій один або більша кількість білків, описаних в цьому документі, експресуються із спільного нуклеотидного сегмента, який також містить інші відкриті рамки зчитування і/або промотори в залежності від типу обраної системи експресії. Наприклад, бактеріальна мультигенна система експресії може використовувати один промотор для керування експресією мульти-зв'язаних/тандемних відкритих рамок зчитування з одного оперону. В іншому прикладі, в рослинній мультигенній системі експресії можуть використовуватися мульти-зв'язані експресійні касети, кожна з яких експресує різний білок або різний агент, такий як одна або більша кількість молекул дцРНК. В ще одному прикладі, в рослинній мультигенній системі експресії можуть використовуватися мульти-незв'язані експресійні касети, кожна з яких експресує різний білок або різний агент, такий як одна або більша кількість молекул дцРНК. Промотор для застосування в описаній в цьому документі рекомбінантній нуклеїновій кислоті може містити повну послідовність промотора або будь-який його варіант або фрагмент, який володіє активністю промотора або активністю генної регуляції.

Конструкція рекомбінантного полінуклеотида або рекомбінантної ДНК, яка містить послідовність, яка кодує токсичний білок BCW, може бути доставлена в клітини-господарі

векторами, наприклад, плазмідною, бакуловірусом, штучною хромосоною, віріоном, космідною, фагмідною, фагом або вірусним вектором. Такі вектори можуть бути використані для досягнення стабільної або тимчасової експресії послідовності, яка кодує токсичний білок BCW, в клітині-господарі, або подальшої експресії в поліпептид. Екзогенна конструкція рекомбінантного полінуклеотида або рекомбінантної ДНК, яка містить послідовність, яка кодує токсичний білок BCW, і яка вводиться в клітину-господаря, називається в цьому документі "трансгеном".

Також в цьому документі запропоновані трансгенні бактерії, трансгенні клітини рослин, трансгенні рослини, гриби і дріжджі, і частини трансгенних рослин, які містять будь-який рекомбінантний полінуклеотид, який експресує будь-яку одну або більшу кількість послідовностей, які кодують токсичний білок BCW, запропонований в цьому документі. Термін "бактеріальна клітина" або "бактерія" може включати в себе, але не обмежується клітиною *Agrobacterium*, *Bacillus*, *Escherichia*, *Salmonella*, *Pseudomonas* или *Rhizobium*. Термін "клітина рослини" або "рослина" може включати в себе, але не обмежується лише клітиною рослини або рослиною: люцерни, банана, ячменю, квасолі, брокколі, капусти, декоративної, моркви, маніоки, рицини, цвітної капусти, селери, нуту, китайської капусти, цитрусових, кокосової пальми, кави, кукурудзи, конюшини, бавовни, гарбузових, огірка, псевдотсуги Мензиса, баклажана, евкалипта, льону, часнику, винограду, хмелю, цибулі-порей, салату-латуку, сосни ладанної, проса, дині, горіха, вівса, оливкового дерева, ріпчастої цибулі, декоративних рослин, пальмових, пасовищних трав, гороху, арахісу, перцю, голубиноного гороху, соснових, картоплі, тополі, гарбуза, сосни променистої, редьки, ріпаку, рису, кореневищ, жита, дикого шафрану, чагарникових, сорго, сосни південної, сої, шпинату, гарбузових, полуниці, цукрового буряка, цукрової тростини, соняшнику, кукурудзи цукрової, амбрового дерева, батату, проса прутувидного, чаю, тютюну, помідора, тритикале, дернової трави, кавуна і пшениці. У деяких варіантах здійснення, запропоновані трансгенні рослини і частини трансгенних рослин, одержані з клітини трансгенної рослини. У деяких варіантах здійснення, трансгенні рослини можуть бути отримані з трансгенного насіння, шляхом розмноження, розрізання, відламування, подрібнення або іншим чином від'єднання частини від рослини. У деяких варіантах здійснення, частина рослини може бути насінням, коробочкою, листом, квіткою, стеблом, коренем або будь-якою їх частиною, або фрагментом частини трансгенної рослини який не регенерується. Як використовується в цьому контексті, частина трансгенної частини рослини "що не регенерується" є частиною, яка не може бути стимульована для утворення цілої рослини, або яка не може бути стимульована для утворення цілої рослини, здатної до статевого і/або безстатевого розмноження. У деяких варіантах здійснення, частина частини рослини, яка не регенерується, є частиною трансгенного насіння, насінневої коробочки, листка, квітки, стебла або кореня.

В цьому документі також запропоновані способи отримання трансгенних рослин, які містять інгібуючу комах або лускокрилих кількість токсичного білка BCW. Такі рослини можуть бути отримані шляхом введення рекомбінантного полінуклеотида, який кодує будь-який з токсичних білків BCW, представлених в цьому документі, в клітину рослини, і відбір рослини, отриманої з вказаної клітини рослини, яка експресує кількість токсичного білка BCW, що пригнічує комах або лускокрилих. Рослини можуть бути отримані з рослинних клітин методами регенерації, з насіння, пилку або меристеми.

Способи трансформації рослин відомі в цій галузі техніки. Наприклад, *Agrobacterium*-опосередкована трансформація, описана в публікаціях патентних заявок США № 2009/0138985A1 (соя), № 2008/0280361A1 (соя), № 2009/0142837A1 (кукурудза), № 2008/0282432 (бавовна), № 2008/0256667 (бавовна), № 2003/0110531 (пшениця), № 2001/0042257 A1 (цукровий буряк), патентах США № 5750871 (канола), № 7026528 (пшениця), і № 6365807 (рис), і в *Agencibia et al. (1998) Transgenic Res. 7: 213-222* (цукрова тростина).

В цьому документі також запропоноване застосування трансгенної рослини, яка експресує кількість токсичного білка BCW, що пригнічує комах або лускокрилих, для боротьби із зараженням комахами або лускокрилими. Будь-яка з вищезгаданих трансгенних рослин може бути використана в способах захисту рослини від зараження комахами або лускокрилими, наведеними в цьому документі. Способи отримання трансгенних рослин, які експресують активні проти лускокрилих білки, такі як білки Cry1A (патент США № 5880275), Cry1B (заявка на патент США № 10/525318), Cry1C (патент США № 6033874), Cry1F, химери Cry1A/F (патенти США № 7070982; 6962705 та 6713063) і білок Cry2Ab (патент США № 7064249) є добре охарактеризованих.

В цьому документі також запропоноване застосування будь-якої з вищевказаних трансгенних клітин-господарів для продукування токсичного білка BCW.

Додаткові аспекти винаходу включають в себе антитіла і способи виявлення

полінуклеотидів, які кодують токсичні білки BCW, або способи розрізнення їх фрагментів та сегментів, способи ідентифікації додаткових членів родини білків, які пригнічують комах, готові форми і способи контролю росту і/або зараження комахами, і також способи для надання такого контролю рослинам та іншим господарям-реципієнтам.

5 У деяких варіантах здійснення, рослинний продукт може включати в себе товарні або інші комерційні продукти, отримані з трансгенної рослини або частини трансгенної рослини, причому товарні або інші продукти можуть відслідковуватись з допомогою комерції шляхом виявлення нуклеотидних сегментів, або експресованої РНК або білків, які кодують або містять характерні частини токсичного білка BCW. Такі товарні або інші комерційні продукти включають в себе, але  
10 не обмежуються лише цими: частини рослини, біомасу, масло, шрот, цукор, корм для тварин, борошно, пластівці, висівки, пух, оброблене насіння і насіння.

Також при цьому запропоновані оброблені рослинні продукти, причому вказаний оброблений продукт містить детектовану кількість рекомбінантного полінуклеотида, який кодує токсичний білок BCW, його сегмент, його фрагмент, який інгібує комах, або будь-яку його  
15 характерну частину. У деяких варіантах здійснення, оброблений продукт вибирають з групи, що складається з: рослинної біомаси, масла, шроту, корму для тварин, борошна, пластівців, висівок, пуху, лушпиння і обробленого насіння. У деяких варіантах здійснення, оброблений продукт є не придатним для регенерації.

Також в цьому документі запропоновані способи боротьби з комахами. У деяких варіантах здійснення, зараження лускокрилими культурних рослин є контрольованим. Такі способи можуть включати в себе вирощування рослини, яка містить кількість токсичного білка BCW, яка інгібує комах або лускокрилих. В окремих випадках здійснення, такі способи можуть додатково включати в себе будь-який один або більшу кількість з: (i) нанесення будь-якої композиції, яка містить або кодує токсичний білок BCW, на рослину або насіння, яке дає рослина; і/або (ii)  
25 трансформацію рослини або рослинної клітини, яка дає рослину, полінуклеотидом, який кодує токсичний білок BCW. У деяких варіантах здійснення, рослина являє собою тимчасово або стабільно трансформовану трансгенну рослину, яка містить трансген, який експресує кількість токсичного білка BCW, яка інгібує комах або лускокрилих. У деяких варіантах здійснення, рослина являє собою не трансгенну рослину, на яку була нанесена композиція, яка містить токсичний білок BCW. У деяких варіантах здійснення таких способів, рослина являє собою рослину кукурудзи або цукрової тростини. У деяких варіантах здійснення, вид *Lepidoptera* являє собою *Agrotis ipsilon*. У деяких варіантах здійснення, вид *Lepidoptera* являє собою *Diatraea saccharalis*. У деяких варіантах здійснення, вид *Lepidoptera* знаходиться на полі з культурною  
30 рослиною.

Збільшення кількості білків, розкритих в цьому документі, або в рослинах, або з допомогою технологічного процесу, може включати в себе культивування рекомбінантних Bt-клітин в умовах для експресії/продукування рекомбінантного поліпептиду/білків. Такий процес може включати в себе отримання шляхом висушування, ліофілізації, гомогенізації, екстракції, фільтрації, центрифугування, седиментації або концентрування культури рекомбінантних Bt-клітин, які експресують/продукують вказаний рекомбінантний поліпептид. Такий процес може привести до отримання Bt-клітинного екстракту, клітинної суспензії, гомогената клітин, лізата клітин, супернатанта клітин, фільтрату клітин або осаду клітин. Отримуючи рекомбінантні поліпептиди/білки, продуковані таким чином, композиція, яка містить рекомбінантні поліпептиди/білки, може містити бактеріальні клітини, бактеріальні спори і параспоральні внутрішньоклітинні тільця, і може бути приготовлена для різних застосувань, включаючи розпорошені сільськогосподарські продукти, які пригнічують комах, або у вигляді готових форм, які пригнічують комах в харчових біоаналізах.  
35

В одному варіанті здійснення, композиція/готова форма, яка пригнічує комах, яка містить розкритий рекомбінантний поліпептид/білок, може додатково містити, щонайменше, один додатковий поліпептид, який демонструє інгібуючу активність по відношенню до тих же видів лускокрилих комах, але відрізняється від рекомбінантного поліпептиду, щоб забезпечити зменшення кількості випадків появи стійкості у лускокрилих комах до токсичного білка BCW або іншої композиції, яка пригнічує лускокрилих. Такий поліпептид вибраний з групи, яка складається з білка, який пригнічує комах, молекули дцРНК, яка пригнічує комах, і хімічної сполуки. Один приклад використання таких рибонуклеотидних послідовностей для боротьби з комахами-шкідниками описаний в публікації заявки на патент США № 2006/0021087. Приклади інших таких композицій включають в себе, але не обмежуються лише цими: Cry1A (патент США № 5880275), Cry1Ab, Cry1Ac, Cry1Ae, Cry1B (заявка на патент США № 10/525318), Cry1C (патент США № 6033874), Cry1E, Cry1F і химери Cry1A/F (патенти США № 7070982; № 6962705 і № 6713063), Cry1G, Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab (патент США №  
50  
55  
60

7064249), Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, ET35, ET66, TIC400, TIC800, TIC807, TIC834, TIC853 і TIC1415. Іншими необмежуваними прикладами є активні проти лускокрилих білки VIP, Axmi і DIG, такі як, але не обмежені Vip3A, VIP3Ab, AXMI-184, AXMI-196, DIG-3, DIG-4, DIG-5 і DIG-11, які можуть бути скомбіновані з білками, описаними в цьому документі.

В інших варіантах здійснення, така композиція може додатково містити, щонайменше, один додатковий поліпептид, який демонструє інгібує по відношенню до комах активність, що не пригнічується іншим токсичним білком BCW, який інгібує комах, для розширення спектру отриманого інгібування комах. Наприклад, для боротьби з жорсткокрилими комахами-шкідниками можна використовувати комбінації токсичних білків BCW, які пригнічують комах, з білками, активними проти жорсткокрилих, такими як, але не обмежуючись лише цими: варіанти Cry1C, варіанти Cry3A, Cry3Bb (патент США № 6501009), Cry34/35, 5307, Axmi184, Axmi205, AxmiR1, TIC407, TIC417, TIC431, TIC901, TIC1201, TIC3131, DIG-10 і eHIP (публікація заявки на патент США № 2010/0017914).

В цій галузі техніки була задокументована можливість комах розвивати стійкість до певних інсектицидів. Одна зі стратегій контролю стійкості комах полягає в використанні трансгенних культур, які експресують два різних агента, що пригнічують комах, які мають різні способи дії. Отже, будь-яка комаха, яка володіє стійкістю до одного з агентів, що пригнічують комах, може контролюватися іншим агентом, який пригнічує комах. Інша стратегія контролю стійкості комах передбачає використання рослин, які не захищені від видів лускокрилих шкідників, щоб створити для них екологічний притулок. Один конкретний приклад описаний в патенті США № 6551962.

Інші варіанти здійснення, розкриті в цьому документі, включають в себе пестицидні хімічні сполуки місцевого застосування, які призначені для боротьби з шкідниками, яких також контролюють з допомогою білків, розкритих в цьому документі, для використання з білками при обробці насіння, розпиленні, поливі, або зрошенні готовими формами, які можна безпосередньо вносити в ґрунт (ґрунтове зрошення), що застосовуються до вирощуваних рослин, які експресують білки, розкриті в цьому документі, або приготовані для застосування до насіння, яке містить один або більшу кількість трансгенів, що кодують один або більшу кількість розкритих білків. Такі готові форми для застосування при обробці насіння можна наносити з різними клейкими речовинами і речовинами для підвищення клейкості, відомими в цій галузі техніки. Такі готові форми можуть містити пестициди, які є синергетичним по СД з розкритими білками, так що готова форма пестицидів має СД, який відрізняється, для боротьби з тими ж або аналогічними шкідниками, які можуть контролюватися розкритими білками, або так, що дія таких пестицидів контролює ширший діапазон господарів-шкідників, таких як види лускокрилих або членистоногих (Hemiptera), або інші види шкідників рослин, такі як види жуків (Coleoptera), які не контролюються ефективно.

Вищевказана композиція/готова форма може додатково містити прийнятний з точки зору сільськогосподарства носій, такий як приманка, порошок, пил, пелети, гранули, спрей, емульсія, колоїдна суспензія, водний розчин, препарат спор/кристалів Bacillus, добриво для насіння, рекомбінантна рослинна клітина/рослинна тканина/насіння/рослина, трансформовані для експресії одного або більшої кількості білків, чи бактерія, трансформована для експресії одного або більшої кількості білків. Залежно від рівня інгібування комах або інсектицидного інгібування, притаманного рекомбінантному поліпептиду, і активності готової форми, що буде застосовуватися в рослинних або харчових біоаналізах, композиція/готова форма може містити різні масові кількості рекомбінантного поліпептиду, наприклад, від 0,0001 % до 0,001 % від 0,01 % до 1 % до 99 % по масі рекомбінантного поліпептиду.

Розкриті в цьому документі білки можна комбінувати в готових формах для місцевого застосування до покривів рослин, до ґрунту, в готових формах для обробки насіння і в готових формах з іншими агентами, токсичними для вільових видів лускокрилих-шкідників. Такі агенти включають в себе, але не обмежуються лише цими: білки Cry1A, химери Cry1B, Cry1C, Cry1F, Cry1A/F і білок Cry2Ab.

#### ПРИКЛАДИ

Наступні розкриті варіанти здійснення просто представляють винахід, який може бути втілений в різних формах. Таким чином, конкретні структурні і функціональні деталі, розкриті в цьому документі, не повинні інтерпретуватися як обмежуючі. Слід розуміти, що повне розкриття кожного джерела, яке цитується в цьому документі, включене в цей документ з допомогою посилання.

#### Приклад 1

Цей приклад описує відкриття і аналіз токсичного білка BCW 001 і конструювання химерних

токсинів BCW 002 і BCW 003.

Було встановлено, що штам EG4384 Bt володіє активністю проти лускокрилих в харчових біоаналізах з використанням препаратів спор з кристалами. Була отримана послідовність генома цього штаму, були оброблені сирі зчитані послідовності, зібрані контіги з оброблених зчитаних послідовностей, ідентифіковані відкриті рамки зчитування, які показали гомологію з білками Cry1, і проаналізовані отримані амінокислотні послідовності. Була ідентифікована конкретна відкрита рамка зчитування, як вказано в SEQ ID NO: 1, яка кодує отриману амінокислотну послідовність білка (BCW 001, SEQ ID NO: 2), яка демонструє нову амінокислотну послідовність в порівнянні з більшістю білків Cry1, відомих в цій галузі техніки. Отриманий білок з відкритої рамки зчитування володіє всіма характеристиками нового білка типу Cry1, оскільки він має довжину 1180 амінокислот, а вирівнювання з відомими білками Cry1 вказує на те, що цей білок має характерну трьоходоменну структуру в межах близько 600-630 амінокислотних амінокислот і характерну структуру амінокислотної послідовності протоксину типу Cry1A. Полінуклеотидна послідовність, яка кодує цю передбачену амінокислотну послідовність, містить відкриту рамку зчитування, яка також є характерною для Cry1, а саме є сайт рестрикції NheI всередині сегмента ДНК, який кодує С-кінцеву область передбаченого домену I токсину, і сайт рестрикції KpnI в межах сегмента ДНК, який кодує N-кінцеву частину передбаченого прототоксिनного домену.

Порівняння амінокислотної послідовності токсину BCW 001 з Cry1Ac показує, що амінокислотний сегмент, який відповідає домену I (амінокислоти від близько позиції 1 до близько позиції 258), демонструє лише близько 67 % ідентичності з таким же сегментом в Cry1Ac, амінокислотний сегмент, який відповідає домену II (амінокислоти від близько позиції 58 до близько позиції 460), демонструє дуже низький відсоток ідентичності з таким же сегментом в Cry1Ac, а амінокислотний сегмент, який відповідає домену III (амінокислоти від близько позиції 460 до близько позиції 607) демонструє близько 63 % ідентичності з сегментом домену III з Cry1Ah2.

Сегмент ДНК, який кодує по суті передбачені домени II і III, був вирізаний по сайтам рестрикції від NheI до KpnI, і замінений на відповідний сегмент кодуєчої послідовності Cry1Ac в векторі експресії, який містить сегмент ДНК, який кодує Cry1Ac, що дало відкриту рамку зчитування, яка складається з, і з'єднання по рамці зчитування в послідовному порядку від 5'кінця до 3'кінця, першого сегмента, який кодує домен I Cry1Ac, другого сегмента, який кодує домен II і III BCW 001, і третього сегмента, який кодує прототоксिनний домен токсичного білка Cry1Ac. Ця химерна конструкція (SEQ ID NO: 3) кодує химерний токсичний білок, позначений в цьому документі як BCW 002 (SEQ ID NO: 4). Незначне зміщення точки розриву між доменом I і доменом II дає відкриту рамку зчитування (SEQ ID NO: 5), яка кодує химерний токсичний білок, який позначений в цьому документі як BCW 003 (SEQ ID NO: 6), який має амінокислотну послідовність, яка відрізняється від BCW002 лише по амінокислотній позиції 259. BCW 003, як і BCW 001, містить треонін (T) в позиції 259, а BCW 002 містить ізолейцин (I) в цій позиції. BCW 001 відрізняється від BCW 002 і BCW 003, головним чином, в домені I токсину, тобто по амінокислотам 1-202, а BCW 002 і BCW 002 практично ідентичні, за винятком різниці I/T в позиції 259.

#### Приклад 2

Цей приклад показує ефективність біологічної активності білків BCW 001, 002 і 003 для контролю лускокрилих шкідників.

Трансформуючі конструкції, які експресують токсичні білки BCW 001, 002 і 003 в *E. coli* або в застосовуваних *Bacillus thuringiensis* або інших *Bacilli*, дозволили провести тестування експресованих білків в біоаналізі і порівняти з білками, відомими в цій галузі техніки, які токсичні для совки-іпсилон, такими як Cry1Fa і Cry1Ac. Спостерігалось, що отримані рекомбінантні штами експресують рекомбінантний білок з активністю проти лускокрилих шкідників. Активність в біоаналізі була особливо сильною при тестуванні проти личинок совки-іпсилон (*Agrotis ipsilon*) і *Helicoverpa zea*. Як вказано вище в ДЕТАЛЬНОМУ ОПИСІ СУТІ ВІНАХОДУ, РІВНІ ТЕХНІКИ та КОРОТКОМУ ОПИСІ СУТІ ВІНАХОДУ, було виявлено дуже мало токсичних білків, які проявляють будь-який помітний рівень біологічної активності проти совки-іпсилон, і тому в цій галузі техніки існує потреба в ідентифікації таких білків для використання в рослинах, для захисту таких рослин від зараження совкою-іпсилон, і для того, щоб переконатись, що існує достатній запас додаткових активних проти совки-іпсилон білків, доступних для подолання будь-якого розвитку стійкості до будь-яких таких активних проти совки-іпсилон білків, які використовуються в даний час, таких як токсичні білки Cry1Fa.

#### Приклад 3

В цьому прикладі продемонстровано, що домени II і III BCW 001 достатньо для передачі

біологічної активності проти совки-іпсилон інших токсичних білків Cry1, коли такі домени замінюють відповідні домени в таких інших токсинах Cry1.

Було ідентифіковано багато химерних токсинів BCW, які мають активність проти лускокрилих, зокрема, дві химери проявляли сильну активність проти BCW, WBC і SCB.

5 Конструкції з нуклеотидними послідовностями, які кодують Cry1Ab, Cry1Ac і Cry1Ca, використовували для конструювання химер, які містять сегменти домену II і домену III BCW 001, які замінюють відповідні домени Cry1Ab, Cry1Ac і Cry1Ca, і отримані нативні і химерні білки тестували в біоаналізі з кристалами спор. Активності проти BCW, FAW і CEW в таких харчових біоаналізах були зведені в таблицю. В експериментальних умовах тестування, Cry1Ac демонстрував активність проти FAW, BCW і CEW, Cry1Ab демонстрував активність проти FAW і CEW, але не проти BCW, а Cry1Ca не демонстрував активності проти FAW, BCW, а також CEW. BCW 001 демонстрував активність проти BCW і CEW, але не FAW. У порівнянні з BCW 003, активність Cry1Ac проти BCW була близько в десять разів меншою ніж у Cry1Ac. Cry1Ab і Cry1Ac, химери, які містять домени II і III BCW 001, демонстрували токсичні властивості при тестуванні в біоаналізі проти FAW, BCW і CEW. Cry1Ac не був токсичним для CEW, а Cry1Ab не був токсичним для BCW. Були сконструйовані химери Cry1Ca/BCW 001, в яких домен III Cry1Ca був замінений відповідним доменом BCW 001, і отриманий химерний токсин демонстрував токсичні властивості проти FAW, BCW і CEW, тоді як токсин Cry1Ca був неефективний при тестуванні проти будь-якого з цих шкідників.

#### 20 Приклад 4

Цей приклад ілюструє токсичні властивості BCW 001, 002 і 003 при тестуванні в біоаналізі проти різноманітних лускокрилих шкідників.

25 Протоколи для біоаналізу і підрахунку смертності і затримки росту комах відомі в цій галузі техніки, приклади яких описані в публікації заявки на патент PCT № WO 2012/139004 і в патенті США № 7927598.

На Фіг. 1 співвідносяться різні токсичні білки BCW 001, 002 і 003 по пестицидній активності до видів комах в харчових біоаналізах. Кожний з цих токсичних білків продемонстрував активність проти лускокрилих комах.

#### 30 Приклад 5

У цьому прикладі описується конструювання штучних послідовностей, які кодують білки згідно цього винаходу для використання в рослинах, приготування рослинних векторів і конструкцій для використання в рослинах, і отримання рослин, які експресують білки згідно цього винаходу.

35 Нуклеотидні послідовності, які кодують білок BCW 001 (SEQ ID NO: 1), білок BCW 002 (SEQ ID NO: 3) і білок BCW 003 (SEQ ID NO: 5), були сконструйовані і синтезовані згідно способів, описаних в патенті США 5500365. Ці сконструйовані кодуючі області, призначені для експресії в рослинах, представлені в цьому документі як SEQ ID NO: 7, яка кодує BCW 001, SEQ ID NO: 9, яка кодує BCW 002, і SEQ ID NO: 11, яка кодує BCW 003.

40 Різноманітні касети для експресії в рослинах конструювали з послідовностями, вказаними в SEQ ID NO: 7, 9 і 11. Такі касети експресії корисні для тимчасової експресії в протопластах рослин або для трансформації рослинних клітин. Типові експресійні касети були сконструйовані з урахуванням можливого розміщення білка в клітині. Один набір касет експресії був сконструйований таким чином, щоб дозволити білку транслюватися і залишатися в цитоплазмі. Інший набір касет експресії був сконструйований так, щоб мати транзитний пептид, суміжний з токсичним білком, щоб забезпечити націлювання на органелу клітини, таку як хлоропласт або 45 пластид. Всі касети експресії були сконструйовані так, щоб починатися з 5'-кінця промотором, який може складатися з декількох промоторних елементів і енансерних елементів, суміжно зв'язаних для посилення експресії трансгена. За промоторною послідовністю зазвичай слідували суміжно одна або декілька лідерних послідовностей в 3' відносно промотору. 50 Послідовність інтрона була розміщена в 3' відносно лідерної послідовності для поліпшення експресії трансгена. Кодуюча послідовність токсину або транзитного пептиду, і кодуюча послідовність токсину були розташовані в 3' по відношенню до промотору, лідерної послідовності та інтронів. Послідовність була розміщена в 3' по відношенню до кодуючої послідовності для полегшення термінації транскрипції і забезпечує послідовності, важливі для поліаденілювання отриманого транскрипта. Всі елементи, описані вище, були розташовані суміжно, часто з додатковою послідовністю, передбаченою для конструювання експресійної касети, такою як сайти рестрикції ендонуклеаз або сайти безлігазного клонування.

55 Для рослин кукурудзи був розроблений набір касет експресії для цитозольної експресії BCW 001, яка містить промотор убіквітину 1 Mexicana, BCW 002, який містить промотор актину 15 Oryza sativa або промотор 35S, і BCW 003, який містить промотор 35S.

Інший набір касет експресії був розроблений для цільової експресії в рослинах кукурудзи токсичних для комах білків BCW 002 і BCW 003, в яких послідовність, яка кодує пептид хлоропласта (наприклад, СТР2), була злита по рамці зчитування на 5'-кінці ДНК сегмента, який кодує токсичні білки BCW, що містять промотор актину 15 *Oryza sativa* або промотор 35S, і послідовність, яка містить промотор 35S.

Касети для експресії в рослинах цукрової тростини, що містять промотор CaMV 35S або промотор PCISV.FLt, функціонально зв'язаний з промотором 35S, конструювали у вигляді векторів трансформації рослин. У деяких випадках, була включена касета, яка експресує націлений на хлоропласти Cry2Ab.

Рослини, які експресують білки згідно цього винаходу, тестували проти личинок BCW, WBC, CEW, SWC і SCB третьої лінки. Касета експресії в цитозолі для BCW 001, і націлена на цитозоль і пластиди касета експресії для BCW 002 і BCW 003 були клоновані і використовувалися для отримання трансгенних рослин кукурудзи, які експресують ці білки. Трансформовані клітини індукували для формування рослин способами, відомими в цій галузі техніки. Біоаналіз з використанням дисків листя рослин проводили аналогічно тому, як описано в патенті США № 8344207. Ступеню пошкодження листа (LDR) присвоювалась оцінка, заснована на відсотку диска листа, з'їденого комахою, за шкалою від 0 (з'їдено 0 %) до 11 (з'їдено понад 50 %). Крок оцінки ступеня збільшується на 5 %. Ізогенну лінію кукурудзи використовували для отримання тканини в якості негативного контролю, і результати оцінювали. Як експресія в пластидах-мішенях, так і цитозольна експресія токсичних для комах білків BCW 002 і BCW 003 знижували пошкодження при харчуванні в порівнянні з нетрансформованим контролем. Результати аналізів дисків з листя по відношенню до цих комах узгоджуються з даними біоаналізу із прикладів, наведених вище. Одна конструкція, яка містить касету BCW 003 для цитозольної експресії, дала 34 трансформанта, 25 з них продемонстрували повний контроль над личинками BCW. Рослини кукурудзи, які експресують BCW 001 і BCW 003 в цитоплазмі, також тестували на CEW, SWC і FAW. Рослини, які експресували BCW 003, демонстрували 100 % контроль CEW і SWC, і значення LDR в діапазоні від 1 до 2. Три трансформанта, які експресували BCW 001, дали рослини, що проявляють ефективність по відношенню до CEW і SWC, і значення LDR в діапазоні від 1 до 3. Це узгоджується з даними харчового біоаналізу, представленими в попередніх прикладах.

Трансгенні рослини цукрової тростини, які експресували BCW 003, були отримані і випробувані проти SCB в біоаналізі. Кожний біоаналіз включав в себе диски листя з цукрової тростини дикого типу в якості негативного контролю, і з позитивного контролю, який експресував високі рівні Cry2Ab. Смертність комах і пошкодження листя вимірювали через чотири (4) дні після зараження. Було виявлено, що диски листя декількох трансформантів цукрової тростини, які експресували BCW 003, демонструють контроль *Diatraea saccharalis* in planta і мають ступінь пошкодження нижче 2, аналогічно позитивному контролю і середній смертності комах 90-100 %.

Трансформанти, які експресували BCW 003 і Cry2Ab, демонстрували кращий контроль SCB в порівнянні з трансформантами, які експресували лише BCW 003.

#### ФОРМУЛА ВИНАХОДУ

1. Полінуклеотидна конструкція, яка містить нуклеотидну послідовність, яка кодує:

(a) інсектицидний білок, який має амінокислотну послідовність, що містить SEQ ID NO: 6 або SEQ ID NO: 12, або;

(b) інсектицидний білок, який має щонайменше 95 % ідентичності з амінокислотною послідовністю (a),

при цьому вказана нуклеотидна послідовність функціонально зв'язана з гетерологічною послідовністю промотору.

2. Білок, токсичний для виду лускокрилих совки-іпсилон (black cutworm), який містить:

(a) амінокислотну послідовність, що містить SEQ ID NO: 6 або SEQ ID NO: 12; або

(b) амінокислотну послідовність, яка має щонайменше 95 % ідентичності з амінокислотною послідовністю (a).

3. Полінуклеотидна конструкція за п. 1, яка **відрізняється** тим, що вказаний інсектицидний білок проявляє активність проти видів лускокрилих, вибраних з групи, що складається з: *Agrotis ipsilon*, *Striacosta albicosta* і *Helicoverpa zea*, *Ostrinia nubilalis*, *Diatraea saccharalis*, *Diatraea grandiosella*, *Trichoplusia ni* та *Pseudoplusia includens*.

4. Білок за п. 2, який **відрізняється** тим, що вказаний інсектицидний білок проявляє біоактивність проти видів лускокрилих, вибраних з групи, що складається з: *Agrotis ipsilon*,

*Striacosta albicosta* і *Helicoverpa zea*, *Ostrinia nubilalis*, *Diatraea saccharalis*, *Diatraea grandiosella*, *Trichoplusia ni* та *Pseudoplusia includens*.

5. Вектор, який містить полінуклеотидну конструкцію за п. 1.

6. Клітина-хазяїн, яка містить полінуклеотидну конструкцію за п. 1, яка **відрізняється** тим, що вказана клітина-хазяїн вибрана з групи, що складається з бактеріальної клітини, дріжджової клітини і рослинної клітини.

7. Клітина-хазяїн за п. 6, яка **відрізняється** тим, що вказану рослинну клітину вибирають з групи, яка складається з рослинної клітини: люцерни, банана, ячменю, квасолі, броколі, капусти, капусти декоративної, моркви, маніоки, рицини, цвітної капусти, селери, нуту, китайської капусти, цитрусових, кокосової пальми, кави, кукурудзи, конюшини, бавовника, гарбузових, огірка, псевдотсуги Мензиса, баклажана, евкаліпта, льону, часнику, винограду, хмелю, цибулі-порей, салату-латуку, сосни ладанної, проса, дині, горіха, вівса, оливкового дерева, ріпчастої цибулі, декоративних рослин, пальмових, пасовищних трав, гороху, арахісу, перцю, голубиног гороху, соснових, картоплі, тополі, гарбуза, сосни променистої, редьки, ріпаку, рису, кореневищ, жита, дикою шафрану, чагарникових, сорго, сосни південної, сої, шпинату, гарбузових, полуниці, цукровою буряка, цукрової тростини, соняшнику, кукурудзи цукрової, амбрового дерева, батату, проса прутувидного, чаю, тютюну, помідора, тритикале, дернової трави, кавуна і пшениці.

8. Рослина, яка містить полінуклеотидну конструкцію за п. 1.

9. Насіння, отримане з рослини за п. 8, яке **відрізняється** тим, що вказане насіння містить визначувану кількість вказаної полінуклеотидної конструкції.

10. Рослина за п. 8, яка **відрізняється** тим, що насіння, пилок, потомство, рослинні клітини, рослинна тканина і товарні продукти, отримані з вказаної рослини, містять визначувану кількість вказаної полінуклеотидної конструкції.

11. Біологічний зразок, який містить придатну для визначення кількість полінуклеотидної конструкції за п. 1.

12. Композиція, яка забезпечує інсектицидно ефективну кількість білка за п. 2 для боротьби з видами лускокрилих шкідників, і:

(а) агент, який відрізняється від вказаного білка і також токсичний для тих же видів лускокрилих, причому вказаний агент вибирають з групи, яка складається з поліпептиду, який має амінокислотну послідовність, яка відрізняється від вказаного білка, молекули РНК і хімічної сполуки; або

(b) агент вибраний з групи, що складається з: Cry1A, Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae, Cry1B, Cry1C, Cry1D, Cry1E, Cry1F, Cry1G, Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab, Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, ET35, ET66, TIC400, TIC800, TIC807, TIC834, TIC853, TIC1415, VIP3A, VIP3Ab, інсектицидних білків AXMI, інсектицидних білків DIG, eNIPs і білків VIP.

13. Композиція за п. 12, яка додатково містить додатковий пестицидний агент, причому вказаний додатковий агент вибирають з групи, що складається з: Cry1C, Cry3A, Cry3B, Cry34, Cry35, Cry51Aa1, ET29, ET33, ET34, ET70, TIC407, TIC417, TIC431, TIC901, TIC1201, TIC3131, 5307, DIG-10, Axmi184, Axmi205 і AxmiR1.

14. Рослина за п. 8, яка додатково містить:

(а) трансгенну подію рослини кукурудзи, вибрану з групи, що складається з: DKB89614-9, MON801, MON802, MON809, MON810, MON863, MON88017, MON89034, трансформанта 4114-3, трансформанта 5307, DAS59122-7, Bt10, Bt11, Bt176, СВН-351, DKB-83614-9, MIR162, MIR604, TC1507, TC6275, трансформанта 676, трансформанта 678, трансформанта 680, трансформанта 98140, DAS40278-9, DKB89790-5, MON21-9, HCEM485, MON832, MON87427, NK603, T14, T25 і VCO01981-5;

(b) трансгенну подію рослини сої, вибрану з групи, що складається з: MON87751, DAS81419-2, MON87701, A2704-12, A2704-21, A5547-127, A5547-35, CV127, DAS44406-6, DAS68416-4, DP356043, FG72, MON4032, ACS-GM003-1, MON87705, MON87708, MON89788, W62, W98 і GFM Cry1A;

(c) трансгенну подію рослини бавовни, вибрану з групи, що складається з: DAS24236-5, DAS21023-5, трансформанта 31707, трансформанта 31803, трансформанта 31807, трансформанта 31808, трансформанта 42317, BNLA-601, COT102, COT67B, трансформанта 1, GHB119, GK12, MON15985, MLS9124, MON1076, MON531, MON757, T303-3, T304-40, SGK321, трансформанта 19-51a, GHB614, LLCotton25, MON88701, MON88702, MON1445, MON1698 і MON88913;

(d) трансгенну подію рослини цукрової тростини трансформанта NXI-1T; і

(e) трансгенну подію рослини рису, вибрану з групи, що складається з: LLRICE06, LLRICE601, LLRIC2E62, GM-A17054 і GM-A17054.

Види комар	BCW 001	BCW 002	BCW 003
<i>A. ipsilon</i>	+	+	+
<i>S. albicosta</i>	+	+	+
<i>H. zea</i>	+	+	+
<i>O. nubilalis</i>	+	+	+
<i>D. saccharalis</i>	HA	HA	+
<i>D. grandiosella</i>	+	HA	+
<i>T. ni</i>	+	HA	+
<i>P. includens</i>	+	HA	+
<i>S. frugiperda</i>	-	-	-

Фір. 1

SEQ ID NO: 2	MEER-NQDQCVFYNCLNHPAYEILEGURIISVGRYFIDI SL SLVELL ISEPVPKXNIIITGL	59
SEQ ID NO: 4	MKNNPH IRECPYFNCLSNPEVENLGGRIETGYTFIDI GLSLTQFLLSEFVPCGAFVI GL	60
SEQ ID NO: 6	MDNRPXNECIEFYNCLSHPEVYVLAGGRIBTGYTFIDISL SLTQFLLSEFVPCGAFVI GL	60
	* * * * *	
SEQ ID NO: 2	LEIYVGFVGFQNDAPLAQVEQLINQRISEAVNFAIQHLEQMARVYHTYATAPAEWERD	119
SEQ ID NO: 4	VDIYWGIFGDSQWDAPLVDIQLINQRISEAVNFAIQHLEQMARVYHTYATAPAEWERD	120
SEQ ID NO: 6	VDIYWGIFGDSQWDAPLVDIQLINQRISEAVNFAIQHLEQMARVYHTYATAPAEWERD	120
	* * * * *	
SEQ ID NO: 2	FENYHLEKAVHTQPTATEYIISGRISVYKIQNPEVQILLVYFAQANLHLSLLKDVVFFGQ	179
SEQ ID NO: 4	FENPALREENAIQFUMHREALTYSIFLLAVQRIQVPLLAVTVQSEALHLSVLDVSVVFGQ	180
SEQ ID NO: 6	FENPALREENAIQFUMHREALTYSIFLLAVQRIQVPLLAVTVQSEALHLSVLDVSVVFGQ	180
	* * * * *	
SEQ ID NO: 2	RWQFSTTVVNHFFNDLREIISTYTGAVRWYNTGLEPVGQDSDROWVRYNQFAPRETLTV	239
SEQ ID NO: 4	RWQFDAMTINSEYMLTFLIGNYDYSVRYVNTGLEPVGQDSDPDWRYNQFAPRETLTV	240
SEQ ID NO: 6	RWQFBAATINRYNDLREIISTYTGAVRWYNTGLEPVGQDSDROWVRYNQFAPRETLTV	240
	* * * * *	
SEQ ID NO: 2	LDIVALFPNVDSEHYPIKIVGQLTTEYTRPVLENFDGSPFGMAQRISQNIQQLADYL	299
SEQ ID NO: 4	LDIVALFPNVDSEHYPIKIVGQLTTEYTRPVLENFDGSPFGMAQRISQNIQQLADYL	300
SEQ ID NO: 6	LDIVALFPNVDSEHYPIKIVGQLTTEYTRPVLENFDGSPFGMAQRISQNIQQLADYL	300
	*	
SEQ ID NO: 2	ESITITDVRAGPNEWEGSQITASPVGFSQPEFAPPLFENAGNAAPVYVSLTGLGTFK	359
SEQ ID NO: 4	NGITLYTQVREDFNYRGGHQIPASVVGFSQPEFAPPLFENAGNAAPVYVSLTGLGTFK	360
SEQ ID NO: 6	ESITITDVRAGPNEWEGSQITASPVGFSQPEFAPPLFENAGNAAPVYVSLTGLGTFK	360
SEQ ID NO: 2	LSEFLYPAIILGSGPNNQELFVLDGTEFSPASLTMLPSTIYKQACTVGLDVIIPQONS	419
SEQ ID NO: 4	LSEFLYPAIILGSGPNNQELFVLDGTEFSPASLTMLPSTIYKQACTVGLDVIIPQONS	420
SEQ ID NO: 6	LSEFLYPAIILGSGPNNQELFVLDGTEFSPASLTMLPSTIYKQACTVGLDVIIPQONS	420
SEQ ID NO: 2	VPPRAGFSHRLSHVTHLSQAAGAVYTLRAPTFYSHQRRSATTNTIADSEITQIPAVKGRS	479
SEQ ID NO: 4	VPPRAGFSHRLSHVTHLSQAAGAVYTLRAPTFYSHQRRSATTNTIADSEITQIPAVKGRS	480
SEQ ID NO: 6	VPPRAGFSHRLSHVTHLSQAAGAVYTLRAPTFYSHQRRSATTNTIADSEITQIPAVKGRS	480
SEQ ID NO: 2	IINNGTVISGPFETCCDLVRLYNAQFNINBAYLAVPIFPQGPSINYYVRYVYASTSSLP	539
SEQ ID NO: 4	IINNGTVISGPFETCCDLVRLYNAQFNINBAYLAVPIFPQGPSINYYVRYVYASTSSLP	540
SEQ ID NO: 6	IINNGTVISGPFETCCDLVRLYNAQFNINBAYLAVPIFPQGPSINYYVRYVYASTSSLP	540
SEQ ID NO: 2	VGVVFNILSHFTTTPADARSLQNLQSDPCEIIGTAGETLPSLSESGIGRPELSTINLIVD	599
SEQ ID NO: 4	VGVVFGDI SHFTTTPADARSLQNLQSDPCEIIGTAGETLPSLSESGIGRPELSTINLIVD	600
SEQ ID NO: 6	VGVVFNILSHFTTTPADARSLQNLQSDPCEIIGTAGETLPSLSESGIGRPELSTINLIVD	600
SEQ ID NO: 2	SEFTFVQATFRAESDLBAAQKAVNALFTSTRQLGKRTDVIDYHIDQVSNLVECLSDSEFY	659
SEQ ID NO: 4	KP6FLPVTATFRAESDLBAAQKAVNALFTSTRQLGKRTDVIDYHIDQVSNLVECLSDSEFY	660
SEQ ID NO: 6	SEFTFVQATFRAESDLBAAQKAVNALFTSTRQLGKRTDVIDYHIDQVSNLVECLSDSEFY	660

Фір. 2

SEQ ID NO: 2	LDEKPELSEKVKHAKPLSDERNLLQDPNFRGINSRQFDKQWRGSDIITIQSGDDVFKENYV	719
SEQ ID NO: 4	LDEKPELSEKVKHAKPLSDERNLLQDPNFRGINSRQFDKQWRGSDIITIQSGDDVFKENYV	720
SEQ ID NO: 6	LDEKPELSEKVKHAKPLSDERNLLQDPNFRGINSRQFDKQWRGSDIITIQSGDDVFKENYV	720
SEQ ID NO: 2	TLPQTFDECYPTTYLYQKIDSEKLEKAYTRYQLRCYIEDSQQLEIYLIIRYNAKHETVWVPGT	779
SEQ ID NO: 4	TLPQTFDECYPTTYLYQKIDSEKLEKAYTRYQLRCYIEDSQQLEIYLIIRYNAKHETVWVPGT	780
SEQ ID NO: 6	TLPQTFDECYPTTYLYQKIDSEKLEKAYTRYQLRCYIEDSQQLEIYLIIRYNAKHETVWVPGT	780
*		
SEQ ID NO: 2	GSLWPLSAQSPFKKCGEPNRCAPHLEWNPDLDCSCSDCEKCAHSHHFFSLDIDVGCETDLN	839
SEQ ID NO: 4	GSLWPLSAQSPFKKCGEPNRCAPHLEWNPDLDCSCSDCEKCAHSHHFFSLDIDVGCETDLN	840
SEQ ID NO: 6	GSLWPLSAQSPFKKCGEPNRCAPHLEWNPDLDCSCSDCEKCAHSHHFFSLDIDVGCETDLN	840
*		
SEQ ID NO: 2	EDLCVWVIFKIKTQDCHARLCHLEFLLEKPLVGCALARVYKRAEKKWRDREKLEWETNIV	899
SEQ ID NO: 4	EDLCVWVIFKIKTQDCHARLCHLEFLLEKPLVGCALARVYKRAEKKWRDREKLEWETNIV	900
SEQ ID NO: 6	EDLCVWVIFKIKTQDCHARLCHLEFLLEKPLVGCALARVYKRAEKKWRDREKLEWETNIV	900
**		
SEQ ID NO: 2	YKEAKESVDALFVNSQYDQLQADTNIAMIIHAADKRVKRIPEAYLDELVIKGVNADISEE	959
SEQ ID NO: 4	YKEAKESVDALFVNSQYDQLQADTNIAMIIHAADKRVKRIPEAYLDELVIKGVNADISEE	960
SEQ ID NO: 6	YKEAKESVDALFVNSQYDQLQADTNIAMIIHAADKRVKRIPEAYLDELVIKGVNADISEE	960
* * *		
SEQ ID NO: 2	LECRIFTAFSLYDARNVIKNGDFNNGLLCWNVYKGVVDVVEEQNHRSVLYVPEWEAEVSGE	1019
SEQ ID NO: 4	LECRIFTAFSLYDARNVIKNGDFNNGLLCWNVYKGVVDVVEEQNHRSVLYVPEWEAEVSGE	1020
SEQ ID NO: 6	LECRIFTAFSLYDARNVIKNGDFNNGLLCWNVYKGVVDVVEEQNHRSVLYVPEWEAEVSGE	1020
* *		
SEQ ID NO: 2	VKVCFGRCYIIRVTAYKEGCGECQVTIHEIENNTDELKFPKNCVREESIYQNTVTQNDYTV	1079
SEQ ID NO: 4	VKVCFGRCYIIRVTAYKEGCGECQVTIHEIENNTDELKFPKNCVREESIYQNTVTQNDYTV	1080
SEQ ID NO: 6	VKVCFGRCYIIRVTAYKEGCGECQVTIHEIENNTDELKFPKNCVREESIYQNTVTQNDYTV	1080
* *		
SEQ ID NO: 2	NQERVGGAYTFRNRCYDEAYEENRQSVFADYASVYEEKVVYTDGGRGDPCEENRCYQDYTFL	1139
SEQ ID NO: 4	NQERVGGAYTFRNRCYDEAYEENRQSVFADYASVYEEKVVYTDGGRGDPCEENRCYQDYTFL	1136
SEQ ID NO: 6	NQERVGGAYTFRNRCYDEAYEENRQSVFADYASVYEEKVVYTDGGRGDPCEENRCYQDYTFL	1136
* * * * * * * * *		
SEQ ID NO: 2	PCYVTKLELYFPSTKQVWIEIGSTECTYIVDSVELLMEE	1180
SEQ ID NO: 4	PCYVTKLELYFPSTKQVWIEIGSTECTYIVDSVELLMEE	1177
SEQ ID NO: 6	PCYVTKLELYFPSTKQVWIEIGSTECTYIVDSVELLMEE	1177
*		

Фіг. 2 (продовження)

ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

<110> Monsanto Technology LLC

<120> ПЕСТИЦИДНИ ВІЛКОВІ ТОКСИНИ, АКТИВНІ ПО ВІДНОШЕННЮ ДО ЛУСКОКРИЛИХ КОМАХ

<130> MONS:434WO

<150> US 62/445,313

<151> 2017-01-12

<160> 12

<170> Bepcia PatentIn 3.5

<210> 1

<211> 3543

<212> ДНК

<213> Bacillus thuringiensis

<220>

<221> CDS

<222> (1)..(3540)

<400> 1

atg gag gaa aat aat cag aat caa tgc gtc cct tat aat tgt ttg aat	48
Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn	
1 5 10 15	
aat cct gca atc gaa ata tta gaa gga gac aga ata tca gtt ggt aac	96
Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn	
20 25 30	
act cca atc gat att tct cta tca ctt gtg gaa ctt ctt att agt gaa	144
Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu	
35 40 45	
ttt gtc cca ggc ggt gga ata ata aca gga ttg ttg aac ata gta tgg	192
Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp	
50 55 60	
gga ttt gta ggg cct tcc caa tgg gac gca ttt ctt got caa gtg gaa	240
Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu	
65 70 75 80	
cag tta att aac caa agg ata tca gaa gct gta aga aat aca gca att	288
Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile	
85 90 95	
cag gaa tta gag gga atg gcg cgg gtt tat aga acc tat got act got	336
Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala	
100 105 110	
ttt gct gag tgg gaa aga gat cct aat aac aca gat cta aga gaa gca	384
Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala	

UA 128156 C2

115					120					125						
gta	cgg	aca	cag	ttt	aca	gca	act	gag	act	tat	atc	agt	gga	aga	ata	432
Val	Arg	Thr	Gln	Phe	Thr	Ala	Thr	Glu	Thr	Tyr	Ile	Ser	Gly	Arg	Ile	
	130					135					140					
tct	gtt	tta	aaa	att	caa	aat	ttt	gaa	gtg	cag	ctg	tta	tcg	gtg	ttt	480
Ser	Val	Leu	Lys	Ile	Gln	Asn	Phe	Glu	Val	Gln	Leu	Leu	Ser	Val	Phe	
	145					150					155				160	
gcc	caa	gct	gcc	aat	tta	cat	tta	tct	tta	tta	aga	gac	gtt	gtg	ttt	528
Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Leu	Leu	Arg	Asp	Val	Val	Phe	
				165					170					175		
ttt	ggg	caa	aga	tgg	ggg	ttt	tca	acg	aca	acc	gta	aat	aat	tac	tac	576
Phe	Gly	Gln	Arg	Trp	Gly	Phe	Ser	Thr	Thr	Thr	Val	Asn	Asn	Tyr	Tyr	
			180					185						190		
aat	gat	tta	aca	gaa	gag	att	agt	acc	tat	aca	gat	tat	gca	gta	cgc	624
Asn	Asp	Leu	Thr	Glu	Glu	Ile	Ser	Thr	Tyr	Thr	Asp	Tyr	Ala	Val	Arg	
		195					200					205				
tgg	tac	aat	acg	gga	tta	gag	cgt	gta	tgg	gga	ccg	gat	tct	aga	gat	672
Trp	Tyr	Asn	Thr	Gly	Leu	Glu	Arg	Val	Trp	Gly	Pro	Asp	Ser	Arg	Asp	
	210					215					220					
tgg	gta	agg	tat	aat	caa	ttt	aga	aga	gag	cta	aca	ctt	act	gta	tta	720
Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Glu	Leu	Thr	Leu	Thr	Val	Leu	
	225					230					235				240	
gat	atc	gtt	gct	cta	ttc	cca	aat	tat	gat	agt	cga	agg	tat	cca	att	768
Asp	Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Ser	Arg	Arg	Tyr	Pro	Ile	
				245					250					255		
cga	aca	gtt	tcc	caa	tta	aca	aga	gaa	att	tat	acg	aac	cca	gta	tta	816
Arg	Thr	Val	Ser	Gln	Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Asn	Pro	Val	Leu	
			260					265					270			
gaa	aat	ttt	gat	ggt	agt	ttt	cgt	gga	atg	gct	cag	aga	ata	gaa	cag	864
Glu	Asn	Phe	Asp	Gly	Ser	Phe	Arg	Gly	Met	Ala	Gln	Arg	Ile	Glu	Gln	
		275					280					285				
aat	att	agg	caa	cca	cat	ctt	atg	gat	atc	ctt	aat	agt	ata	acc	att	912
Asn	Ile	Arg	Gln	Pro	His	Leu	Met	Asp	Ile	Leu	Asn	Ser	Ile	Thr	Ile	
	290					295					300					
tat	act	gat	gtg	cat	aga	ggc	ttt	aat	tat	tgg	tca	ggg	cat	caa	ata	960
Tyr	Thr	Asp	Val	His	Arg	Gly	Phe	Asn	Tyr	Trp	Ser	Gly	His	Gln	Ile	
	305					310					315				320	
aca	gct	tct	cct	gta	ggg	ttt	tca	gga	cca	gaa	ttc	gca	ttc	cct	tta	1008
Thr	Ala	Ser	Pro	Val	Gly	Phe	Ser	Gly	Pro	Glu	Phe	Ala	Phe	Pro	Leu	
				325					330					335		
ttt	ggg	aat	gcg	gga	aat	gca	gct	cca	ccc	gta	ctt	gtc	tca	tta	act	1056
Phe	Gly	Asn	Ala	Gly	Asn	Ala	Ala	Pro	Pro	Val	Leu	Val	Ser	Leu	Thr	
			340					345					350			

ggt ttg ggg att ttt aga aca tta tct tca cct tta tat aga aga att Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile 355 360 365	1104
ata ctt ggt tca ggc cca aat aat cag gaa ctg ttt gtc ctt gat gga Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly 370 375 380	1152
acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act ata Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile 385 390 395 400	1200
tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccg cca cag Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln 405 410 415	1248
gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt cat Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His 420 425 430	1296
gtt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga gct Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala 435 440 445	1344
cca acg ttt tct tgg cag cat cgc agt got acg aca act aat ata att Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile 450 455 460	1392
gca gcg gat agt att act caa att cct got gtt aaa gga cgt tct att Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile 465 470 475 480	1440
att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc gat Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp 485 490 495	1488
ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca tac Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr 500 505 510	1536
ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt gtt Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg Val 515 520 525	1584
cgt gtt cgt tat gct tct aca tct tca etc cct gta gat gta gtt ttc Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val Phe 530 535 540	1632
gga aat att agt cat cct act aca ttg cca gcc act gcc aga tca tta Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser Leu 545 550 555 560	1680
gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga act Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly Thr 565 570 575	1728

UA 128156 C2

ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta tct	1776
Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu Ser	
580 585 590	
act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act gca	1824
Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr Ala	
595 600 605	
acc ttt gaa gca gaa tcg gat tta gaa aga gca caa aag gcg gtg aat	1872
Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn	
610 615 620	
gcg ctg ttt act tct aca aac caa cta ggg ata aaa aca gat gtg acg	1920
Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val Thr	
625 630 635 640	
gat tat cat att gat caa gtg tcc aat tta gtg gag tgt tta tcg gat	1968
Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp	
645 650 655	
gaa ttt tat ctg gat gaa aag cga gaa ttg tcc gag aaa gtc aaa cat	2016
Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His	
660 665 670	
gcg aag cga ctc agt gat gag cga aat tta ctt caa gat cca aac ttc	2064
Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe	
675 680 685	
agg ggc atc aat aga caa cca gat cgt ggc tgg aga gga agt acg gat	2112
Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp	
690 695 700	
att acc atc caa gga gga gat gac gta ttc aaa gag aat tac gtc aca	2160
Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr	
705 710 715 720	
cta cca ggt acc ttt gat gag tgc tat cca acg tat tta tat caa aaa	2208
Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys	
725 730 735	
ata gat gag tcg aaa tta aaa gcc tat acc cgt tat caa tta aga ggg	2256
Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly	
740 745 750	
tat atc gag gat agt caa gac tta gaa atc tat tta att cgc tac aat	2304
Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn	
755 760 765	
gca aaa cat gaa aca gta aat gtg cca ggt acg ggt tcc tta tgg ccg	2352
Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro	
770 775 780	
ctt tca gcc caa agt cca atc gga aag tgt gga gag ccg aat cga tgc	2400
Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys	
785 790 795 800	
gcg cca cac ctt gaa tgg aat cct gat tta cac tgt tcc tgc aga gac	2448

UA 128156 C2

Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	His	Cys	Ser	Cys	Arg	Asp		
				805					810					815			
ggg	gaa	aaa	tgt	gct	cat	cat	tct	cat	cat	ttc	tcc	ttg	gac	att	gat		2496
Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	Asp		
			820					825					830				
gtt	gga	tgt	aca	gac	tta	aat	gag	gat	tta	ggt	gta	tgg	gtg	ata	ttc		2544
Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly	Val	Trp	Val	Ile	Phe		
		835					840					845					
aag	att	sag	acg	caa	gat	ggc	cat	gca	aga	cta	gga	aat	cta	gag	ttt		2592
Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu	Gly	Asn	Leu	Glu	Phe		
	850					855					860						
ctc	gaa	gag	aaa	cca	tta	gta	ggg	gaa	gca	cta	gct	ogt	gtg	aaa	aga		2640
Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	Arg		
865					870					875					880		
gcg	gag	aaa	aaa	tgg	aga	gac	aaa	cgc	gaa	aaa	tta	caa	ttg	gaa	aca		2688
Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Gln	Leu	Glu	Thr		
			885					890					895				
aat	atc	gtt	tat	aaa	gag	gca	aaa	gaa	tct	gta	gat	gct	tta	ttt	gta		2736
Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe	Val		
			900					905					910				
aac	tct	caa	tat	gat	caa	tta	caa	gcg	gat	acg	aac	atc	gcg	atg	att		2784
Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr	Asn	Ile	Ala	Met	Ile		
		915				920						925					
cat	gcg	gca	gat	aaa	cgt	gtt	cat	aga	atc	cga	gaa	gcg	tac	ctt	cca		2832
His	Ala	Ala	Asp	Lys	Arg	Val	His	Arg	Ile	Arg	Glu	Ala	Tyr	Leu	Pro		
	930					935					940						
gag	tta	tct	gtg	att	ccg	ggt	gta	aat	gca	gac	att	tcc	gaa	gaa	tta		2880
Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Asp	Ile	Ser	Glu	Glu	Leu		
945					950					955				960			
gaa	ggg	ogt	att	ttc	act	gca	ttc	tct	cta	tat	gat	gcg	aga	aat	gtc		2928
Glu	Gly	Arg	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr	Asp	Ala	Arg	Asn	Val		
			965						970				975				
att	aaa	aat	ggc	gat	ttc	aat	aat	ggc	tta	tta	tgc	tgg	aac	gtg	aaa		2976
Ile	Lys	Asn	Gly	Asp	Phe	Asn	Asn	Gly	Leu	Leu	Cys	Trp	Asn	Val	Lys		
			980					985					990				
ggg	cat	gta	gat	gta	gaa	gaa	caa	aat	aac	cac	cgt	tcc	gtc	ctt	gtt		3024
Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	His	Arg	Ser	Val	Leu	Val		
		995					1000					1005					
gtt	ccg	gaa	tgg	gaa	gca	gaa	gtg	tca	caa	gag	gtt	cgt	gtc	tgt			3069
Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val	Cys			
	1010					1015					1020						
ccg	ggg	cgt	ggc	tat	atc	ctt	cgt	gtc	aca	gcg	tac	aag	gag	gga			3114
Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu	Gly			

UA 128156 C2

1025	1030	1035	
tat gga gaa ggt tgc gta acc	att cat gag atc gag	aac aat aca	3159
Tyr Gly Glu Gly Cys Val Thr	Ile His Glu Ile Glu	Asn Asn Thr	
1040	1045	1050	
gac gaa ctg aag ttt agc aac	tgt gta gaa gag gaa	gtc tat cca	3204
Asp Glu Leu Lys Phe Ser Asn	Cys Val Glu Glu Glu	Val Tyr Pro	
1055	1060	1065	
aac aac acg gta acg tgt aat	gat tat act gca aat	caa gaa gaa	3249
Asn Asn Thr Val Thr Cys Asn	Asp Tyr Thr Ala Asn	Gln Glu Glu	
1070	1075	1080	
tat gag ggt acg tac act tct	cgt aat cga gga tat	gac gaa gcc	3294
Tyr Glu Gly Thr Tyr Thr Ser	Arg Asn Arg Gly Tyr	Asp Glu Ala	
1085	1090	1095	
tat gaa agc aat tct tct gta	cca gct gag tat gcg	tca gtc tat	3339
Tyr Glu Ser Asn Ser Ser Val	Pro Ala Glu Tyr Ala	Ser Val Tyr	
1100	1105	1110	
gaa gaa aaa gtg tat aca gat	gga cga aga ggg aat	cct tgt gaa	3384
Glu Glu Lys Val Tyr Thr Asp	Gly Arg Arg Gly Asn	Pro Cys Glu	
1115	1120	1125	
tct aac aga gga tat ggg gat	tac aca cca cta cca	gct ggc tat	3429
Ser Asn Arg Gly Tyr Gly Asp	Tyr Thr Pro Leu Pro	Ala Gly Tyr	
1130	1135	1140	
gtg aca aaa gaa tta gag tac	ttc cca gaa acc gat	aag gta tgg	3474
Val Thr Lys Glu Leu Glu Tyr	Phe Pro Glu Thr Asp	Lys Val Trp	
1145	1150	1155	
att gag att gga gaa aca gaa	gga aca ttc att gtg	gat agt gtg	3519
Ile Glu Ile Gly Glu Thr Glu	Gly Thr Phe Ile Val	Asp Ser Val	
1160	1165	1170	
gaa tta ctc ctt atg gag gaa	taa		3543
Glu Leu Leu Leu Met Glu Glu			
1175	1180		

<210> 2  
 <211> 1180  
 <212> PRT  
 <213> Bacillus thuringiensis

<400> 2

Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn  
 1 5 10 15

Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn  
 20 25 30

UA 128156 C2

Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu  
35 40 45

Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp  
50 55 60

Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu  
65 70 75 80

Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile  
85 90 95

Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala  
100 105 110

Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala  
115 120 125

Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile  
130 135 140

Ser Val Leu Lys Ile Gln Asn Phe Glu Val Gln Leu Leu Ser Val Phe  
145 150 155 160

Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe  
165 170 175

Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr  
180 185 190

Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg  
195 200 205

Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp  
210 215 220

Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu  
225 230 235 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile  
245 250 255

UA 128156 C2

Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu  
260 265 270

Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln  
275 280 285

Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile  
290 295 300

Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile  
305 310 315 320

Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu  
325 330 335

Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr  
340 345 350

Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile  
355 360 365

Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly  
370 375 380

Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile  
385 390 395 400

Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln  
405 410 415

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

Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala  
435 440 445

Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile  
450 455 460

Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile  
465 470 475 480

Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp



UA 128156 C2

Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys  
725 730 735

Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly  
740 745 750

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn  
755 760 765

Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro  
770 775 780

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys  
785 790 795 800

Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp  
805 810 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp  
820 825 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe  
835 840 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe  
850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg  
865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr  
885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val  
900 905 910

Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile  
915 920 925

His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro  
930 935 940

UA 128156 C2

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu  
 945 950 955 960

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val  
 965 970 975

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys  
 980 985 990

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val  
 995 1000 1005

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 1010 1015 1020

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1025 1030 1035

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1040 1045 1050

Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1055 1060 1065

Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu  
 1070 1075 1080

Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
 1085 1090 1095

Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr  
 1100 1105 1110

Glu Gln Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu  
 1115 1120 1125

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1130 1135 1140

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1145 1150 1155

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1160 1165 1170

Glu Leu Leu Leu Met Glu Glu  
 1175 1180

<210> 3  
 <211> 3534  
 <212> ДНК  
 <213> Штучна послідовність

<220>  
 <223> Повність синтетична ДНК, яка кодує BCW 002

<220>  
 <221> CDS  
 <222> (1)..(3531)

<400> 3  
 atg gat aac aat ccg aac atc aat gaa tgc att cct tat aat tgt tta 48  
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15  
 agt aac cct gaa gta gaa gta tta ggt gga gaa aga ata gaa aat ggt 96  
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30  
 tac acc cca atc gat att tcc ttg tcc cta acg caa ttt ctt ttg agt 144  
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
 35 40 45  
 gaa ttt gtt ccc ggt gct gga ttt gtg tta gga cta gtt gat ata ata 192  
 Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
 50 55 60  
 tgg gga att ttt ggt ccc tct caa tgg gac gca ttt ctt gta caa att 240  
 Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
 65 70 75 80  
 gaa cag tta att aac caa aga ata gaa gaa ttc gct agg aac caa gcc 288  
 Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95  
 att tct aga tta gaa gga cta agc aat ctt tat caa att tac gca gaa 336  
 Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
 100 105 110  
 tct ttt aga gag tgg gaa gca gat cct act aat cca gca tta aga gaa 384  
 Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
 115 120 125  
 gag atg cgt att caa ttc aat gac atg aac agt gcc ctt aca acc gct 432  
 Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
 130 135 140

UA 128156 C2

att cct ctt ttg gca gtt caa aat tat caa gtt cct ctt tta tca gta	480
Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val	
145 150 155 160	
tat gtt caa gct gca aat tta cat tta tca gtt ttg aga gat gtt tca	528
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser	
165 170 175	
gtg ttt gga caa agg tgg gga ttt gat gcc gcg act atc aat agt cgt	576
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg	
180 185 190	
tat aat gat tta act agg ctt att ggc aac tat aca gat tat gct gta	624
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val	
195 200 205	
cgc tgg tac aat acg gga tta gaa cgt gta tgg gga ccg gat tct aga	672
Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg	
210 215 220	
gat tgg gta agg tat aat caa ttt aga aga gaa tta aca cta act gta	720
Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val	
225 230 235 240	
tta gat atc gtt gct ctg ttc ccg aat tat gat agt aga aga tat cca	768
Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro	
245 250 255	
att cga ata gtt tcc caa tta aca aga gaa att tat acg aac cca gta	816
Ile Arg Ile Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val	
260 265 270	
tta gaa aat ttt gat ggt agt ttt cgt gga atg gct cag aga ata gaa	864
Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu	
275 280 285	
cag aat att agg caa cca cat ctt atg gat atc ctt aat agt ata acc	912
Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr	
290 295 300	
att tat act gat gtg cat aga ggc ttt aat tat tgg tca ggg cat caa	960
Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln	
305 310 315 320	
ata aca gct tct cct gta ggg ttt tca gga cca gaa ttc gca ttc cct	1008
Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro	
325 330 335	
tta ttt ggg aat gcg gga aat gca gct cca ccc gta ctt gtc tca tta	1056
Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu	
340 345 350	
act ggt ttg ggg att ttt aga aca tta tct tca cct tta tat aga aga	1104
Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg	
355 360 365	

UA 128156 C2

att ata ctt ggt tca ggc cca aat aat cag gaa ctg ttt gtc ctt gat	1152
Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp	
370 375 380	
gga acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act	1200
Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr	
385 390 395 400	
ata tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccc cca	1248
Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro	
405 410 415	
cag gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt	1296
Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser	
420 425 430	
cat gtt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga	1344
His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg	
435 440 445	
gct cca acg ttt tct tgg cag cat cgc agt gct acg aca act aat ata	1392
Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile	
450 455 460	
att gca gcg gat agt att act caa att cct gct gtt aaa gga cgt tct	1440
Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser	
465 470 475 480	
att att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc	1488
Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly	
485 490 495	
gat ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca	1536
Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala	
500 505 510	
tac ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt	1584
Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg	
515 520 525	
gtt cgt gtt cgt tat gct tct aca tct tca ctc cct gta gat gta gtt	1632
Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val	
530 535 540	
ttc gga aat att agt cat cct act aca ttc cca gcc act gcc aga tca	1680
Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser	
545 550 555 560	
tta gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga	1728
Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly	
565 570 575	
act ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta	1776
Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu	
580 585 590	
tct act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act	1824

UA 128156 C2

Ser	Thr	Ile	Asn	Leu	Ile	Val	Asp	Arg	Phe	Glu	Phe	Ile	Pro	Val	Thr		
		595					600					605					
gca	acc	ttt	gaa	gca	gaa	tgg	gat	tta	gaa	aga	gca	caa	aag	gcg	gtg		1972
Ala	Thr	Phe	Glu	Ala	Glu	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val		
		610				615					620						
aat	gcg	ctg	ttt	act	tct	aca	sac	caa	cta	ggg	ata	aaa	aca	gat	gtg		1920
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Ile	Lys	Thr	Asp	Val		
		625			630					635					640		
acg	gat	tat	cat	att	gat	caa	gtg	tcc	aat	tta	gtg	gag	tgt	tta	tgg		1968
Thr	Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Glu	Cys	Leu	Ser		
				645					650					655			
gat	gaa	ttt	tat	ctg	gat	gaa	aag	cga	gaa	ttg	tcc	gag	aaa	gtc	aaa		2016
Asp	Glu	Phe	Tyr	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys		
			660					665					670				
cat	gcg	aag	cga	ctc	agt	gat	gag	cga	aat	tta	ctt	caa	gat	cca	sac		2064
His	Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn		
			675				680					685					
ttc	agg	ggc	atc	aat	aga	caa	cca	gat	cgt	ggc	tgg	aga	gga	agt	acg		2112
Phe	Arg	Gly	Ile	Asn	Arg	Gln	Pro	Asp	Arg	Gly	Trp	Arg	Gly	Ser	Thr		
		690				695					700						
gat	att	acc	atc	caa	gga	gga	gat	gac	gta	ttc	aaa	gag	aat	tac	gtc		2160
Asp	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val		
					710					715					720		
aca	cta	cca	ggt	acc	ttt	gat	gag	tgc	tat	cca	aca	tat	ttg	tat	caa		2208
Thr	Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln		
				725					730					735			
aaa	atc	gat	gaa	tca	aaa	tta	aaa	gcc	ttt	acc	cgt	tat	caa	tta	aga		2256
Lys	Ile	Asp	Gln	Ser	Lys	Leu	Lys	Ala	Phe	Thr	Arg	Tyr	Gln	Leu	Arg		
			740					745					750				
ggg	tat	atc	gaa	gat	agt	caa	gac	tta	gaa	atc	tat	tta	att	cgc	tac		2304
Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Leu	Ile	Arg	Tyr		
			755				760						765				
aat	gca	aaa	cat	gaa	aca	gta	aat	gtg	cca	ggt	acg	ggt	tcc	tta	tgg		2352
Asn	Ala	Lys	His	Glu	Thr	Val	Asn	Val	Pro	Gly	Thr	Gly	Ser	Leu	Trp		
			770			775						780					
cgg	ctt	tca	goc	caa	agt	cca	atc	gga	aag	tgt	gga	gag	cgg	aat	cga		2400
Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly	Glu	Pro	Asn	Arg		
					790					795					800		
tgc	gcg	cca	cac	ctt	gaa	tgg	aat	cct	gac	tta	gat	tgt	tgg	tgt	agg		2448
Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg		
				805					810					815			
gat	gga	gaa	aag	tgt	gcc	cat	cat	tgg	cat	cat	ttc	tcc	tta	gac	att		2496
Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile		

UA 128156 C2

	820		825		830	
gat gta gga tgt aca gac tta aat gag gac cta ggt gta tgg gtg atc						2544
Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile						
	835		840		845	
ttt aag att aag acg caa gat ggg cac gca aga cta ggg aat cta gag						2592
Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu						
	850		855		860	
ttt ctc gaa gag aaa cca tta gta gga gaa gca cta gct cgt gtg aaa						2640
Phe Leu Gln Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys						
	865		870		875	880
aga gcg gag aaa aaa tgg aga gac aaa cgt gaa aaa ttg gaa tgg gaa						2688
Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu						
		885		890		895
aca aat atc gtt tat aaa gag gca aaa gaa tct gta gat gct tta ttt						2736
Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe						
		900		905		910
gta aac tct caa tat gat caa tta caa gcg gat acg aat att gcc atg						2784
Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met						
		915		920		925
att cat gcg gca gat aaa cgt gtt cat agc att cga gaa gct tat ctg						2832
Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu						
		930		935		940
ccg gag ctg tct gtg att ccg ggt gtc aat gcg gct att ttt gaa gaa						2880
Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu						
		945		950		955
ttt gaa ggg cgt att ttc act gca ttc tcc cta tat gat gcg aga aat						2928
Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn						
		965		970		975
gtc att aaa aat ggt gat ttt aat aat gcc tta tcc tgc tgg aac gtg						2976
Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val						
		980		985		990
aaa ggg cat gta gat gta gaa gaa caa aac aac caa cgt tcc gtc ctt						3024
Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu						
		995		1000		1005
gtt gtt ccg gaa tgg gaa gca gaa gtg tca caa gaa gtt cgt gtc						3069
Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val						
		1010		1015		1020
tgt ccg ggt cgt gcc tat atc ctt cgt gtc aca gcg tac aag gag						3114
Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu						
		1025		1030		1035
gga tat gga gaa ggt tgc gta acc att cat gag atc gag aac aat						3159
Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn						
		1040		1045		1050

UA 128156 C2

aca gac gaa ctg aag ttt agc aac tgc gta gaa gag gaa atc tat 3204  
 Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065

cca aat aac acg gta acg tgt aat gat tat act gta aat caa gaa 3249  
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Glu Glu  
 1070 1075 1080

gaa tac gga ggt gcg tac act tct cgt aat cga gga tat aac gaa 3294  
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095

gct cct tcc gta cca gct gat tat gcg tca gtc tat gaa gaa aaa 3339  
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110

tcg tat aca gat gga cga aga gag aat cct tgt gaa ttt aac aga 3384  
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125

ggg tat agg gat tac acg cca cta cca gtt ggt tat gtg aca aaa 3429  
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140

gaa tta gaa tac ttc cca gaa acc gat aag gta tgg att gag att 3474  
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155

gga gaa acg gaa gga aca ttt atc gtg gac agc gtg gaa tta ctc 3519  
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170

ctt atg gag gaa tag 3534  
 Leu Met Glu Glu  
 1175

<210> 4  
 <211> 1177  
 <212> PRT  
 <213> Штучна послідовність

<220>  
 <223> Синтетична конструкція

<400> 4

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser

UA 128156 C2

			35					40				45			
Glu	Phe	Val	Pro	Gly	Ala	Gly	Phe	Val	Leu	Gly	Leu	Val	Asp	Ile	Ile
	50					55					60				
Trp	Gly	Ile	Phe	Gly	Pro	Ser	Gln	Trp	Asp	Ala	Phe	Leu	Val	Gln	Ile
65					70				75					80	
Glu	Gln	Leu	Ile	Asn	Gln	Arg	Ile	Glu	Glu	Phe	Ala	Arg	Asn	Gln	Ala
				85					90					95	
Ile	Ser	Arg	Leu	Glu	Gly	Leu	Ser	Asn	Leu	Tyr	Gln	Ile	Tyr	Ala	Glu
			100					105					110		
Ser	Phe	Arg	Glu	Trp	Glu	Ala	Asp	Pro	Thr	Asn	Pro	Ala	Leu	Arg	Glu
		115					120					125			
Glu	Met	Arg	Ile	Gln	Phe	Asn	Asp	Met	Asn	Ser	Ala	Leu	Thr	Thr	Ala
	130					135					140				
Ile	Pro	Leu	Leu	Ala	Val	Gln	Asn	Tyr	Gln	Val	Pro	Leu	Leu	Ser	Val
145					150					155					160
Tyr	Val	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Val	Leu	Arg	Asp	Val	Ser
				165					170					175	
Val	Phe	Gly	Gln	Arg	Trp	Gly	Phe	Asp	Ala	Ala	Thr	Ile	Asn	Ser	Arg
			180					185					190		
Tyr	Asn	Asp	Leu	Thr	Arg	Leu	Ile	Gly	Asn	Tyr	Thr	Asp	Tyr	Ala	Val
		195					200					205			
Arg	Trp	Tyr	Asn	Thr	Gly	Leu	Glu	Arg	Val	Trp	Gly	Pro	Asp	Ser	Arg
	210					215					220				
Asp	Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Gln	Leu	Thr	Leu	Thr	Val
225					230					235					240
Leu	Asp	Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Ser	Arg	Arg	Tyr	Pro
				245					250					255	
Ile	Arg	Ile	Val	Ser	Gln	Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Asn	Pro	Val
			260					265					270		

UA 128156 C2

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu  
 275 280 285

Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr  
 290 295 300

Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln  
 305 310 315 320

Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro  
 325 330 335

Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu  
 340 345 350

Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg  
 355 360 365

Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp  
 370 375 380

Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr  
 385 390 395 400

Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro  
 405 410 415

Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser  
 420 425 430

His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg  
 435 440 445

Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile  
 450 455 460

Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser  
 465 470 475 480

Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly  
 485 490 495

UA 128156 C2

Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala  
 500 505 510  
 Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg  
 515 520 525  
 Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val  
 530 535 540  
 Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser  
 545 550 555 560  
 Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly  
 565 570 575  
 Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu  
 580 585 590  
 Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr  
 595 600 605  
 Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val  
 610 615 620  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val  
 625 630 635 640  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 645 650 655  
 Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys  
 660 665 670  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 675 680 685  
 Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr  
 690 695 700  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 705 710 715 720

UA 128156 C2

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 725 730 735  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg  
 740 745 750  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 755 760 765  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 770 775 780  
 Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg  
 785 790 795 800  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg  
 805 810 815  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 820 825 830  
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
 835 840 845  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
 850 855 860  
 Phe Leu Glu Gln Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys  
 865 870 875 880  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu  
 885 890 895  
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 900 905 910  
 Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met  
 915 920 925  
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu  
 930 935 940  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu



Leu Met Glu Glu  
1175

<210> 5  
<211> 3534  
<212> ДНК  
<213> Штучна послідовність

<220>  
<223> Повністю синтетична послідовність, яка кодує BCW 003

<220>  
<221> CBS  
<222> (1)..(3531)

<400> 5  
atg gat aac aat ccg aac atc aat gaa tgc att cct tat aat tgt tta 48  
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
1 5 10 15  
agt aac cct gaa gta gaa gta tta ggt gga gaa aga ata gaa act ggt 96  
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
20 25 30  
tac acc cca atc gat att tcc ttg tgg cta acg caa ttt ctt ttg agt 144  
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
35 40 45  
gaa ttt gtt ccc ggt gct gga ttt gtg tta gga cta gtt gat ata ata 192  
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
50 55 60  
tgg gga att ttt ggt ccc tct caa tgg gac gca ttt ctt gta caa att 240  
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
65 70 75 80  
gaa cag tta att aac caa aga ata gaa gaa ttc gct agg aac caa gcc 298  
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
85 90 95  
att tct aga tta gaa gga cta agc aat ctt tat caa att tac gca gaa 336  
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
100 105 110  
tct ttt sga gag tgg gaa gca gat cct act aat cca gca tta aga gaa 384  
Ser Phe Arg Gln Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
115 120 125  
gag atg cgt att caa ttc aat gac atg aac agt gcc ctt aca acc gct 432  
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
130 135 140  
att cct ctt ttg gca gtt caa aat tat caa gtt cct ctt tta tca gta 480

UA 128156 C2

Ile	Pro	Leu	Leu	Ala	Val	Gln	Asn	Tyr	Gln	Val	Pro	Leu	Leu	Ser	Val	
145					150					155					160	
tat	ggt	caa	gct	gca	aat	tta	cat	tta	tca	ggt	ttg	aga	gat	gtt	tca	528
Tyr	Val	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Val	Leu	Arg	Asp	Val	Ser	
			165						170					175		
gtg	ttt	gga	caa	agg	tgg	gga	ttt	gat	gcc	gcg	act	atc	aat	agt	ogt	576
Val	Phe	Gly	Gln	Arg	Trp	Gly	Phe	Asp	Ala	Ala	Thr	Ile	Asn	Ser	Arg	
			180					185					190			
tat	aat	gat	tta	act	agg	ctt	att	ggc	aac	tat	aca	gat	tat	gct	gta	624
Tyr	Asn	Asp	Leu	Thr	Arg	Leu	Ile	Gly	Asn	Tyr	Thr	Asp	Tyr	Ala	Val	
		195					200					205				
cgc	tgg	tac	aat	acg	gga	tta	gaa	ogt	gta	tgg	gga	ccg	gat	tct	aga	672
Arg	Trp	Tyr	Asn	Thr	Gly	Leu	Glu	Arg	Val	Trp	Gly	Pro	Asp	Ser	Arg	
	210					215					220					
gat	tgg	gta	agg	tat	aat	caa	ttt	aga	aga	gaa	tta	aca	cta	act	gta	720
Asp	Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Glu	Leu	Thr	Leu	Thr	Val	
	225				230					235				240		
tta	gat	atc	ggt	gct	ctg	ttc	ccg	aat	tat	gat	agt	aga	aga	tat	cca	768
Leu	Asp	Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Ser	Arg	Arg	Tyr	Pro	
				245					250					255		
att	cga	aca	ggt	tcc	caa	tta	aca	aga	gaa	att	tat	acg	aac	cca	gta	816
Ile	Arg	Thr	Val	Ser	Gln	Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Asn	Pro	Val	
			260					265					270			
tta	gaa	aat	ttt	gat	ggt	agt	ttt	cgt	gga	atg	got	cag	aga	ata	gaa	864
Leu	Glu	Asn	Phe	Asp	Gly	Ser	Phe	Arg	Gly	Met	Ala	Gln	Arg	Ile	Glu	
		275					280					285				
cag	aat	att	agg	caa	cca	cat	ctt	atg	gat	atc	ctt	aat	agt	ata	acc	912
Gln	Asn	Ile	Arg	Gln	Pro	His	Leu	Met	Asp	Ile	Leu	Asn	Ser	Ile	Thr	
	290					295					300					
att	tat	act	gat	gtg	cat	aga	ggc	ttt	aat	tat	tgg	tca	ggg	cat	caa	960
Ile	Tyr	Thr	Asp	Val	His	Arg	Gly	Phe	Asn	Tyr	Trp	Ser	Gly	His	Gln	
	305				310					315				320		
ata	aca	gct	tct	ccg	gta	ggg	ttt	tca	gga	cca	gaa	ttc	gca	ttc	ccg	1008
Ile	Thr	Ala	Ser	Pro	Val	Gly	Phe	Ser	Gly	Pro	Glu	Phe	Ala	Phe	Pro	
				325					330					335		
tta	ttt	ggg	aat	gcg	gga	aat	gca	got	cca	ccc	gta	ctt	gtc	tca	tta	1056
Leu	Phe	Gly	Asn	Ala	Gly	Asn	Ala	Ala	Pro	Pro	Val	Leu	Val	Ser	Leu	
			340					345					350			
act	ggt	ttg	ggg	att	ttt	aga	aca	tta	tct	tca	ccg	tta	tat	aga	aga	1104
Thr	Gly	Leu	Gly	Ile	Phe	Arg	Thr	Leu	Ser	Ser	Pro	Leu	Tyr	Arg	Arg	
		355					360					365				
att	ata	ctt	ggt	tca	ggc	cca	aat	aat	cag	gaa	ctg	ttt	gtc	ctt	gat	1152
Ile	Ile	Leu	Gly	Ser	Gly	Pro	Asn	Asn	Gln	Glu	Leu	Phe	Val	Leu	Asp	

UA 128156 C2

370	375	380	
gga acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act			1200
Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr			
385	390	395	400
ata tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccg cca			1248
Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro			
	405	410	415
cag gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt			1296
Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser			
	420	425	430
cat gtt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga			1344
His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg			
	435	440	445
gct cca acg ttt tct tgg cag cat cgc agt gct acg aca act aat ata			1392
Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile			
	450	455	460
att gca gcg gat agt att act caa att cct gct gtt aaa gga cgt tct			1440
Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser			
	465	470	475
att att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc			1488
Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly			
	485	490	495
gat ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca			1536
Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala			
	500	505	510
tac ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt			1584
Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg			
	515	520	525
gtt cgt gtt cgt tat gct tct aca tct tca ctc cct gta gat gta gtt			1632
Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val			
	530	535	540
ttc gga aat att agt cat cct act aca ttc cca gcc act gcc aga tca			1680
Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser			
	545	550	555
tta gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga			1728
Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly			
	565	570	575
act ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta			1776
Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu			
	580	585	590
tct act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act			1824
Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr			
	595	600	605

gca acc ttt gaa gca gaa tcg gat tta gaa aga gca caa aag gcg gtg Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val 610 615 620	1972
aat gcg ctg ttt act tct ace aac caa cta ggg ata aaa aca gat gtg Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val 625 630 635 640	1920
acg gat tat cat att gat caa gtg tcc aat tta gtg gag tgt tta tcg Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser 645 650 655	1968
gat gaa ttt tat ctg gat gaa aag cga gaa ttg tcc gag aaa gtc aaa Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys 660 665 670	2016
cat gcg aag cga ctg agt gat gag cga aat tta ctt caa gat cca aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn 675 680 685	2064
ttc agg ggc atc aat aga caa cca gat cgt ggc tgg aga gga agt acg Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr 690 695 700	2112
gat att acc atc caa gga gga gat gac gta ttc aaa gag aat tac gtc Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val 705 710 715 720	2160
aca cta cca ggt acc ttt gat gag tgc tat cca aca tat ttg tat caa Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln 725 730 735	2208
aaa atc gat gaa tca aaa tta aaa gcc ttt acc cgt tat caa tta aga Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg 740 745 750	2256
ggg tat atc gaa gat agt caa gac tta gaa atc tat tta att cgc tac Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr 755 760 765	2304
aat gca aaa cat gaa aca gta aat gtg cca ggt acg ggt tcc tta tgg Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp 770 775 780	2352
ccg ctt tca gcc caa agt cca atc gga aag tgt gga gag ccg aat cga Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg 785 790 795 800	2400
tgc gcg cca cac ctt gaa tgg aat cct gac tta gat tgt tcg tgt agg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg 805 810 815	2448
gat gga gaa aag tgt gcc cat cat tcg cat cat ttc tcc tta gac att Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile 820 825 830	2496

UA 128156 C2

gat gta gga tgt aca gac tta aat gag gac cta ggt gta tgg gtg atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845	2544
ttt aag att aag acg caa gat ggg cac gca aga cta ggg aat cta gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Gln 850 855 860	2582
ttt ctc gaa gag aaa cca tta gta gga gaa gca cta gct cgt gtg aaa Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys 865 870 875 880	2640
aga gcg gag aaa aaa tgg aga gac aaa cgt gaa aaa ttg gaa tgg gaa Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu 885 890 895	2688
aca aat atc gtt tat aaa gag gca aaa gaa tct gta gat got tta ttt Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe 900 905 910	2736
gta aac tct caa tat gat caa tta caa gcg gat acg aat att gcc atg Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met 915 920 925	2784
att cat gcg gca gat aaa cgt gtt cat agc att cga gaa got tat ctg Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu 930 935 940	2832
cct gag ctg tct gtg att ccg ggt gtc aat gcg gct att ttt gaa gaa Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu 945 950 955 960	2880
tta gaa ggg cgt att ttc act gca ttc tcc cta tat gat gcg aga aat Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn 965 970 975	2928
gtc att aaa aat ggt gat ttt aat aat ggc tta tcc tgc tgg aac gtg Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val 980 985 990	2976
aaa ggg cat gta gat gta gaa gaa caa aac aac caa cgt tog gtc ott Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu 995 1000 1005	3024
gtt gtt ccg gaa tgg gaa gca gaa gtg tca caa gaa gtt cgt gtc Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val 1010 1015 1020	3069
tgt ccg ggt cgt ggc tat atc ett cgt gtc aca gcg tac aag gag Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu 1025 1030 1035	3114
gga tat gga gaa ggt tgc gta acc att cat gag atc gag aac aat Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn 1040 1045 1050	3159
aca gac gaa ctg aag ttt agc aac tgc gta gaa gag gaa atc tat	3204

UA 128156 C2

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065  
 cca aat aac acg gta acg tgt aat gat tat act gta aat caa gaa 3249  
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Glu Glu  
 1070 1075 1080  
 gaa tac gga ggt gcg tac act tct cgt aat cga gga tat aac gaa 3284  
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090  
 gct cct tcc gta cca gct gat tat gcg tca gtc tat gaa gaa aaa 3319  
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110  
 tcg tat aca gat gga cga aga gag aat cct tgt gaa ttt aac aga 3384  
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125  
 ggg tat agg gat tac acg cca cta cca gtt ggt tat gtg aca aaa 3429  
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140  
 gaa tta gaa tac ttc cca gaa acc gat aag gta tgg att gag att 3474  
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155  
 gga gaa acg gaa gga aca ttt atc gtg gac agc gtg gaa tta ctc 3519  
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170  
 ctt atg gag gaa tag 3534  
 Leu Met Glu Glu  
 1175

<210> 6  
 <211> 1177  
 <212> PRT  
 <213> Штучна послідовність  
 <220>  
 <223> Синтетична конструкція  
 <400> 6

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15  
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30  
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Glu Phe Leu Leu Ser  
 35 40 45

UA 128156 C2

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
 130 135 140

Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val  
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser  
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg  
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val  
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg  
 210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val  
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro  
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val  
 260 265 270

UA 128156 C2

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu  
 275 280 285  
 Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr  
 290 295 300  
 Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln  
 305 310 315 320  
 Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro  
 325 330 335  
 Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu  
 340 345 350  
 Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg  
 355 360 365  
 Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp  
 370 375 380  
 Gly Thr Gln Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr  
 385 390 395 400  
 Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro  
 405 410 415  
 Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser  
 420 425 430  
 His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg  
 435 440 445  
 Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile  
 450 455 460  
 Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser  
 465 470 475 480  
 Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly  
 485 490 495  
 Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala

UA 128156 C2

				500						505										510	
Tyr	Leu	Glu	Val	Pro	Ile	Phe	Phe	Gln	Ser	Pro	Ser	Thr	Asn	Tyr	Arg						
		515						520					525								
Val	Arg	Val	Arg	Tyr	Ala	Ser	Thr	Ser	Ser	Leu	Pro	Val	Asp	Val	Val						
		530				535					540										
Phe	Gly	Asn	Ile	Ser	His	Pro	Thr	Thr	Phe	Pro	Ala	Thr	Ala	Arg	Ser						
545					550					555					560						
Leu	Asp	Asn	Leu	Gln	Ser	Asn	Asp	Phe	Gly	Tyr	Ile	Asp	Ile	Ala	Gly						
				565					570					575							
Thr	Phe	Leu	Pro	Ser	Leu	Gly	Pro	Ser	Ile	Gly	Ile	Arg	Pro	Met	Leu						
			580					585					590								
Ser	Thr	Ile	Asn	Leu	Ile	Val	Asp	Arg	Phe	Glu	Phe	Ile	Pro	Val	Thr						
		595					600					605									
Ala	Thr	Phe	Glu	Ala	Glu	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val						
		610				615					620										
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Ile	Lys	Thr	Asp	Val						
625					630					635					640						
Thr	Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Glu	Cys	Leu	Ser						
				645					650					655							
Asp	Glu	Phe	Tyr	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys						
			660					665					670								
His	Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn						
		675					680					685									
Phe	Arg	Gly	Ile	Asn	Arg	Gln	Pro	Asp	Arg	Gly	Trp	Arg	Gly	Ser	Thr						
		690				695					700										
Asp	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val						
705					710					715					720						
Thr	Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln						
				725					730						735						

UA 128156 C2

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg  
 740 745 750  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 755 760 765  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 770 775 780  
 Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg  
 785 790 795 800  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg  
 805 810 815  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 820 825 830  
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
 835 840 845  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
 850 855 860  
 Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys  
 865 870 875 880  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu  
 885 890 895  
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 900 905 910  
 Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met  
 915 920 925  
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu  
 930 935 940  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
 945 950 955 960

UA 128156 C2

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val  
 980 985 990

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu  
 995 1000 1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 1010 1015 1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1025 1030 1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1040 1045 1050

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065

Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu  
 1070 1075 1080

Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095

Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110

Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170

Leu Met Glu Glu  
1175

<210> 7  
<211> 3543  
<212> ДНК  
<213> Штучна послідовність

<220>  
<223> Повністю синтетична послідовність для застосування в рослині, яка кодує BCW 001

<220>  
<221> CDS  
<222> (1)..(3540)  
<223> BCW 001

<400> 7  
atg gag gag aac aac cag aac cag tgt gtc cca tac aac tgc ctc aac 48  
Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn  
1 5 10 15  
aac cca gcc atc gaa atc ctt gag ggc gac cga att tca gtc ggc aac 96  
Asn Pro Ala Ile Glu Ile Leu Gln Gly Asp Arg Ile Ser Val Gly Asn  
20 25 30  
acg ccc atc gac atc tcc ctg agt ctt gtg gaa ctc ctc atc tcc gag 144  
Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu  
35 40 45  
ttc gtc cct ggc ggc ggc ata atc acc ggt ctg ctc aac atc gtg tgg 192  
Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp  
50 55 60  
gga ttc gtg ggc cca tcc cag tgg gat gcg ttc ctg gcc caa gtg gag 240  
Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu  
65 70 75 80  
cag ctc atc aac cag agg atc tcc gag gcc gtc cgc aat acc gcg atc 288  
Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile  
85 90 95  
caa gag ctg gag ggc atg gcc cgc gtg tac cgc acc tac gcc acc gcc 336  
Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala  
100 105 110  
ttt gct gaa tgg gag cgc gac ccg aac aac act gac ctg cgc gag gcc 384  
Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala  
115 120 125  
gtc cga acc cag ttc acg gcg acc gag acc tac atc agc ggc cgg atc 432  
Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile  
130 135 140  
tca gtg ctc aag atc cag aac ttc gag gtg cag ctc cta tcc gtc ttc 480

UA 128156 C2

Ser Val Leu Lys Ile Gln Asn Phe Glu Val Gln Leu Leu Ser Val Phe	
145	150 155 160
ggc cag gcc gcc aac ttg cac ctg agc etc ctg cgg gac gtt gtg ttc	528
Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe	
	165 170 175
ttc gcc cag cgg tgg gcc ttc tct act acg acc gtg aac aac tac tac	576
Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Val Asn Asn Tyr Tyr	
	180 185 190
aac gac ctg acg gag gaa atc agc acc tac aca gat tac gca gtt cgt	624
Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg	
	195 200 205
tgg tac aac acc gcc ctt gag cgc gtg tgg gcc cgg gac tcc cgc gat	672
Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp	
	210 215 220
tgg gtc cgc tac aac cag ttc cgc cgc gag ctg acg ctt aca gtg ctg	720
Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu	
	225 230 235 240
gac atc gtc gca ctc ttt cct aac tac gac tcc agg cgc tat ccc atc	768
Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile	
	245 250 255
agg aca gtg tca cag ctc acc cgc gag atc tac aca aac cgg gtg ctt	816
Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu	
	260 265 270
gag aac ttc gac gcc agc ttc cgt gcc atg gcc cag cgc att gaa cag	864
Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln	
	275 280 285
aac atc cgc cag ccg cac ctt atg gac atc ttg aac agt atc act atc	912
Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile	
	290 295 300
tac acc gac gtc cac aga gcc ttc aac tac tgg agc gga cac cag atc	960
Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile	
	305 310 315 320
aca gcc agc ccg gta gcc ttc tgg ggt cca gag ttc gcc ttc ccg ctg	1008
Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu	
	325 330 335
ttt ggg aac gct gcc aat gcc gcc ccg ccc gtg ctg gtc agc ctc act	1056
Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr	
	340 345 350
ggt ctc gcc atc ttc cgc aca ctt tcc tgg ccg ctg tac agg agg atc	1104
Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile	
	355 360 365
atc ctc ggg tcc ggt ccg aac aac cag gag ctg ttc gtg ctc gac ggg	1152
Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly	

UA 128156 C2

370	acc gag ttc agt ttc gcc agc ctc acg acg aac ctc ccg tcc acc atc	1200
	Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile	
	385 390 395 400	
	tat cgc cag cgc gga acg gtc gat tcc ctg gat gtt atc cca ccg caa	1248
	Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln	
	405 410 415	
	gac aat tct gtg ccg ccg agg gcc ggg ttc tcc cac ccg ctg tct cac	1296
	Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His	
	420 425 430	
	gtg act atg ctt tca cag gcc gcc gga gcc gtg tac acg ctc cgt gcg	1344
	Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala	
	435 440 445	
	cct act ttc tcc tgg cag cac cgc agc gcg acc acg acc aac atc atc	1392
	Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile	
	450 455 460	
	gca gca gac tcc atc acc cag atc ccg gcc gtt aag ggc cgc agc atc	1440
	Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile	
	465 470 475 480	
	atc aac aac gga act gtc atc agc ggt ccg ggc ttc acg ggc ggc gac	1488
	Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp	
	485 490 495	
	ctg gtc ccg ctc tac aac gca gac ttc aac atc aat aac cgc gct tat	1536
	Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr	
	500 505 510	
	ctt gaa gta cct atc ttc ttc cag agc ccg agc act aac tac ccg gtt	1584
	Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg Val	
	515 520 525	
	cgc gtc cgc tac gcc agc acc tcc agc ctc cct gtc gat gtc gtg ttc	1632
	Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val Phe	
	530 535 540	
	gga aac ata agc cat ccg acc acg ttc cca gcc acg gct agg agc ctg	1680
	Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser Leu	
	545 550 555 560	
	gac aac cta cag agc aac gac ttc ggc tac atc gac atc gcg ggc acc	1728
	Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly Thr	
	565 570 575	
	tct ctg cca agc ctg ggt ccg tct atc ggc atc cgc ccg atg ctg agc	1776
	Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu Ser	
	580 585 590	
	act atc aac cta att gtg gac ccg ttc gag ttt atc ccg gtg acg gca	1824
	Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr Ala	
	595 600 605	

UA 128156 C2

acg ttc gag gcg gag tct gac ctc gaa agg gca cag aag gcc gtg aac	1872
Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn	
610 615 620	
gcc ctg ttc acg agc acc aac cag ctt ggc att aag act gat gtc acc	1920
Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val Thr	
625 630 635 640	
gac tac cac att gac caa gtc agc aac ctg gtg gag tgc ctc tcg gac	1968
Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp	
645 650 655	
gag ttc tat ctt gat gag aaa cgg gaa cta agc gag aag gtg aag cac	2016
Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His	
660 665 670	
gca aag cgc ttg agc gac gag cgg aac tta ctc cag gac cct aac ttc	2064
Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe	
675 680 685	
cgt ggg att aac cgc cag ccg gat cgc ggg tgg cgc ggc tca acg gac	2112
Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp	
690 695 700	
atc acc atc cag gcc gcc gat gac gtc ttc aag gag aac tac gtg acc	2160
Ile Thr Ile Gln Gly Gln Asp Asp Val Phe Lys Glu Asn Tyr Val Thr	
705 710 715 720	
ctc cct gcc acg ttc gac gag tgc tac ccg acg tac ctt tat cag aag	2208
Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys	
725 730 735	
att gac gaa agc aag ctg aaa gcc tac acc cgc tac cag ttg cgc gcc	2256
Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly	
740 745 750	
tac atc gag gac tct caa gac ctg gag atc tac ttg att cga tac aac	2304
Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn	
755 760 765	
ggg aaa cac gag acc gtc aac gtg ccg gcc act ggg agc ctg tgg ccg	2352
Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro	
770 775 780	
ttg tct gca caa agt ccg atc gcc aag tgc gcc gag cca aac cgg tgc	2400
Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys	
785 790 795 800	
gct ccg cac ctg gag tgg aac cca gac ctt cat tgc tcc tgt agg gat	2448
Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp	
805 810 815	
ggc gag aag tgc gct cac cac agc cat cac ttc agc ctc gac att gac	2496
Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp	
820 825 830	

UA 128156 C2

gtc ggc tgt acc gac ctt aat gag gat ctg ggt gtg tgg gtg atc ttc Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe 835 840 845	2544
aag atc aag acc cag gac ggt cac gcc cgg ttg ggc aat ctg gag ttc Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe 850 855 860	2592
ctg gag gag aag ccg ctg gtt ggc gag gct ctc gcg cgg gtc aag cgg Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg 865 870 875 880	2640
gcg gag aag aag tgg cgg gac aaa cgc gag aag ctc cag tta gag acg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr 885 890 895	2688
aac atc gtg tac aag gag gcg aag gaa tcc gtg gac gca cta ttc gtg Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val 900 905 910	2736
aac agc cag tac gac caa ctc cag gcc gac acc aac atc gcc atg att Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile 915 920 925	2784
cac gca gcc gac aag agg gtg cac cgc atc cgc gaa gcc tac ctt ccc His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro 930 935 940	2832
gaa ctt tcg gtc atc cca gcc gtc aac gct gac atc tcg gag gaa ttg Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu 945 950 955 960	2880
gag ggc aga atc ttc acg gcc ttc tot ttg tac gat gcc agy aac gtc Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val 965 970 975	2928
atc aag aac gcc gac ttc aac aac gcc ctg ctg tgc tgg aac gtg aag Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys 980 985 990	2976
ggc cac gtg gac gtc gag gag cag aac aac cac aga tca gtc ctg gtg Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val 995 1000 1005	3024
gtg ccc gag tgg gaa gcc gaa gtc tca caa gaa gtc cgg gtg tgc Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys 1010 1015 1020	3069
cct gga cgc ggg tac att ctc cgc gtg acc gcc tac aag gag gcc Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly 1025 1030 1035	3114
tac ggt gag gcc tgc gtg acc atc cac gag atc gag aac aac acc Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr 1040 1045 1050	3159
gac gag ctg aaa ttc agt aac tgt gtt gag gag gag gtg tac ccg	3204

UA 128156 C2

Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1055 1060 1065  
  
 aac aac acc gtc acc tgc aac gac tac act gcg aac cag gag gaa 3249  
 Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu  
 1070 1075 1080  
  
 tac gag ggc acg tac acg agc cgc aat cgc ggg tac gac gag gcg 3294  
 Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
 1085 1090 1095  
  
 tac gag agc aac tcc agc gtc ccg gcc gag tac gcc tcc gtg tac 3339  
 Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr  
 1100 1105 1110  
  
 gag gag aag gtt tac acc gac ggg agg cgt gcc aac ccg tgc gag 3384  
 Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu  
 1115 1120 1125  
  
 agc aac aga ggc tac ggc gat tac act ccg ctt ccc gct ggc tac 3429  
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1130 1135 1140  
  
 gtg acg aaa gag ctg gag tac ttc cca gag acc gac aag gtg tgg 3474  
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1145 1150 1155  
  
 atc gag atc gga gaa acg gag ggc acg ttc ata gtg gac tcc gtt 3519  
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1160 1165 1170  
  
 gag ctg ctg ctc atg gag gag tag 3543  
 Glu Leu Leu Leu Met Glu Glu  
 1175 1180

<210> 8  
 <211> 1180  
 <212> PRT  
 <213> Штучна послідовність  
  
 <220>  
 <223> Синтетична конструкція  
  
 <400> 8

Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn  
 1 5 10 15

Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn  
 20 25 30

Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu  
 35 40 45

UA 128156 C2

Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp  
 50 55 60

Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu  
 65 70 75 80

Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile  
 85 90 95

Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala  
 100 105 110

Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala  
 115 120 125

Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile  
 130 135 140

Ser Val Leu Lys Ile Gln Asn Phe Glu Val Gln Leu Leu Ser Val Phe  
 145 150 155 160

Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe  
 165 170 175

Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr  
 180 185 190

Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg  
 195 200 205

Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp  
 210 215 220

Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu  
 225 230 235 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile  
 245 250 255

Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu  
 260 265 270

UA 128156 C2

Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln  
 275 280 285  
 Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile  
 290 295 300  
 Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile  
 305 310 315 320  
 Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu  
 325 330 335  
 Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr  
 340 345 350  
 Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile  
 355 360 365  
 Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly  
 370 375 380  
 Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile  
 385 390 395 400  
 Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln  
 405 410 415  
 Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His  
 420 425 430  
 Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala  
 435 440 445  
 Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile  
 450 455 460  
 Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile  
 465 470 475 480  
 Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp  
 485 490 495  
 Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr

UA 128156 C2

500					505					510					
Leu	Glu	Val	Pro	Ile	Phe	Phe	Gln	Ser	Pro	Ser	Thr	Asn	Tyr	Arg	Val
		515					520					525			
Arg	Val	Arg	Tyr	Ala	Ser	Thr	Ser	Ser	Leu	Pro	Val	Asp	Val	Val	Phe
	530					535					540				
Gly	Asn	Ile	Ser	His	Pro	Thr	Thr	Phe	Pro	Ala	Thr	Ala	Arg	Ser	Leu
545					550					555					560
Asp	Asn	Leu	Gln	Ser	Asn	Asp	Phe	Gly	Tyr	Ile	Asp	Ile	Ala	Gly	Thr
				565					570					575	
Phe	Leu	Pro	Ser	Leu	Gly	Pro	Ser	Ile	Gly	Ile	Arg	Pro	Met	Leu	Ser
			580						585					590	
Thr	Ile	Asn	Leu	Ile	Val	Asp	Arg	Phe	Glu	Phe	Ile	Pro	Val	Thr	Ala
		595					600						605		
Thr	Phe	Glu	Ala	Glu	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val	Asn
	610					615						620			
Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Ile	Lys	Thr	Asp	Val	Thr
625					630					635					640
Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Glu	Cys	Leu	Ser	Asp
				645					650						655
Glu	Phe	Tyr	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys	His
			660					665						670	
Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn	Phe
		675						680					685		
Arg	Gly	Ile	Asn	Arg	Gln	Pro	Asp	Arg	Gly	Trp	Arg	Gly	Ser	Thr	Asp
	690						695					700			
Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	Thr
705					710					715					720
Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	Lys
				725					730						735

Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly  
740 745 750

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn  
755 760 765

Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro  
770 775 780

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys  
785 790 795 800

Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp  
805 810 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp  
820 825 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe  
835 840 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe  
850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg  
865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr  
885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val  
900 905 910

Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile  
915 920 925

His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro  
930 935 940

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu  
945 950 955 960

UA 128156 C2

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val  
 965 970 975

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys  
 980 985 990

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Va  
 995 1000 1005

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 1010 1015 1020

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1025 1030 1035

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1040 1045 1050

Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1055 1060 1065

Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu  
 1070 1075 1080

Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
 1085 1090 1095

Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr  
 1100 1105 1110

Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu  
 1115 1120 1125

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1130 1135 1140

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1145 1150 1155

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1160 1165 1170

Glu Leu Leu Leu Met Glu Glu  
 1175 1180

<210> 9  
 <211> 3534  
 <212> ДНК  
 <213> Штучна послідовність

<220>  
 <223> Повністю синтетична послідовність, яка кодує СВ 002 для застосування в рослинах

<220>  
 <221> CDS  
 <222> (1)..(3531)  
 <223> BCW 002

<400> 9  
 atg gac aac aac ccg aac atc aac gag tgc atc ccc tac aac tgc ctc 48  
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15  
 tcc aac ccg gag gtc gag gtg ctg ggc ggc gaa agy atc gag acc ggc 96  
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30  
 tac act ccc atc gac atc agc ctc agc ctg acc cag ttc ctg ctc tct 144  
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
 35 40 45  
 gag ttc gtg ccc ggc gcg ggg ttc gtt ctc ggc ctg gtc gac atc atc 192  
 Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
 50 55 60  
 tgg ggc atc ttc ggt ccg agc cag tgg gac gcc ttt ctc gtt cag att 240  
 Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
 65 70 75 80  
 gag cag ctg atc aac cag cgc atc gag gag ttc gcc cgc aac cag gcg 288  
 Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95  
 atc tcc ccg ctg gag ggc ctc tcc aac ctg tac caa atc tac gcc gag 336  
 Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
 100 105 110  
 agc ttc ccg gag tgg gaa gcc gat ccg acc aac ccc gct ctc agy gag 384  
 Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
 115 120 125  
 gag atg ccg att cag ttc aac gac atg aac tcc gct ctc acg act gcc 432  
 Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
 130 135 140  
 atc cca ctc ctc gct gtg cag aac tac caa gtg ccg ctc ctg tcc gtg 480

UA 128156 C2

Ile	Pro	Leu	Leu	Ala	Val	Gln	Asn	Tyr	Gln	Val	Pro	Leu	Leu	Ser	Val	
145					150					155					160	
tac	gtg	cag	gcc	gcc	aat	ctg	cac	ctc	tcc	gtc	ctc	egg	gac	gtt	agc	528
Tyr	Val	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Val	Leu	Arg	Asp	Val	Ser	
				165					170					175		
gtg	ttc	ggg	cag	cgc	tgg	ggc	ttc	gac	gcc	gct	acc	atc	aac	tcc	cgt	576
Val	Phe	Gly	Gln	Arg	Trp	Gly	Phe	Asp	Ala	Ala	Thr	Ile	Asn	Ser	Arg	
			180					185					190			
tac	aac	gat	ctc	act	cgc	ctc	atc	ggc	aac	tac	acc	gac	tat	gcc	gtg	624
Tyr	Asn	Asp	Leu	Thr	Arg	Leu	Ile	Gly	Asn	Tyr	Thr	Asp	Tyr	Ala	Val	
		195					200					205				
cgc	tgg	tac	aac	act	ggt	ctt	gag	aga	gtc	tgg	ggc	ccg	gac	agc	cgc	672
Arg	Trp	Tyr	Asn	Thr	Gly	Leu	Glu	Arg	Val	Trp	Gly	Pro	Asp	Ser	Arg	
	210					215					220					
gac	tgg	gtg	cgc	tac	aac	cag	ttc	cgg	cgc	gag	ctg	acc	ctc	acc	gtg	720
Asp	Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Glu	Leu	Thr	Leu	Thr	Val	
225					230					235					240	
ctc	gac	atc	gta	gcc	ctc	ttt	ccc	aac	tac	gac	tcc	egg	cgc	tac	ccg	768
Leu	Asp	Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Ser	Arg	Arg	Tyr	Pro	
				245					250					255		
att	cgc	atc	gtc	agc	cag	ctc	acc	agg	gag	atc	tac	acc	aac	cct	gtg	816
Ile	Arg	Ile	Val	Ser	Gln	Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Asn	Pro	Val	
			260					265					270			
ctg	gag	aac	ttc	gac	ggc	tcc	ttt	cgc	ggg	atg	gcc	caa	cgc	ata	gag	864
Leu	Glu	Asn	Phe	Asp	Gly	Ser	Phe	Arg	Gly	Met	Ala	Gln	Arg	Ile	Glu	
		275					280					285				
cag	aac	atc	cgc	caa	cct	cat	ctg	atg	gac	atc	ctt	aat	tct	atc	acc	912
Gln	Asn	Ile	Arg	Gln	Pro	His	Leu	Met	Asp	Ile	Leu	Asn	Ser	Ile	Thr	
	290					295					300					
atc	tac	act	gac	gtt	cat	cgc	ggg	ttt	aac	tac	tgg	tcg	ggc	cac	caa	960
Ile	Tyr	Thr	Asp	Val	His	Arg	Gly	Phe	Asn	Tyr	Trp	Ser	Gly	His	Gln	
305					310					315					320	
atc	act	ggg	tcg	ccc	gtt	ggt	ttc	tcc	ggc	ccg	gag	ttc	ggc	ttc	cct	1008
Ile	Thr	Ala	Ser	Pro	Val	Gly	Phe	Ser	Gly	Pro	Glu	Phe	Ala	Phe	Pro	
				325					330					335		
ctg	ttc	gga	aac	ggg	ggc	aat	gcc	gct	cca	ccc	gta	ttg	gtg	agc	ctg	1056
Leu	Phe	Gly	Asn	Ala	Gly	Asn	Ala	Ala	Pro	Pro	Val	Leu	Val	Ser	Leu	
			340					345					350			
acc	ggc	ctc	ggc	atc	ttc	cgt	aca	ctg	tct	agc	cct	ctg	tac	aga	agg	1104
Thr	Gly	Leu	Gly	Ile	Phe	Arg	Thr	Leu	Ser	Ser	Pro	Leu	Tyr	Arg	Arg	
		355					360					365				
atc	att	ctt	ggc	agc	ggt	ccc	aat	aac	cag	gaa	ctc	ttc	gtg	ttg	gac	1152
Ile	Ile	Leu	Gly	Ser	Gly	Pro	Asn	Asn	Gln	Glu	Leu	Phe	Val	Leu	Asp	



UA 128156 C2

gcc act ttc gag gct gag tcc gac cta gag cgt gct cag aag gca gtc 1872  
 Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val  
 610 615 620

aat gct ctg ttt acc tcc acc aat cag ctc ggc att aag acc gat gtg 1920  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val  
 625 630 635 640

acc gat tac cac att gac caa gtc tca aac ctc gtt gag tgc ctc tgg 1968  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 645 650 655

gat gag ttc tac ctt gat gag aag agg gag ctt tca gag aaa gtt aag 2016  
 Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys  
 660 665 670

cac gct aag aga ctc tgg gac gaa cgc aat ctg ttg caa gat ccc aac 2064  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 675 680 685

ttc aga ggg atc aac cgt cag cca gac cgg gga tgg cgc ggg tcc acg 2112  
 Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr  
 690 695 700

gac atc act atc cag gcc ggt gat gac gtc ttc aag gag aac tac gtg 2160  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 705 710 715 720

acc ctg ccg gcc acc ttt gac gaa tgc tac ccc act tac ctc tac cag 2208  
 Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 725 730 735

aag att gac gag tcc aag ctc aag gcg ttc aca cgc tac cag ctc agg 2256  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg  
 740 745 750

ggt tac atc gag gac tcc caa gac ctg gaa atc tac ctg atc cgc tac 2304  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 755 760 765

aac gct aag cac gag act gtc aac gtg ccc gcc acc gcc agc ctg tgg 2352  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 770 775 780

ccc ttg tcc gct cag agc cca atc gcc aag tgc gcc gag ccc aac cgc 2400  
 Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg  
 785 790 795 800

tgc gcg ccc cac ctg gaa tgg aac ccc gac ctc gac tgt agc tgc cgc 2448  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg  
 805 810 815

gac gga gag aag tgc gcg cat cac tcc cac cac ttc agc ctc gac atc 2496  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 820 825 830

UA 128156 C2

gac gtc ggt tgc acc gac ctt aac gag gat ctg ggc gtt tgg gtg atc	2544
Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile	
835 840 845	
ttc aag atc aag act cag gac ggc cac gcc cgc ctg gga aac ctg gag	2592
Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu	
850 855 860	
ttc ctg gag gag aag ccc ctg gtt ggc gag gcc ctg gcc cgc gtc aag	2640
Phe Leu Gln Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys	
865 870 875 880	
agg gcc gag aag aaa tgg cgc gac aag cgc gag aag ctg gag tgg gag	2688
Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu	
885 890 895	
acc aac atc gtg tac aag gaa ggc aag gag tca gtt gac gcc ctg ttc	2736
Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe	
900 905 910	
gtc aac agc cag tac gac cag ctg cag gca gac aca aac atc gct atg	2784
Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met	
915 920 925	
atc cat ggc gcc gac aag cgc gtc cac tcc atc cgc gag gcg tac ctg	2832
Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu	
930 935 940	
ccc gag ctg tcc gtc atc ccc ggc gtc aac gcc gcg atc ttt gag gag	2880
Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu	
945 950 955 960	
ctg gag ggc cgc atc ttc acc gcc ttc tcc ctg tac gac gca cgc aac	2928
Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn	
965 970 975	
gtt atc aag aat ggc gac ttc aac aac ggg ctg tcc tgc tgg aat gtc	2976
Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val	
980 985 990	
aag gcc cac gtg gac gtc gag gag cag aac aac cag cgc tca gtc ctg	3024
Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu	
995 1000 1005	
gtc gtc cag gag tgg gag gcc gaa gtc agc cag gaa gtc cgc gtc	3069
Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val	
1010 1015 1020	
tgc cct gga cgc ggg tac atc ctg cgc gtc act gcc tac aag gaa	3114
Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu	
1025 1030 1035	
ggc tac gga gag gcc tgc gtc acc atc cat gag atc gaa aac aac	3159
Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn	
1040 1045 1050	
acg gat gag ctt aag ttc agc aac tgt gtt gaa gag gaa atc tac	3204

UA 128156 C2

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065  
 ccg aac aac acg gtc acc tgc aat gat tac acc gtc aac cag gag 3249  
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu  
 1070 1075 1080  
 gaa tac ggt gga gct tac acc toc cgc aac agg ggc tac aac gag 3294  
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095  
 gca ccc tct gtc ccg gcc gac tac gct tca gtc tac gaa gag aag 3339  
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110  
 tcg tac acc gac gga cgc aga gag aac ccg tgt gag ttc aac cgc 3384  
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125  
 ggc tac cgc gat tac acc ccg ctg cct gtc ggg tac gtc acc aaa 3429  
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140  
 gag ctg gaa tac ttc cca gag acc gac aaa gtc tgg att gag atc 3474  
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155  
 ggc gag acc gag gcc acg ttc atc gty gac toc gtc gaa ctc att 3519  
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170  
 ctg atg gaa gag tga 3534  
 Leu Met Glu Glu  
 1175

<210> 10  
 <211> 1177  
 <212> PRT  
 <213> Штучна послідовність  
 <220>  
 <223> Синтетична конструкція  
 <400> 10

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15  
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30  
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
 35 40 45

UA 128156 C2

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
 130 135 140

Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val  
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser  
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg  
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val  
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg  
 210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val  
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro  
 245 250 255

Ile Arg Ile Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val  
 260 265 270

UA 128156 C2

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu  
 275 280 285  
 Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr  
 290 300  
 Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln  
 305 310 315 320  
 Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro  
 325 330 335  
 Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu  
 340 345 350  
 Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg  
 355 360 365  
 Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Gln Leu Phe Val Leu Asp  
 370 375 380  
 Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr  
 385 390 395 400  
 Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro  
 405 410 415  
 Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser  
 420 425 430  
 His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg  
 435 440 445  
 Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile  
 450 455 460  
 Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser  
 465 470 475 480  
 Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly  
 485 490 495  
 Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala



UA 128156 C2

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg  
740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg  
785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg  
805 810 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
820 825 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
835 840 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
850 855 860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys  
865 870 875 880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu  
885 890 895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met  
915 920 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu  
930 935 940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
945 950 955 960

UA 128156 C2

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 965 970 975  
 Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val  
 980 985 990  
 Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Le  
 995 1000 1005  
 Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 1010 1015 1020  
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1025 1030 1035  
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1040 1045 1050  
 Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065  
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu  
 1070 1075 1080  
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095  
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110  
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125  
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140  
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155  
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170

Leu Met Glu Glu  
1175

<210> 11  
<211> 3534  
<212> ДНК  
<213> Штучна послідовність

<220>  
<223> Повністю синтетична нуклеотидна послідовність, яка кодує BCW 003 для застосування в рослинах

<220>  
<221> CDS  
<222> (1)..(3531)  
<223> BCW 003

<400> 11  
atg gac aac aac ccg aac atc aac gag tgc atc ccc tac aac tgc ctc 48  
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
1 5 10 15  
tcc aac ccg gag gtc gag gtg ctg ggc ggc gaa agg atc gag acc ggc 96  
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
20 25 30  
tac act ccc atc gac atc agc ctc agc ctg acc cag ttc ctg ctc tct 144  
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
35 40 45  
gag ttc gtg ccc ggc gcg ggg ttc gtt ctc ggc ctg gtc gac atc atc 192  
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
50 55 60  
tgg ggc atc ttc ggt ccg agc cag tgg gac gcc ttt ctc gtt cag att 240  
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
65 70 75 80  
gag cag ctg atc aac cag cgc atc gag gag ttc gcc cgc aac cag gcg 288  
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
85 90 95  
atc tcc ccg ctg gag ggc ctc tcc aac ctg tac caa atc tac gcc gag 336  
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu  
100 105 110  
agc ttc ccg gag tgg gaa gcc gat ccg acc aac ccc gct ctc agg gag 384  
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu  
115 120 125  
gag atg ccg att cag ttc aac gac atg aac tcc gct ctc acg act gcc 432  
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
130 135 140  
atc cca ctc ctc gct gtg cag aac tac caa gtg ccg ctc ctg tcc gtg 480

Ile	Pro	Leu	Leu	Ala	Val	Gln	Asn	Tyr	Gln	Val	Pro	Leu	Leu	Ser	Val	
145					150					155					160	
tac	gtg	cag	gcc	gcc	aat	ctg	cac	ctc	tcc	gtc	ctc	cgg	gac	gtt	agc	528
Tyr	Val	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Val	Leu	Arg	Asp	Val	Ser	
			165					170						175		
gtg	ttc	ggg	cag	cgc	tgg	ggc	ttc	gac	gcc	gct	acc	atc	aac	tcc	cgt	576
Val	Phe	Gly	Gln	Arg	Trp	Gly	Phe	Asp	Ala	Ala	Thr	Ile	Asn	Ser	Arg	
			180					185					190			
tac	aac	gat	ctc	act	cgc	ctc	atc	ggc	aac	tac	acc	gac	tat	gcc	gtg	624
Tyr	Asn	Asp	Leu	Thr	Arg	Leu	Ile	Gly	Asn	Tyr	Thr	Asp	Tyr	Ala	Val	
		195					200					205				
cgc	tgg	tac	aac	act	ggg	ctt	gag	aga	gtc	tgg	ggc	cgg	gac	agc	cgc	672
Arg	Trp	Tyr	Asn	Thr	Gly	Leu	Glu	Arg	Val	Trp	Gly	Pro	Asp	Ser	Arg	
	210					215					220					
gac	tgg	gtg	cgc	tac	aac	cag	ttc	cgg	cgc	gag	ctg	acc	ctc	acc	gtg	720
Asp	Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Glu	Leu	Thr	Leu	Thr	Val	
					230					235					240	
ctc	gac	atc	gta	gcc	ctc	ttt	ccc	aac	tac	gac	tcc	cgg	cgc	tac	cgg	768
Leu	Asp	Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Ser	Arg	Arg	Tyr	Pro	
				245					250					255		
att	cgc	acc	gtc	agc	cag	ctc	acc	agg	gag	atc	tac	acc	aac	cct	gtg	816
Ile	Arg	Thr	Val	Ser	Gln	Leu	Thr	Arg	Gln	Ile	Tyr	Thr	Asn	Pro	Val	
			260					265					270			
ctg	gag	aac	ttc	gac	ggc	tcc	ttt	cgc	ggg	atg	gcc	caa	cgc	ata	gag	864
Leu	Glu	Asn	Phe	Asp	Gly	Ser	Phe	Arg	Gly	Met	Ala	Gln	Arg	Ile	Glu	
		275					280					285				
cag	aac	atc	cgc	caa	cct	cat	ctg	atg	gac	atc	ctt	aat	tct	atc	acc	912
Gln	Asn	Ile	Arg	Gln	Pro	His	Leu	Met	Asp	Ile	Leu	Asn	Ser	Ile	Thr	
		290				295					300					
atc	tac	act	gac	gtt	cat	cgc	ggg	ttt	aac	tac	tgg	tgg	ggc	cac	caa	960
Ile	Tyr	Thr	Asp	Val	His	Arg	Gly	Phe	Asn	Tyr	Trp	Ser	Gly	His	Gln	
				305		310				315					320	
atc	act	ggg	tgg	ccc	gtt	ggg	ttc	tcc	ggc	ccg	gag	ttc	ggg	ttc	cct	1008
Ile	Thr	Ala	Ser	Pro	Val	Gly	Phe	Ser	Gly	Pro	Glu	Phe	Ala	Phe	Pro	
				325					330					335		
ctg	ttc	gga	aac	ggg	ggc	aat	gac	gct	cca	ccc	gta	ttg	gtg	agc	ctg	1056
Leu	Phe	Gly	Asn	Ala	Gly	Asn	Ala	Ala	Pro	Pro	Val	Leu	Val	Ser	Leu	
			340					345					350			
acc	ggc	ctc	ggc	atc	ttc	cgt	aca	ctg	tct	agc	cct	ctg	tac	aga	agg	1104
Thr	Gly	Leu	Gly	Ile	Phe	Arg	Thr	Leu	Ser	Ser	Pro	Leu	Tyr	Arg	Arg	
			355				360					365				
atc	att	ctt	ggc	agc	ggg	ccc	aat	aac	cag	gaa	ctc	ttc	gtg	ttg	gac	1152
Ile	Ile	Leu	Gly	Ser	Gly	Pro	Asn	Asn	Gln	Glu	Leu	Phe	Val	Leu	Asp	

UA 128156 C2

370	ggc acc gag ttc agc ttc gcc agt ctt acg acc aat ttg ccc toc aca	375	ggc agt ctt acg acc aat ttg ccc toc aca	380		1200
	Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr					
	385		390		395	400
	atc tat cgc cag cgc ggt act gtg gac tcc ctt gat gtg ata cca cct					1248
	Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro					
		405		410		415
	cag gac aac tct gtc cca cct cgc gcc ggt ttc tcc cac cgc ctc agc					1296
	Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser					
		420		425		430
	cac gtc act atg ctg agt cag get gcg gga gcc gtg tac acc ctt cgg					1344
	His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg					
		435		440		445
	gct ccg acg ttt agc tgg cag cac agg agc gcg act acc acg aac atc					1392
	Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile					
		450		455		460
	att gcg gct gac tcc atc act caa atc cct gcc gtt aag ggt cgc tcc					1440
	Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser					
		465		470		475
	atc atc aac aat ggg aca gtg atc tcg gga ccg gcc ttc acc gcc ggt					1488
	Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly					
		485		490		495
	gac ctg gtg agg ctg tac aac gcg gac ttc aac atc aac aac agg gcg					1536
	Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala					
		500		505		510
	tac ctc gaa gtc ccg atc ttc ttc cag tcg ccc agc acg aac tat cgt					1584
	Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg					
		515		520		525
	gtc agg gtc cgg tac gcc tca acc tca tcc ctc ccg gtc gat gtg gtc					1632
	Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val					
		530		535		540
	ttc gcc aac atc agc cac ccg acc acg ttt ccg gct acc gcc cga tcc					1680
	Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser					
		545		550		555
	ctg gac aat ctg caa agc aac gat ttc gcc tac att gac att gcc ggg					1728
	Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly					
		565		570		575
	acg ttc ctc ccg agc ctc gcc cca tcc atc gcc atc cgg ccc atg ctc					1776
	Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu					
		580		585		590
	tcc acc atc aac ctg atc gtg gat cgg ttt gag ttc atc cca gtg aca					1824
	Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr					
		595		600		605

UA 128156 C2

gcc	act	ttc	gag	gct	gag	tcc	gac	cta	gag	cgt	gct	cag	aag	gca	gtc	1872
Ala	Thr	Phe	Glu	Ala	Gln	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val	
	610					615					620					
aat	gct	ctg	ttt	acc	tcc	acc	aat	cag	ctc	ggc	att	aag	acc	gat	gtg	1920
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Ile	Lys	Thr	Asp	Val	
	625				630					635					640	
acc	gat	tac	cac	att	gac	caa	gtc	tca	aac	ctc	gtt	gag	tgc	ctc	tgg	1968
Thr	Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Glu	Cys	Leu	Ser	
				645					650					655		
gat	gag	ttc	tac	ctt	gat	gag	aag	agg	gag	ctt	tca	gag	aaa	gtt	aag	2016
Asp	Glu	Phe	Tyr	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys	
		660						665					670			
cac	gct	aag	aga	ctc	tgg	gac	gaa	cgc	aat	ctg	ttg	caa	gat	ccc	aac	2064
His	Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn	
		675					680					685				
ttc	aga	ggg	atc	aac	cgt	cag	cca	gac	cgg	gga	tgg	cgc	ggg	tcc	acg	2112
Phe	Arg	Gly	Ile	Asn	Arg	Gln	Pro	Asp	Arg	Gly	Trp	Arg	Gly	Ser	Thr	
	690					695					700					
gac	atc	act	atc	cag	ggc	ggt	gat	gac	gtc	ttc	aag	gag	aac	tac	gtg	2160
Asp	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	
	705				710					715					720	
acc	ctg	ccg	ggc	acc	ttt	gac	gaa	tgc	tac	ccc	act	tac	ctc	tac	cag	2208
Thr	Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	
				725					730					735		
aag	att	gac	gag	tcc	aag	ctc	aag	gog	ttc	aca	cgc	tac	cag	ctc	agg	2256
Lys	Ile	Asp	Glu	Ser	Lys	Leu	Lys	Ala	Phe	Thr	Arg	Tyr	Gln	Leu	Arg	
		740						745					750			
ggt	tac	atc	gag	gac	tcc	caa	gac	ctg	gaa	atc	tac	ctg	atc	cgc	tac	2304
Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Leu	Ile	Arg	Tyr	
		755					760						765			
aac	gct	aag	cac	gag	act	gtc	aac	gtg	ccc	ggc	acc	ggc	agc	ctg	tgg	2352
Asn	Ala	Lys	His	Glu	Thr	Val	Asn	Val	Pro	Gly	Thr	Gly	Ser	Leu	Trp	
	770					775					780					
ccc	ttg	tcc	gct	cag	agc	cca	atc	ggc	aag	tgc	ggc	gag	ccc	aac	cgc	2400
Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly	Glu	Pro	Asn	Arg	
	785				790					795				800		
tgc	ggg	ccc	cac	ctg	gaa	tgg	aac	ccc	gac	ctc	gac	tgt	agc	tgc	cgc	2448
Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg	
				805					810					815		
gac	gga	gag	aag	tgc	gcg	cat	cac	tcc	cac	cac	ttc	agc	ctc	gac	atc	2496
Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	
			820					825						830		

UA 128156 C2

gac gtc ggt tgc acc gac ott aac gag gat ctg ggc gtt tgg gtg atc	2544
Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile	
835 840 845	
ttc aag atc aag act cag gac ggc cac gcc cgc ctg gga aac ctg gag	2592
Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu	
850 855 860	
ttc ctg gag gag aag ccc ctc gtt ggc gag gcc ctg gcc cgc gtc aag	2640
Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys	
865 870 875 880	
agg gcc gag aag aaa tgg cgc gac aag cgc gag aag ctg gag tgg gag	2688
Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu	
885 890 895	
acc aac atc gtg tac aag gaa gcg aag gag tca gtt gac gcc ctg ttc	2736
Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe	
900 905 910	
gtc aac agc cag tac gac cag ctc cag gca gac aca aac atc gct atg	2784
Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met	
915 920 925	
atc cat gcg gcc gac aag cgc gtc cac tcc atc cgc gag gcg tac ctg	2832
Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu	
930 935 940	
ccc gag ctg tcc gtc atc ccc gcc gtc aac gcc gcg atc ttt gag gag	2880
Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu	
945 950 955 960	
ctg gag gcc cgc atc ttc acc gcc ttc tcc ctc tac gac gca cgc aac	2928
Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn	
965 970 975	
gtt atc aag aat gcc gac ttc aac aac ggg ctg tcc tgc tgg aat gtc	2976
Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val	
980 985 990	
aag gcc cac gtg gac gtc gag gag cag aac aac cag cgc tca gtc ctg	3024
Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu	
995 1000 1005	
gtc gtc ccg gag tgg gag gcc gaa gtc agc cag gaa gtc cgc gtc	3069
Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val	
1010 1015 1020	
tgc cct gga cgc ggg tac atc ctg cgc gtc act gcc tac aag gaa	3114
Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu	
1025 1030 1035	
ggc tac gga gag gcc tgc gtc acc atc cat gag atc gaa aac aac	3159
Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn	
1040 1045 1050	
acg gat gag ctt aag ttc agc aac tgt gtt gaa gag gaa atc tac	3204

UA 128156 C2

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065

ccg aac aac acg gtc acc tgc aat gat tac acc gtc aac cag gag 3249  
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu  
 1070 1075 1080

gaa tac ggt gga gct tac acc tcc cgc aac agg ggc tac aac gag 3294  
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095

gca ccc tct gtc ccg gcc gac tac gct tca gtc tac gaa gag aag 3339  
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110

tgg tac acc gac gga cgc aga gag aac ccg tgt gag ttc aac cgc 3384  
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125

ggc tac cgc gat tac acc ccg ctg cct gtc ggg tac gtc acc aaa 3429  
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140

gag ctg gaa tac ttc cca gag acc gac aaa gtc tgg att gag atc 3474  
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155

ggc gag acc gag ggc acg ttc atc gtg gac tcc gtc gaa ctc ctt 3519  
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170

ctg atg gaa gag tga 3534  
 Leu Met Glu Glu  
 1175

<210> 12  
 <211> 1177  
 <212> PRT  
 <213> Штучна послідовність

<220>  
 <225> Синтетична конструкція

<400> 12

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu  
 1 5 10 15

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly  
 20 25 30

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser  
 35 40 45

UA 128156 C2

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile  
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile  
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Gln  
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Gln  
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala  
 130 135 140

Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val  
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser  
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg  
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val  
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg  
 210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val  
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro  
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val  
 260 265 270

UA 128156 C2

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu  
 275 280 285  
 Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr  
 290 295 300  
 Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln  
 305 310 315 320  
 Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro  
 325 330 335  
 Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu  
 340 345 350  
 Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg  
 355 360 365  
 Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp  
 370 375 380  
 Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr  
 385 390 395 400  
 Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro  
 405 410 415  
 Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser  
 420 425 430  
 His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg  
 435 440 445  
 Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile  
 450 455 460  
 Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser  
 465 470 475 480  
 Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly  
 485 490 495  
 Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala

UA 128156 C2

			500					505						510		
Tyr	Leu	Glu	Val	Pro	Ile	Phe	Phe	Gln	Ser	Pro	Ser	Thr	Asn	Tyr	Arg	
		515						520					525			
Val	Arg	Val	Arg	Tyr	Ala	Ser	Thr	Ser	Ser	Leu	Pro	Val	Asp	Val	Val	
	530					535					540					
Phe	Gly	Asn	Ile	Ser	His	Pro	Thr	Thr	Phe	Pro	Ala	Thr	Ala	Arg	Ser	
545					550					555					560	
Leu	Asp	Asn	Leu	Gln	Ser	Asn	Asp	Phe	Gly	Tyr	Ile	Asp	Ile	Ala	Gly	
				565					570					575		
Thr	Phe	Leu	Pro	Ser	Leu	Gly	Pro	Ser	Ile	Gly	Ile	Arg	Pro	Met	Leu	
			580					585					590			
Ser	Thr	Ile	Asn	Leu	Ile	Val	Asp	Arg	Phe	Glu	Phe	Ile	Pro	Val	Thr	
		595					600					605				
Ala	Thr	Phe	Glu	Ala	Glu	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val	
	610					615					620					
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Ile	Lys	Thr	Asp	Val	
625					630					635					640	
Thr	Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Glu	Cys	Leu	Ser	
			645						650					655		
Asp	Glu	Phe	Tyr	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys	
		660						665					670			
His	Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn	
		675					680					685				
Phe	Arg	Gly	Ile	Asn	Arg	Gln	Pro	Asp	Arg	Gly	Trp	Arg	Gly	Ser	Thr	
690						695					700					
Asp	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	
705					710					715					720	
Thr	Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	
				725					730					735		

UA 128156 C2

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg  
 740 745 750  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 755 760 765  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 770 775 780  
 Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg  
 785 790 795 800  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg  
 805 810 815  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 820 825 830  
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
 835 840 845  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
 850 855 860  
 Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys  
 865 870 875 880  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu  
 885 890 895  
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 900 905 910  
 Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met  
 915 920 925  
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu  
 930 935 940  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
 945 950 955 960

UA 128156 C2

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val  
 980 985 990

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu  
 995 1000 1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 1010 1015 1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1025 1030 1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1040 1045 1050

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr  
 1055 1060 1065

Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu  
 1070 1075 1080

Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu  
 1085 1090 1095

Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys  
 1100 1105 1110

Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg  
 1115 1120 1125

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys  
 1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile  
 1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu  
 1160 1165 1170

Leu Met Glu Glu  
 1175

---

Комп'ютерна верстка Л. Бурлак

---

ДО "Український національний офіс інтелектуальної власності та інновацій", вул. Дмитра Годзенка, 1, м. Київ – 42, 01601