



УКРАЇНА

(19) UA (11) 92622 (13) C2

(51) МПК (2009)

C12N 15/54

C12N 15/82

C12N 9/10

A01H 5/00

C12N 5/10

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІОПИС
ДО ПАТЕНТУ НА ВИНАХІД

(54) ГЕНИ GRG23 ТА GRG51, ЯКІ НАДАЮТЬ СТІЙКІСТЬ ДО ГЕРБІЦИДІВ

1

(21) a200808448

(22) 01.12.2006

(24) 25.11.2010

(86) PCT/US2006/045908, 01.12.2006

(31) 60/741,166

(32) 01.12.2005

(33) US

(31) 60/817,799

(32) 30.06.2006

(33) US

(46) 25.11.2010, Бюл.№ 22, 2010 р.

(72) ПЕТЕРС ЧЕРІЛ Л., US, ХАЙНСОН ДЖІЛЛ, US,
ХАММЕР ФІЛІП Е., US, ВАНДЕ БЕРГ БРАЙАН, US,
СХОУТЕН ЛАУРА КУПЕР, US, КАРП БРАЙАН, US

(73) АТЕНІКС КОРПОРЕЙШН, US

(56) PIPKE R ET AL: "DEGRADATION OF THE
PHOSPHONATE HERBICIDE GLYPHOSATE BY
ARTHROBACTER-ATROCZYANEUS ATCC 13752"
APPLIED AND ENVIRONMENTAL MICROBIOLOGY,
vol. 54, no. 5, 1988, pages 1293-1296PIPKE R ET AL: "ISOLATION AND
CHARACTERIZATION OF A MUTANT OF
ARTHROBACTER-SP STRAIN GLP-1 WHICH
UTILIZES THE HERBICIDE GLYPHOSATE AS ITS
SOLE SOURCE OF PHOSPHORUS AND
NITROGEN" APPLIED AND ENVIRONMENTAL
MICROBIOLOGY, vol. 54, no. 11, 1988, pages 2868-
2870,PADGETTE S R ET AL: "SITE-DIRECTED
MUTAGENESIS OF A CONSERVED REGION OF
THE 5-ENOLPYRUVYL-SHIKIMATE-3-PHOSPHATE
SYNTHASE ACTIVE SITE" JOURNAL OF
BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY
OF BIOLOGICAL CHEMISTS, BIRMINGHAM,,
US, vol. 266, no. 33, 25 November 1991 (1991-11-
25), pages 22364-22369(57) 1. Ізольована молекула нуклеїнової кислоти,
вибрана із групи, яка складається з:

а) молекули нуклеїнової кислоти, яка містить нуклеотидну послідовність SEQ ID NO: 1, 3 або 5, або комплементарну до такої;

б) молекули нуклеїнової кислоти, яка містить нуклеотидну послідовність, яка має принаймні 80 %

2

ідентичність послідовності з нуклеотидною послідовністю SEQ ID NO: 1, 3 або 5, або комплементарну до такої;

с) нуклеотидної послідовності стійкості до гербіциду вставки ДНК плазмиди, депонованої під Номером Доступу NRRL B-30888 або B-30949, або комплементарної до такої;

d) молекули нуклеїнової кислоти, яка кодує поліпептид, що містить амінокислотну послідовність SEQ ID NO: 2, 4 або 6; і

е) молекули нуклеїнової кислоти, яка містить нуклеотидну послідовність, яка кодує поліпептид, яка має принаймні 80 % ідентичність амінокислотної послідовності з амінокислотною послідовністю SEQ ID NO: 2, 4 або 6.

2. Ізольована молекула нуклеїнової кислоти за п. 1, де вказана нуклеотидна послідовність є синтетичною послідовністю, що була сконструйована для експресії в рослині.

3. Клітина-хазяїн, яка містить вектор, що містить молекулу нуклеїнової кислоти за п. 1.

4. Клітина-хазяїн за п. 3, яка є бактеріальною клітиною-хазяїном.

5. Клітина-хазяїн за п. 3, яка є рослинною клітиною.

6. Трансформоване насіння рослини, яка містить клітину-хазяїна за п. 5.

7. Ізольований поліпептид, вибраний із групи, яка складається з:

а) поліпептиду, що містить амінокислотну послідовність SEQ ID NO: 2, 4 або 6;

б) поліпептиду, кодованого нуклеотидною послідовністю SEQ ID NO: 1, 3 або 5;

с) поліпептиду, що містить амінокислотну послідовність, яка має принаймні 80 % ідентичність послідовності з амінокислотною послідовністю SEQ ID NO: 2, 4 або 6, де вказаний поліпептид має активність стійкості до гербіциду;

d) поліпептиду, що кодується нуклеотидною послідовністю, що принаймні на 80 % ідентична до нуклеотидної послідовності SEQ ID NO: 1, 3 або 5, де вказаний поліпептид має активність стійкості до гербіциду; і,

(13) C2

(11) 92622

(19) UA

е) поліпептиду, що кодується нуклеотидною послідовністю стійкості до гербіциду вставки ДНК плазмиди, депонованою під Номером Доступу NRRL B-30888 або NRRL B-30949.

8. Поліпептид за п. 7, який додатково містить гетерологічну амінокислотну послідовність.

9. Рослина, що має стабільно вбудовану в її геном ДНК-конструкцію, що містить нуклеотидну послідовність, яка кодує білок, що має активність стійкості до гербіциду, де вказана нуклеотидна послідовність вибрана із групи, яка складається з:

а) нуклеотидної послідовності SEQ ID NO: 1, 3 або 5;

б) нуклеотидної послідовності, що має принаймні 80 % ідентичність послідовності з нуклеотидною послідовністю SEQ ID NO: 1, 3 або 5, де вказана нуклеотидна послідовність кодує поліпептид, який має активність стійкості до гербіциду;

с) нуклеотидної послідовності, яка кодує поліпептид, що містить амінокислотну послідовність SEQ ID NO: 2, 4 або 6;

д) нуклеотидної послідовності, яка кодує поліпептид, що має принаймні 80 % ідентичність амінокислотної послідовності з амінокислотною послідовністю SEQ ID NO: 2, 4 або 6, де вказаний поліпептид має активність стійкості до гербіциду; і,

є) нуклеотидної послідовності стійкості до гербіциду вставки ДНК плазмиди, депонованої під Номером Доступу NRRL B-30888 або NRRL B-30949;

де вказана нуклеотидна послідовність є функціонально зв'язаною із промотором, що керує експресією кодувальної послідовності, у рослинній клітині.

10. Рослина за п. 9, де вказаною рослиною є рослинна клітина.

11. Рослина, що має стабільно вбудовану в її геном ДНК-конструкцію, що містить нуклеотидну послідовність, яка кодує поліпептид EPSPS, де вказаний поліпептид EPSPS має K_m у відношенні PEP від приблизно 1 до приблизно 150 мк, і K_i (гліфосат)/ K_m (PEP) від приблизно 500 до приблизно 1000, де вказана рослина проявляє стійкість до гліфосатного гербіциду.

12. Рослина за п. 11, де вказаною рослиною є рослина сої.

13. Рослина за п. 11, де вказаною рослиною є зернова рослина.

14. Рослина за п. 11, де вказана рослина вибрана із групи, яка складається з кукурудзи, сорго звичайного, пшениці, соняшника, томату, хрестоцвітних, перців, картоплі, бавовни, рису, сої, цукрового буряка, цукрового очерету, тютюну, ячменю й олійних культур.

Галузь техніки

Даний винахід розкриває нові гени, що кодують стійкість до гербіцидів, які застосовні в біології рослин, при селекції сільськогосподарських культур і для культур рослинних клітин.

Попередній рівень техніки

N-фосфометилглутамин, який звичайно називають гліфосатом, є важливим агрономічним хімікатом. Гліфосат інгібує фермент, що перетворює фосфоенолпіровиноградну кислоту (PEP) і 3-фосфошикімову кислоту (S3P) в 5-енолпірувіл-3-фосфошикімову кислоту. Інгібування цього ферменту (5-енолпірувілшикімат-3-фосфатсинтаза; яка позначається в даному описі як "EPSPS"), знищує рослинні клітини за допомогою відключення шляху метаболізму шікімату, пригнічуючи в такий спосіб біосинтез ароматичних кислот.

Оскільки гербіциди класу гліфосату пригнічують біосинтез ароматичних амінокислот, вони не тільки знищують рослинні клітини, але і є токсичними для бактеріальних клітин. Гліфосат пригнічує багато бактеріальних синтаз EPSP, і в такий спосіб є токсичним для цих бактерій. Однак певні бактеріальні синтази EPSP мають високу стійкість до гліфосату.

Рослинні клітини, стійкі до токсичності гліфосату, можуть бути отримані за допомогою трансформування рослинних клітин для експресії гліфосат-стійких бактеріальних EPSP-синтаз. Зокрема, бактеріальний ген зі штаму CP4 *Agrobacterium tumefaciens* був використаний для надання рослинним клітинам стійкості до гербіцидів після експресії в рослинах. Мутована EPSP-синтаза зі штаму CT7 *Salmonella typhimurium* надає стійкість до гліфосату бактеріальним клітинам і надає стійкість

до гліфосату рослинним клітинам (Патенти США №№ 4535060; 4769061 і 5094945). Однак існує потреба в інших генах стійкості до гербіцидів.

Кінетична активність EPSPS може бути оцінена за допомогою вимірювання вивільнення фосфату. Вивільнення фосфату визначається з використанням спряженого аналізу для флуоресцентного виявлення фосфату на основі утворення N-ацетил-3,7-дигідроксифеноксацину (Amplex® Red), як відомо в галузі техніки (Vazquez et al. (2003) *Analytical Biochemistry* 320: 292-298). Опубліковані умови аналізу можуть привести до насичення аналізу в експериментах, де фосфат вивільняється дуже швидко. Для вимірювання кінетичної активності EPSPS необхідні додаткові способи.

Суть винаходу

Представлено композиції й способи надання стійкості або толерантності до гербіциду бактеріям, рослинам, рослинним клітинам, тканинам і насінням. Композиції включають молекули нуклеїнових кислот, які кодують поліпептиди стійкості або толерантності до гербіциду, вектори, які містять такі молекули нуклеїнових кислот, і клітини-хазяї, які містять вектори. Композиції також включають антитіла до поліпептидів стійкості або толерантності до гербіциду. Як відзначено, нуклеотидні послідовності за винаходом можуть бути використані в конструкціях ДНК або експресійних касетах для трансформації й експресії в організмах, що включають мікроорганізми й рослини. Композиції також містять трансформовані бактерії, рослини, рослинні клітини, тканини й насіння. Крім того, представлені способи одержання поліпептидів, які

кодуються синтетичними нуклеотидами за винаходом.

Представлено ізольовані молекули нуклеїнових кислот і варіанти таких, які кодують поліпептиди стійкості або толерантності до гербіциду. Додатково, охоплені послідовності амінокислоти й варіанти таких, що кодуються полінуклеотидами, які надають стійкість або толерантність до гербіциду. Даний винахід представляє молекули ізольованих нуклеїнових кислот, що містять нуклеотидну послідовність, представлену в SEQ ID NO:1, 3 або 5, нуклеотидну послідовність, яка кодує амінокислотну послідовність, представлену в SEQ ID NO:2, 4 або 6, нуклеотидну послідовність стійкості до гербіциду, депоновану в бактерії-хазяїні у вигляді номерів доступу NRRL B-30888 або NRRL B-30949, так само, як і варіанти й фрагменти таких. Нуклеотидні послідовності, які комплементарні до нуклеотидних послідовностей за винаходом, або ті, які гібридизуються з послідовностями за винаходом, також охоплені.

Також представлені способи вимірювання ферментативної кінетичної активності із застосуванням флуорогенних субстратів.

Опис фігур креслень

На Фіг.1 показане вирівнювання GRG23 ORF1 амінокислотної послідовності (SEQ ID NO:2) і GRG51 (SEQ ID NO:6) з *Bacillus clausii* (SEQ ID NO:7), *Rubrobacter xylanophilus* (SEQ ID NO:8), *Escherichia coli* (SEQ ID NO:11), штамом CP4 *Agrobacterium* sp. (SEQ ID NO:10) і *Zea mays* (SEQ ID NO:9).

На Фіг.2 показаний графік розсіювання ферментативної активності GRG23 (вісь Y) як функції концентрації PEP (вісь X) при концентраціях гліфосату 0, 3, 5 і 10 mM.

На Фіг.3 показаний графік розсіювання K_m (app.) (вісь Y) як функції концентрації гліфосату (вісь X). Точка перетину осі X представляє K_i для гліфосату.

Докладний опис

Дані винаходи тепер будуть описані більш докладно з посиланням на додані креслення, у яких представлені деякі, але не всі, варіанти здійснення винаходів. Дійсно, ці винаходи можуть бути втілені в багатьох різних формах і не повинні бути розглянуті як обмежені варіантами здійснення, представленими тут; скоріше ці способи здійснення представлені так, щоб це розкриття задовольнило застосовні законні вимоги. Подібні номери у всій заявці стосуються подібних елементів.

Багато модифікацій і інших варіантів здійснення винаходів, представлених тут, будуть зрозумілі фахівцям в галузі техніки, до якої ці винаходи належать, що мають перевагу ідеї винаходу, представленого в попередніх описах і супутніх кресленнях. Тому, слід розуміти, що винаходи не повинні бути обмежені конкретними розкритими варіантами здійснення, і що модифікації й інших варіантів здійснення призначені для включення в обсяг домагань. Хоча тут використовуються певні терміни, вони застосовані тільки в типовому й описовому змісті, а не з метою обмеження.

Даний винахід стосується композицій і способів регуляції в організмах стійкості до гербіциду, особливо в рослинах або рослинних клітинах.

Способи включають трансформацію організмів нуклеотидними послідовностями, що кодують ген стійкості до гліфосату за винаходом. Нуклеотидні послідовності за винаходом придатні для одержання рослин, які проявляють підвищену толерантність до гербіцидного гліфосату. Таким чином, представлені трансформовані бактерії, рослини, рослинні клітини, рослинні тканини й насіння. Композиції включають нуклеїнові кислоти й білки, що мають відношення до толерантності до гербіциду в мікроорганізмах і рослинах, так само як і трансформовані бактерії, рослини, рослинні тканини й насіння. Розкриті нуклеотидні послідовності гена стійкості до гліфосату (grg23 і grg51) і амінокислотні послідовності білків, які кодуються ними. Послідовності знаходять застосування в конструюванні експресійних векторів для наступної трансформації рослин, які представляють інтерес, як зони для виділення інших генів стійкості до гліфосату, як маркери селекції, і т.п. Таким чином, під "геном стійкості до гліфосату за винаходом" мається на увазі нуклеотидна послідовність, наведена в SEQ ID NO:1 або 3, і варіанти й фрагменти таких (SEQ ID NO:5, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30 і 32), які кодують поліпептид стійкості або толерантності до гліфосату. Аналогічно, "поліпептид стійкості до гліфосату за винаходом" являє собою поліпептид, який має амінокислотну послідовність, наведену в SEQ ID NO:2 або 4, і варіанти й фрагменти таких (SEQ ID NO:6, 15, 17, 19, 21, 23, 25, 27, 29, 31 і 33), які надають клітині-хазяїнові стійкості або толерантності до гліфосату.

Плазміди, що містять нуклеотидні послідовності стійкості до гербіциду за винаходом, депоновані в постійній колекції Колекції Культур Служби сільськогосподарських досліджень, Північної Регіональної Науково-дослідної лабораторії (NRRL) 18 листопада 2005 за номером доступу NRRL B-30888 (grg23), і 26 червня 2006 за номером доступу NRRL B-30949 (grg51). Це депонування буде підтримуватися відповідно до Будапештського Договору про Міжнародне Визнання Депонування Мікроорганізмів для цілей Патентної Процедури. Це депонування було зроблено для зручності фахівців в галузі техніки, і не допускається, що депонування є необхідним за 35 кодексу законів США §112.

Під "гліфосатом" мається на увазі будь-яка гербіцидна форма N-фосфометилгліцину (включаючи будь-яку сіль такого) і інші форми, з яких отримується аніон гліфосату in planta. "Білок стійкості до гербіциду" або білок, який є продуктом експресії "молекули нуклеїнових кислот, що кодує стійкість до гербіциду", включає білки, які надають клітині здатність бути толерантною до більш високої концентрації гербіциду, ніж клітини, які не експресують білок, або переносити певну концентрацію гербіциду протягом більшого проміжку часу, ніж клітини, які не експресують білок. "Білок стійкості до гліфосату" включає білок, який надає клітині здатності залишатися толерантною до більш високої концентрації гліфосату, ніж клітини, які не експресують білок, або бути толерантною до певної концентрації гліфосату протягом більш довгого проміжку часу, ніж клітини, які не експресують білок. Під виразом "бути толерантним" або "толера-

нтність" мається на увазі або виживання, або виконання найважливіших клітинних функцій, таких як синтез білка й дихання, таким способом, який важко відрізнити від необроблених клітин.

Ізольовані молекули нуклеїнових кислот і варіанти й фрагменти таких

Один аспект винаходу стосується ізольованих молекул нуклеїнових кислот, які містять нуклеотидні послідовності, що кодують білки й поліпептиди стійкості до гербіциду або біологічно активні частини таких, так само як і молекули нуклеїнових кислот, що підходять для застосування як гібридизаційні зонди для ідентифікації нуклеїнових кислот, які кодують стійкість до гербіциду. Як використовуються в даному описі, термін "молекула нуклеїнової кислоти" призначений включати молекули ДНК (наприклад, кДНК або геномної ДНК) і молекули РНК (наприклад, мРНК) і аналоги ДНК або РНК, отримані з використанням нуклеотидних аналогів. Молекула нуклеїнової кислоти може бути одностанцюговою або дволанцюговою.

Нуклеотидні послідовності, які кодують білки за даним винаходом, включають послідовності, наведені в SEQ ID NO:1, 3 і 5, нуклеотидну послідовність стійкості до гербіциду, введена в бактеріального хазяїна під номерами доступу NRRL B-30888 і NRRL B-30949, і варіанти, фрагменти, і комплементарні до таких. Під терміном "комплементарний" мається на увазі нуклеотидна послідовність, яка достатньо комплементарна до даної нуклеотидної послідовності таким чином, що вона може гібридизуватися з даною нуклеотидною послідовністю, утворюючи, таким чином, стійкий дуплекс. Відповідна амінокислотна послідовність для білка стійкості до гербіциду, яка кодується цими нуклеотидними послідовностями, наведена в SEQ ID NO:2, 4 або 6. Винахід також охоплює молекули нуклеїнових кислот, які містять нуклеотидні послідовності, що кодують неповнорозмірні білки стійкості до гербіциду, і комплементарні до таких.

"Ізольована" або "очищена" молекула нуклеїнової кислоти або білка або біологічно активна частина такої, є по суті вільною від іншого клітинного матеріалу або культурального середовища, якщо виробляється за допомогою рекомбінантних методик, або є по суті вільною від хімічних попередників або інших хімічних реагентів, якщо синтезується хімічно. Переважно, "ізольована" нуклеїнова кислота вільна від послідовностей (переважно послідовностей, які кодують білок), які в природі примикають до нуклеїнової кислоти (тобто, послідовності, розташовані на 5' і 3' кінцях нуклеїнової кислоти) у геномній ДНК організму, з якого отримана нуклеїнова кислота. Для винаходу, термін "ізольована" застосовно до молекул нуклеїнових кислот, виключає ізольовані хромосоми. Наприклад, у різних варіантах здійснення, ізольована молекула нуклеїнової кислоти, яка кодує стійкість до гліфосату, може містити менше, ніж приблизно 5 т.п.н., 4 т.п.н., 3 т.п.н., 2 т.п.н., 1 т.п.н., 0,5 т.п.н. або 0,1 т.п.н. нуклеотидної послідовності, що у природі примикає до молекули нуклеїнової кислоти в геномній ДНК клітини, з якої отримана нуклеїнова кислота. Білок стійкості до гербіциду, який по суті не містить клітинного матеріалу, включає одержання білка, що має менше, ніж приблизно

30%, 20%, 10% або 5% (суха вага) білка стійкості до не гербіциду (також позначається у даному описі як "забруднюючий білок").

Молекули нуклеїнових кислот, які є фрагментами цих нуклеотидних послідовностей, які кодують стійкість до гербіциду, також входять до обсягу даного винаходу. Під "фрагментом" розуміється частина нуклеотидної послідовності, яка кодує білок стійкості до гербіциду. Фрагмент нуклеотидної послідовності може кодувати біологічно активну частину білка стійкості до гербіциду, або це може бути фрагмент, який може бути використаний як гібридизаційний зонд або ПЛР-праймер з використанням способів, розкритих нижче. Молекули нуклеїнових кислот, які є фрагментами нуклеотидної послідовності стійкості до гербіциду, містять принаймні приблизно 15, 20, 50, 75, 100, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950 послідовних нуклеотидів, або до числа нуклеотидів, присутніх у повнорозмірній нуклеотидній послідовності, яка кодує стійкість до гербіциду, розкритій в даному описі (наприклад, 1892 нуклеотидів для SEQ ID NO:1, 1259 нуклеотидів для SEQ ID NO:3, і 1242 нуклеотидів для SEQ ID NO:5). Під "послідовними" нуклеотидами маються на увазі нуклеотидні залишки, які є безпосередньо прилеглими один до одного.

Фрагменти нуклеотидних послідовностей за даним винаходом головним чином будуть кодувати фрагменти білка, які зберігають біологічну активність повнорозмірного білка стійкості до гліфосату; тобто, активність стійкості до гербіциду. Під виразом "зберігає активність стійкості до гербіциду", мається на увазі, що фрагмент буде мати принаймні приблизно 30%, принаймні приблизно 50%, принаймні приблизно 70% або принаймні приблизно 80% активності стійкості до гербіциду повнорозмірних білків стійкості до гліфосату, які розкриті в даному описі як SEQ ID NO:2, 4 або 6. Способи вимірювання активності стійкості до гербіциду відомі в галузі техніки. Див., наприклад, Патенти США №№ 4535060 і 5188642, кожний з яких включений у даний опис повністю за допомогою посилання.

Фрагмент нуклеотидної послідовності, яка кодує стійкість до гербіциду, що кодує біологічно активну частину білка за винаходом, буде кодувати принаймні приблизно 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400 послідовних амінокислот, або аж до загального числа амінокислот у повнорозмірному білку стійкості до гербіциду за даним винаходом (наприклад, 436 амінокислот для SEQ ID NO:2, 413 амінокислот для SEQ ID NO:4 і 413 амінокислот для SEQ ID NO:6).

Білки стійкості до гербіциду за даним винаходом кодуються нуклеотидною послідовністю, достатньо ідентичною до нуклеотидної послідовності SEQ ID NO:1, 3 або 5. Під терміном "достатньо ідентична" мають на увазі амінокислотну або нуклеотидну послідовність, яка має принаймні приблизно 60% або 65% ідентичність послідовності, приблизно 70% або 75% ідентичність послідовності, приблизно 80% або 85% ідентичність послідов-

ності, приблизно 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% або 99% ідентичність послідовності при зіставленні з послідовністю порівняння з використанням однієї із програм вирівнювання, описаних у даному описі, з використанням стандартних параметрів. Кваліфікований фахівець у цій галузі техніки визнає, що ці значення можуть бути підходящим чином пристосовані для визначення відповідних ідентичностей білків, які кодуються двома нуклеотидними послідовностями за допомогою прийняття в увагу виродженості кодонів, подібності амінокислот, розташування рамки читання, і т.п.

Щоб визначити відсоток ідентичності двох послідовностей амінокислот або двох нуклеїнових кислот, послідовності вирівнюють із метою оптимального порівняння. Відсоток ідентичності між двома послідовностями є функцією числа ідентичних положень, які є загальними для послідовностей (тобто, відсоток ідентичності = число ідентичних положень/загальна кількість положень (наприклад, положення, які перекриваються) \times 100). В одному варіанті здійснення дві послідовності мають однакову довжину. Відсоток ідентичності між двома послідовностями може бути визначений з використанням методик, подібних до описаних нижче, з або без допущення пропусків. При обчисленні відсотка ідентичності, як правило, підраховують точні співпадіння.

Визначення ідентичності відсотка між двома послідовностями може бути досягнуте з використанням математичного алгоритму. Необмежувальним прикладом математичного алгоритму, який використовують для порівняння двох послідовностей, є алгоритм Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, змінений як в Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Такий алгоритм включений в BLASTN і програми BLASTX від Altschul et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST нуклеотидні пошуки можуть бути виконані програмою BLASTN, число очок = 100, довжина слова = 12, для одержання нуклеотидних послідовностей, гомологічних до GDC-подібних молекул нуклеїнових кислот за винаходом. BLAST пошуки білка можуть бути виконані програмою BLASTX, число очок = 50, довжина слова = 3, для одержання амінокислотних послідовностей, гомологічних до молекул білка стійкості до гербіциду за винаходом. Щоб одержати, з метою порівняння, вирівнювання з розривами, може бути використаний Gapped BLAST, як описано в Altschul et al. (1997) *Nucleic Acid Res.* 25:3389-3402. Як альтернатива, для виконання повторного пошуку, який виявляє віддалені подібності між молекулами, може бути застосований PSI-Blast. Див. Altschul et al. (1997) вище. При використанні BLAST, Gapped BLAST і PSI-BLAST програм, можна використовувати параметри за замовчуванням відповідних програм (наприклад, BLASTX і BLASTN). Див. www.ncbi.nlm.nih.gov. Іншим необмежувальним прикладом математичного алгоритму, який використовують для порівняння послідовностей, є алгоритм ClustalW (Higgins et al. (1994) *Nucleic Acid Res.* 22:4673-4680). ClustalW порівнює послідовності й вирівнює амінокислотні або ДНК-послідовності повністю, і в такий спосіб

може надати дані про консервативність послідовності всієї амінокислотної послідовності. Алгоритм ClustalW використовуються в декількох комерційно доступних пакетах програм для аналізу ДНК/амінокислот, таких як модуль ALIGNX Набір Програми Vector NTI (Корпорація Invitrogen, Carlsbad, CA). Після вирівнювання амінокислотних послідовностей ClustalW може бути визначений відсоток амінокислотної ідентичності. Необмежувальним прикладом програми, придатної для аналізу вирівнювань ClustalW, є GeneDoc™. Genedoc™ (Karl Nicholas) дозволяє оцінювати амінокислотну подібність (або ДНК) і ідентичність між декількома білками. Іншим необмежувальним прикладом математичного алгоритму, який використовують для порівняння послідовностей, є алгоритм Myers і Miller (1988) CABIOS 4:11-17. Такий алгоритм включений у програму ALIGN (версія 2.0), що є частиною пакета програм для вирівнювання послідовностей GCG (доступний від Accelrys, Inc., Сан-Дієго, Каліфорнія). Використовуючи програму ALIGN для порівняння амінокислотних послідовностей, можуть бути застосовані таблиця ваги залишків PAM12 0, штраф довжини розриву 12 і штраф розриву 4.

Якщо не вказано інакше, Версія 10 GAP, яка використовує алгоритм Needleman і Wunsch (1970) *J. Mol. Biol.* 48 (3):443-453, буде використана для визначення ідентичності послідовності або подібності з використанням наступних параметрів: % ідентичності й % подібності для нуклеотидної послідовності з використанням GAP Weight 50 і Length Weight 3 і `nwsgapdna.cmp` матриці очок; % ідентичності або % подібності для амінокислотної послідовності з використанням GAP Weight 1 і Length Weight 2, і програма нарахування очок BLOSUM62. Також можуть бути використані еквівалентні програми. Під "еквівалентною програмою" мають на увазі будь-яку програму порівняння послідовностей, яка для будь-яких двох розглянутих послідовностей, робить вирівнювання, що має ідентичні співпадіння нуклеотидних залишків і ідентичний відсоток ідентичності послідовностей у порівнянні з відповідним вирівнюванням, зробленим за допомогою GAP Версії 10.

Винахід також охоплює молекули варіантів нуклеїнової кислоти. "Варіанти" нуклеотидних послідовностей, які кодують стійкість до гербіциду, включають ті послідовності, які кодують білок стійкості до гербіциду, розкритий у даному описі, але який консервативно відрізняється через виродженість генетичного коду, так само, як і ті, які досить ідентичні, як обговорено вище (наприклад, SEQ ID NO:5, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30 і 32 є варіантами SEQ ID NO:1). Природні алельні варіанти можуть бути ідентифіковані з використанням відомих методів молекулярної біології, таких як полімеразна ланцюгова реакція (PCR) і методи гібридизації, як викладено загалом нижче. Варіантні нуклеотидні послідовності також включають синтетично отримані нуклеотидні послідовності, які були отримані, наприклад, з використанням сайт-спрямованого мутагенезу, але які усе ще кодують білки стійкості до гербіциду, розкриті в даному винаході, як обговорено нижче.

Варіантні білки, охоплені даним винаходом, є біологічно активними, якщо вони зберігають бажану біологічну активність природного білка, тобто, активність стійкості до гербіциду. Під "зберігає активність стійкості до гербіциду" розуміють, що варіант буде мати принаймні приблизно 30%, принаймні приблизно 50%, принаймні приблизно 70% або принаймні приблизно 80% активності стійкості до гербіциду природного білка. Способи вимірювання активності стійкості до гербіциду добре відомі в галузі техніки. Див., наприклад, Патенти США №№ 4535060 і 5188642, кожний з яких включений у даний винахід повністю за допомогою посилаючого.

Кваліфікований фахівець надалі оцінить, що в нуклеотидні послідовності за винаходом можуть бути введені зміни за допомогою мутації, у такий спосіб приводять до змін в амінокислотній послідовності кодованих білків стійкості до гербіциду, без зміни біологічної активності білків. Таким чином, ізольовані варіантні молекули нуклеїнових кислот можуть бути створені за допомогою введення однієї або декількох нуклеотидних замін, додавань або делецій у відповідну нуклеотидну послідовність, розкриту в даному описі, таким чином, що в кодований білок будуть введені одна або декілька амінокислотних замін, додавань або делецій. Мутації можуть бути введені за допомогою стандартних методик, таких як сайт-спрямований мутагенез і ПЛР-опосередкований мутагенез. Такі варіантні нуклеотидні послідовності також входять до обсягу даного винаходу.

Наприклад, консервативні амінокислотні заміни можуть бути зроблені в одному або декількох передбачених несуттєвих амінокислотних залишках. "Несуттєвим" амінокислотним залишком є залишок, який може бути змінений у послідовності білка дикого типу стійкості до гербіциду без змін біологічної активності, тоді як "істотний" амінокислотний залишок необхідний для біологічної активності. "Консервативна амінокислотна заміна" є такою, у якій амінокислотний залишок заміщений на амінокислотний залишок, який має подібний боковий ланцюг. Сімейства амінокислотних залишків, що мають подібні бокові ланцюги, були визначені в галузі техніки. Ці сімейства включають амінокислоти з основними боковими ланцюгами (наприклад, лізин, аргінін, гістидин), кислотними боковими ланцюгами (наприклад, аспарагінова кислота, глутамінова кислота), незарядженими полярними боковими ланцюгами (наприклад, гліцин, аспарагін, глутамін, серин, треонін, тирозин, цистеїн), неполярними боковими ланцюгами (наприклад, аланін, валін, лейцин, ізолейцин, пролін, фенілаланін, метіонін, триптофан), бета-розгалуженими боковими ланцюгами (наприклад, треонін, валін, ізолейцин) і ароматичними боковими ланцюгами (наприклад, тирозин, фенілаланін, триптофан, гістидин). Амінокислотні заміни можуть бути зроблені в неконсервативних ділянках, які зберігають функцію. Головним чином, такі заміни не були б зроблені для консервативних амінокислотних залишків або для амінокислотних залишків, що знаходяться у межах консервативного мотиву, де такі залишки є істотними для активності білка. Однак, кваліфікований фахівець у цій галузі техні-

ки зрозумів би, що функціональні варіанти можуть мати малі консервативні або неконсервативні зміни в консервативних залишках.

Lys-22, Arg-124, Asp-313, Arg-344, Arg-386 і Lys-411 є консервативними залишками EPSP-синтази з *E. coli* (Schonbrunn et al. (2001) Proc. Natl. Acad. Sci. USA 98:1376-1380). Консервативні залишки, важливі для EPSP-синтазної активності також включають Arg-100, Asp-242 і Asp-384 (Selvapandian et al. (1995) FEBS Letters 374:253-256). Arg-27 зв'язується з S3P (Shuttleworth et al. (1999) Biochemistry 38:296-302).

Як альтернатива, варіантні нуклеотидні послідовності можуть бути отримані за допомогою введення мутацій випадковим чином уздовж всієї або частини кодувальної послідовності, наприклад, за допомогою мутагенезу насичення, і отримані мутанти можуть бути перевірені на здатність забезпечувати активність стійкості до гербіциду для ідентифікації мутантів, які зберігають активність. Після мутагенезу, кодований білок може бути експресований рекомбінантним способом, і активність білка може бути визначена з використанням стандартних методів тестування.

Використовуючи способи, такі як ПЛР, гібридизація й т.п., можуть бути ідентифіковані відповідні послідовності стійкості до гербіциду, такі послідовності, які мають істотну ідентичність із послідовностями за винаходом. Див., наприклад, Sambrook і Russell (2001) Molecular Cloning: Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) і Innis, et al. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, St.-Louis, MO).

У способі гібридизації вся або частина нуклеотидної послідовності стійкості до гербіциду може бути використана для скринингу кДНК або геномних бібліотек. Способи створення таких кДНК і геномних бібліотек є загальновідомими в галузі техніки й розкриті в Sambrook and Russell (2001) вище. Так звані гібридизаційні зонди можуть бути фрагментами геномної ДНК, фрагментами кДНК, фрагментами РНК або іншими олігонуклеотидами, і можуть бути помічені групою, яка виявляється, такою як ³²P або будь-яким іншим маркером, який виявляється, таким як інші радіоізотопи, флуоресцентна сполука, фермент або кофактор ферменту. Зонди для гібридизації можуть бути отримані за допомогою мічення синтетичних олігонуклеотидів, ґрунтуючись на відомій(их) нуклеотидній(их) послідовності(ях), що кодує(ють) стійкість до гербіциду, розкритої(их) у даному описі. Додатково можуть бути використані вироджені праймери, сконструйовані на основі консервативних нуклеотидів або амінокислотних залишків у нуклеотидній послідовності або кодованій амінокислотній послідовності. Зонд, як правило, містить ділянку нуклеотидної послідовності, яка гібридується в жорстких умовах із принаймні приблизно 12, принаймні приблизно 25, принаймні приблизно 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600 або 1800 послідовними нуклеотидами нуклеотидної послідовності(ей), що кодує стійкість до гербіциду за винаходом, або фрагмента або варіанта таких. Способи для одержання зондів для гібридизації є

загальновідомими в галузі техніки й розкриті в Sambrook and Russell (2001) вище, і Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY), кожний з яких включений у даний опис за допомогою посилання.

Наприклад, повна послідовність стійкості до гербіциду, розкрита в даному описі, або одна або більше частин такої, можуть бути використані як зонд, здатний специфічно гібридуватися з відповідними послідовностями стійкості до гербіциду й з матричними РНК. Для досягнення специфічної гібридизації при різних умовах, такі зонди включають послідовності, які є унікальними і являють собою принаймні приблизно 10 нуклеотидів у довжину й принаймні приблизно 20 нуклеотидів у довжину. Такі зонди можуть бути використані для ампліфікації відповідних послідовностей стійкості до гербіциду з вибраного організму за допомогою ПЛР. Ця методика може бути використана для виділення додаткових кодувальних послідовностей, з бажаного організму або як діагностичний тест для визначення присутності кодувальних послідовностей, в організмі. Методи гібридизації включають гібридизаційне скринування висіяних на чашки бібліотек ДНК (або бляшок, або колоній; див., наприклад, Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Гібридизація таких послідовностей може бути виконана в жорстких умовах. Під "жорсткими умовами" або "жорсткими умовами гібридизації" мають на увазі умови, у яких зонд гібридується зі своєю послідовністю-мішенню в детектовано більшому ступені, ніж з іншими послідовностями (наприклад, принаймні з 2-кратним перевищенням фону). Жорсткі умови є залежними від послідовності й будуть різні при різних обставинах. За допомогою керування жорсткістю гібридизації та/або умовами промивання, послідовності-мішені, які на 100% комплементарні до зонду, можуть бути ідентифіковані (homologous probing). Як альтернатива, жорсткі умови можуть бути відрегульовані для дозволу деякої невідповідності в послідовностях так, щоб були виявлені більш низькі ступені подібності (heterologous probing). Звичайно, зонд становить у довжину менше, ніж приблизно 1000 нуклеотидів, або менше, ніж приблизно 500 нуклеотидів у довжину.

Як правило, жорсткими умовами будуть такі, при яких концентрація солі менше, ніж приблизно 1,5 М іон Na, як правило, від приблизно 0,01 до 1,0 М концентрації іона Na (або інших солей) при рН від 7,0 до 8,3 і температурі принаймні приблизно 30°C для коротких зондів (наприклад, від 10 до 50 нуклеотидів) і принаймні приблизно 60°C для довгих зондів (наприклад, більше, ніж 50 нуклеотидів). Жорсткі умови можуть також бути досягнуті додаванням агентів для дестабілізації, таких як формамід. Ілюстративні умови низької жорсткості включають гібридизацію в буферному розчині 30%-35%-ого формаміду, 1 М NaCl, 1%-ий SDS (натрію додецилсульфат) при 37°C, і промивання в 1X-2X SSC (20X SSC = 3,0 М NaCl/0,3 М тринатрію цитрат) при 50-55°C. Ілюстративні умови помірної жорсткості включають гібридизацію в 40%-45%-

ому формаміді, 1,0 М NaCl, 1%-ому SDS при 37°C, і промивання в 0,5X-1X SSC при 55-60°C. Ілюстративні умови високої жорсткості включають гібридизацію в 50%-ому формаміді, 1 М NaCl, 1%-ому SDS при 37°C, і промивання в 0,1X SSC при 60-65°C. Необов'язково, буфери для промивання можуть містити від приблизно 0.1% до приблизно 1% SDS. Тривалість гібридизації становить звичайно менше ніж приблизно 24 години, звичайно від приблизно 4 до приблизно 12 годин.

Специфічність, як правило, є функцією постгібридизаційних промивань, критичними факторами є іонна сила й температура розчину для заключного промивання. Для ДНК-ДНК гібридів, T_m може бути наближена до рівняння Meinkoth and Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81,5^\circ\text{C} + 16,6 (\log M) + 0,41 (\%GC) - 0,61 (\% \text{ форм.}) - 500/L$; де M є молярністю одновалентних катіонів, $\%GC$ - відсоток нуклеотидів гуанозину й цитозину в ДНК, $\% \text{ форм.}$ - відсоток формаміду в розчині гібридизації, і L - довжина гібриду в парах азотистих основ. T_m є температурою (при певній іонній силі й рН), при якій 50% комплементарної послідовності-мішені гібридується з ідеально відповідною пробою. T_m зменшується приблизно на 1°C з кожним 1% невідповідності; таким чином, T_m гібридизація та/або умови промивання можуть бути відрегульовані для гібридизації з послідовностями бажаної ідентичності. Наприклад, якщо потрібно знайти послідовності з $\geq 90\%$ -ою ідентичністю, T_m може бути зменшена на 10°C . Звичайно жорсткі умови вибираються, щоб бути приблизно на 5°C нижче, ніж температура точки плавлення (T_m) для специфічної послідовності й комплементарної до неї при певній іонній силі та рН. Однак, при дуже жорстких умовах гібридизація та/або промивання можуть проводитися при 1, 2, 3 або 4°C нижче, ніж температура точки плавлення (T_m); при помірно жорстких умовах гібридизація та/або промивання можуть проводитися при 6, 7, 8, 9, або 10°C нижче ніж температура точки плавлення (T_m); в умовах низької жорсткості гібридизація та/або промивання можуть проводитися при 11, 12, 13, 14, 15 або 20°C нижче ніж температура точки плавлення (T_m). Застосовуючи рівняння, гібридизацію й склади промивання й бажану T_m , середньому фахівцеві зрозуміло, що розмаїтість у жорсткості гібридизації та/або розчинів промивання по суті описано. Якщо бажаний ступінь невідповідності приводить до T_m менше, ніж 45°C (водний розчин) або 32°C (розчин формаміду), переважно збільшити концентрацію SSC так, щоб можна було використати більш високу температуру. Великий посібник з гібридизації нуклеїнових кислот знайдено в Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); і Ausubel et al., eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). Див. Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Ізольовані білки й варіанти й фрагменти таких

Білки стійкості до гербіциду також охоплені обсягом даного винаходу. Під "білком стійкості до гербіциду" або "білком толерантності до гербіциду"

мають на увазі білок, який має амінокислотну послідовність, представлену в SEQ ID NO:2, 4 або 6. Фрагменти, біологічно активні частини й варіанти таких також представлені й можуть бути використані для здійснення на практиці способів за даним винаходом.

"Фрагменти" або "біологічно активні частини" включають поліпептидні фрагменти, що містять частину амінокислотної послідовності, яка кодує білок стійкості до гербіциду, як наведено в SEQ ID NO:2, 4, або 6, і які зберігають активність стійкості до гербіциду. Біологічно активна частина білка стійкості до гербіциду може бути поліпептидом, що, наприклад, має 10, 25, 50, 100 або більше амінокислот у довжину. Такі біологічно активні частини можуть бути отримані за допомогою рекомбінантних методик і оцінені на активність стійкості до гербіциду. Способи вимірювання активності стійкості до гербіциду добре відомі в галузі техніки. Див., наприклад, Патенти США №№ 4535060 і 5188642, кожний з яких включений у даний опис повністю за допомогою посилання. Як використовуються в даному описі, фрагмент містить принаймні 8 послідовних амінокислот SEQ ID NO:2, 4 або 6. Однак винахід охоплює інші фрагменти, такі як, наприклад, будь-який фрагмент у білку, більший, ніж приблизно 10, 20, 30, 50, 100, 150, 200, 250, 300, 350 або 400 амінокислот.

Під "варіантами" маються на увазі білки або поліпептиди, що мають амінокислотну послідовність, яка принаймні приблизно на 60%, 65%, принаймні приблизно на 70%, 75%, принаймні приблизно на 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% або 99%, ідентична до амінокислотної послідовності SEQ ID NO:2, 4, або 6 (наприклад, SEQ ID NO:6, 15, 17, 19, 21, 23, 25, 27, 29, 31 і 33 є варіантами SEQ ID NO:2). Варіанти також включають поліпептиди, які кодуються молекулою нуклеїнової кислоти, що гібридується з молекулою нуклеїнової кислоти SEQ ID NO:1, 3 або 5, або комплексарна до таких, у жорстких умовах. Варіанти включають поліпептиди, які відрізняються за амінокислотною послідовністю внаслідок мутагенезу. Варіантні білки, охоплені даним винаходом, є біологічно активними, так що вони продовжують мати бажану біологічну активність вихідного білка, тобто, зберігають активність стійкості до гербіциду. Способи вимірювання активності стійкості до гербіциду добре відомі в галузі техніки. Див., наприклад, Патенти США №№ 4535060 і 5188642, кожний з яких включений у даний опис повністю за допомогою посилання.

Бактеріальні гени, такі як ген grg23 або grg51 за даним винаходом, досить часто мають множинні ініціюючі кодони метіоніну поблизу початку відкритої рамки читування. Часте ініціювання трансляції в одному або декількох із цих старт-кодонів приведе до утворення функціонального білка. Ці старт-кодони можуть включати ATG-кодони. Однак, бактерії, такі як *Bacillus* sp., також впізнають кодон GTG як старт-кодон, і білки, які починають трансляватися в кодонах GTG, містять метіонін у першій амінокислоті. Крім того, це не часто визнається а *priori*, які із цих кодонів використовуються в природі в бактерії. Таким чином, зрозуміло, що використання одного з додаткових кодонів ме-

тіоніну може привести до утворення варіантів grg23 або grg51, які мають стійкість до гербіциду. Ці білки стійкості до гербіциду входять до обсягу даного винаходу й можуть бути використані в способах за даним винаходом.

Також охоплені антитіла до поліпептидів за даним винаходом, або до варіантів або фрагментів таких. Способи одержання антитіл добре відомі в галузі техніки (див., наприклад, Harlow and Lane (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY); Патент США № 4196265).

Змінені або поліпшені варіанти

Визнано, що послідовність ДНК grg23 або grg51 може бути змінена за допомогою різних способів, і що ці зміни можуть привести до ДНК послідовностей, які кодують білки з амінокислотними послідовностями, відмінними від тих, які кодуються grg23 або grg51. Цей білок може бути змінений різними шляхами, включаючи амінокислотні заміни, делеції, усикання й інсерції. Способи таких маніпуляцій є загальновідомими в галузі техніки. Наприклад, варіанти амінокислотної послідовності білка GRG23 або GRG51 можуть бути отримані за допомогою мутацій у ДНК. Це також може бути досягнуте за допомогою однієї з декількох форм мутагенезу та/або за допомогою спрямованої еволюції. У деяких аспектах, зміни, які кодуються в амінокислотній послідовності, не будуть по суті впливати на функцію білка. Такі варіанти будуть мати бажану активність стійкості до гербіциду. Однак, зрозуміло, що здатність GRG23 або GRG51 надавати стійкість до гербіциду може бути поліпшена за допомогою використання таких методик у композиціях за даним винаходом. Наприклад, GRG23 або GRG51 можуть експресуватися в клітинах-хазяїнах, які проявляють високі рівні помилкового включення під час реплікації ДНК, таких як XL-1 Red (Stratagene, La Jolla, CA). Після відтворення в таких штаммах, ДНК grg23 або grg51 може бути ізольована (наприклад, за допомогою одержання плазмідної ДНК або за допомогою ПЛР-ампліфікації й клонування ПЛР-фрагмента, який отримують, у вектор) і культивуватися в немутагенних штаммах. Клоні, що містять мутації в grg23 або grg51, можуть бути ідентифіковані за допомогою вимірювання поліпшеної стійкості до гербіциду, такому як гліфосат, наприклад, за допомогою вирощування клітин на зростаючих концентраціях гліфосату й тестування для пошуку клонів, які мають толерантність до зростаючих концентрацій гліфосату.

Як альтернатива, можуть бути зроблені зміни з послідовністю білка багатьох білків з аміно- або карбокси-кінця, по суті не впливаючи на активність. Ці зміни можуть включати інсерції, делеції або зміни, що вводять за допомогою сучасних молекулярних способів, таких як ПЛР, включаючи ПЛР-ампліфікацію, що змінює або подовжує послідовність, яка кодує білок, за допомогою включення послідовностей, які кодують амінокислоти в олігонуклеотидах, що використовують при ПЛР-ампліфікації. Як альтернатива, додані послідовності білка можуть включати повні послідовності, які кодують білок, такі як ті, які звичайно застосовуються в галузі техніки для одержання гібридних

білків. Такі гібридні білки часто застосовуються для (1) збільшення експресії білка, який представляє інтерес; (2) введення зв'язувального домену, ферментативної активності або епітопу для поліпшення або очищення білка, або виявлення білка або інших експериментальних застосувань, відомих в галузі техніки; або, (3) адресній секреції або трансляції білка в субклітинну органелу, таку як периплазматичний простір грамнегативних бактерій або ендоплазматична сітка еукаріотичних клітин, остання з яких часто приводить до глікозилування білка.

Варіантні нуклеотидні й амінокислотні послідовності за даним винаходом також охоплюють послідовності, отримані за допомогою мутагенних і рекомбінантних процедур, таких як перетасування ДНК. За допомогою такої процедури одна або декілька різних кодувальних ділянок, білків стійкості до гербіциду може бути використана для створення нового білка стійкості до гербіциду, який має бажані властивості. Таким чином, бібліотеки рекомбінантних полінуклеотидів формуються із числа споріднених полінуклеотидних послідовностей, що включають ділянки послідовності, які мають істотну ідентичність послідовності й можуть бути гомологічно рекомбіновані *in vitro* або *in vivo*. Наприклад, застосовуючи цей підхід, фрагменти послідовності, яка кодує домен, що представляє інтерес, можуть бути перетасовані між геном стійкості до гербіциду за винаходом й іншими відомими генами стійкості до гербіциду для одержання нового гена, що кодує білок з поліпшеною властивістю, яка представляє інтерес, такою як збільшення активності стійкості до гліфосату. Стратегії для такого перетасування ДНК відомі в галузі техніки. Див., наприклад, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer et al. (1997) *Nature Biotech.* 15:436-438; Moore et al. (1997) *J. Mol. Biol.* 272:336-347; Zhang et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer et al. (1998) *Nature* 391:288-291; і Патенти США №№ 5605793 і 5837458.

Трансформація бактеріальних або рослинних клітин

Представлене в даному описі є новими ізольованими генами, які надають стійкість до гербіциду. Також представлено є амінокислотними послідовностями білків GRG23 і GRG51. Білок, який утворюється при трансляції цього гена, дозволяє клітинам функціонувати в присутності концентрацій гербіциду, які інакше були б токсичні для клітин, включаючи рослинні клітини й бактеріальні клітини.

В одному аспекті винаходу, ген *grg23* або *grg51* застосовний як маркерний ген для оцінки трансформації бактеріальних або рослинних клітин.

За допомогою конструювання *grg23* або *grg51* для (1) експресії з бактеріальним промотором, що стимулює транскрипцію в тестованому організмі, (2) правильної трансляції для утворення незмінного пептиду GRG23 або GRG51, і (3) переміщення клітин в іншу токсичну концентрацію гербіциду, можуть бути ідентифіковані клітини, які були трансформовані ДНК за допомогою їх стійкості до гербіциду. Під "промотором" мають на увазі послідов-

ність нуклеїнової кислоти, яка функціонує, щоб направити транскрипцію розташованої далі за напрямом кодувальної послідовності. Промотор, разом з іншими послідовностями нуклеїнової кислоти, які регулюють транскрипцію й трансляцію (які також називаються "регуляторні послідовності"), необхідний для експресії послідовності ДНК, яка представляє інтерес.

Трансформація бактеріальних клітин здійснюється за допомогою однієї з декількох методик, відомих в галузі техніки, включаючи, але не обмежуючись електропорацією або хімічною трансформацією (Див., наприклад, Ausubel (ed.) (1994) *Current Protocols in Molecular Biology* (John Wiley and Sons, Inc., Indianapolis, IN)). Маркери, що несуть стійкість до токсичних речовин, застосовуються для ідентифікації трансформованих клітин (які захопили і експресують тестовану ДНК) від нетрансформованих клітин (тих, які не містять або не експресують тестовану ДНК). В одному аспекті винаходу, ген *grg23* або *grg51* застосовується як маркерний ген для оцінки трансформації бактеріальних або рослинних клітин.

Трансформація рослинних клітин може здійснюватися подібним способом. Під "рослиною" мають на увазі цілі рослини, органи рослини (наприклад, листи, стебла, корінь, і т.д.), насіння, рослинні клітини, пагони, ембріони й потомство такої. Рослинні клітини можуть бути диференційованими або недиференційованими (наприклад, каллус, суспензійні культури клітин, протопласти, клітини листя, клітини кореня, клітини флоєми, пилки). "Трансгенні рослини" або "трансформовані рослини" або "стабільно трансформовані" рослини, клітини або тканини належать до рослин, які увібрали або інтегрували послідовності екзогенної нуклеїнової кислоти або фрагменти ДНК у рослинну клітину. Під "стабільною трансформацією" мають на увазі, що нуклеотидна конструкція, введена в рослину, інтегрує в геном рослини й здатна до того, щоб бути успадкованою потомством такої.

Ген *grg23* або *grg51* за винаходом може бути модифікований для одержання або посилення експресії в рослинних клітинах. Послідовності стійкості до гербіциду за винаходом можуть бути представлені в експресійних касетах для експресії в рослині, що представляє інтерес. "Касета для експресії в рослині" включає конструкції ДНК, які здатні приводити до експресії білка з відкритої рамки зчитування в рослинній клітині. Касета буде включати в 5'-3' напрямку транскрипції, ділянку ініціації транскрипції (тобто, промотор), функціонально зв'язану з послідовністю ДНК за винаходом, і ділянку термінації транскрипції й трансляції (тобто, ділянку термінації), які функціонують у рослинах. Касета може додатково містити принаймні один додатковий ген для котрансформації в організм, такий як ген маркера селекції. Як альтернатива, додатковий ген(и) може бути представлений у декількох касетах для експресії. Така касета для експресії представлена з багатьма рестриктивними сайтами для інсерції послідовності стійкості до гербіциду під транскрипційною регуляцією регуляторних ділянок.

Промотор може бути нативним або аналогічним, або стороннім або гетерологічним, до росли-

ни-хазяїна та/або до послідовності ДНК за винаходом. Додатково, промотор може бути послідовністю, що зустрічається в природі, або, на противагу, синтетичною послідовністю. Якщо промотор є "нативним" або "гомологічним" до рослини-хазяїна, це означає, що промотор походить із нативної рослини, у яку введений промотор. Якщо промотор є "стороннім" або "гетерологічним" до послідовності ДНК за винаходом, це означає, що промотор не походить або не є промотором, що зустрічається в природі, функціонально зв'язаним з послідовністю ДНК за винаходом. "Гетерологічний" звичайно стосується послідовностей нуклеїнової кислоти, які не є ендегенними для клітини або частиною нативного геному, у якому вони знаходяться, і були додані до клітини за допомогою інфекції, трансфекції, мікроін'єкції, електропорації, мікропроекції, або т.п. Під "функціонально зв'язаний" мають на увазі функціональний зв'язок між промотором і другою послідовністю, якщо послідовність промотору ініціює й опосередковує транскрипцію з послідовності ДНК, що відповідає другій послідовності. Звичайно, "функціонально зв'язаний" означає, що зв'язані послідовності нуклеїнової кислоти є послідовними й, при необхідності приєднання двох послідовних ділянок, що кодуєть білок, послідовні в одній і тій же самій рамці зчитування.

Часто такі конструкції будуть також містити 5' і 3' нетрансльовані ділянки. Такі конструкції можуть містити "сигнальну послідовність" або "лідерну послідовність" для полегшення котрансляції або посттрансляційного транспорту пептиду, що представляє інтерес, у певні внутрішньоклітинні структури, такі як хлоропласт (або інший пластид), ендоплазматична сітка або апарат Гольджі, або бути секретованим. Наприклад, ген може в результаті конструювання містити сигнальний пептид для полегшення передачі пептиду в ендоплазматичну сітку. Під "сигнальною послідовністю" мають на увазі послідовність, яка відома або, як вважають, приведе до котрансляційного або посттрансляційного транспорту пептиду через клітинну мембрану. В еукаріот цей транспорт звичайно задіює секрецію в апарат Гольджі, з деяким глікозилюванням. Під "лідерною послідовністю" розуміють будь-яку послідовність, яка при трансляції приводить до амінокислотної послідовності, достатньої для запуску котрансляційного транспорту пептидного ланцюга в субклітинну органелу. Таким чином, вказане вище включає транспорт, спрямований лідерною послідовністю, та/або глікозилювання за допомогою проходів в ендоплазматичну сітку, проходів до вакуоль, пластид, включаючи хлоропласти, мітохондрії, і т.п. Касета для експресії в рослині може також бути складена таким чином, щоб містити інтрон, процесинг мРНК якого необхідний для експресії.

Під "3' нетрансльованою ділянкою" мають на увазі нуклеотидну послідовність, розташовану за ходом транскрипції кодувальної послідовності. Послідовності сигналу поліаденілювання та інші послідовності, що кодуєть регуляторні сигнали, здатні впливати на додавання поліаденіловокислих хвостів до 3-кінця попередника мРНК, є 3' нетрансльованими ділянками. Під "5' нетрансльованою ділянкою" розуміють нуклеотидну

послідовність, розташовану проти ходу транскрипції кодувальної послідовності.

Інші розташовані проти ходу транскрипції або за ходом транскрипції нетрансльовані елементи включають енхансери. Енхансерами є нуклеотидні послідовності, які приводять до збільшення експресії промоторної ділянки. Енхансери відомі в галузі техніки й включають, але не обмежені, ділянкою енхансера SV40 і енхансерним елементом 35S.

Ділянка термінації може бути нативною по відношенню до ділянки ініціації транскрипції, може бути нативною по відношенню до послідовності стійкості до гербіциду за даним винаходом, або може бути отримана з іншого джерела. Зручні ділянки термінації доступні з Ті-плазмиди *A. tumefaciens*, такі як ділянки термінації октопінсинтази й нопалінсинтази. Див. також Guerineau et al. (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon et al. (1991) *Genes Dev.* 5:141-149; Mogen et al. (1990) *Plant Cell* 2:1261-1272; Munroe et al. (1990) *Gene* 91:151-158; Ballas et al. (1989) *Nucleic Acids Res.* 17:7891-7903; і Joshi et al. (1987) *Nucleic Acid Res.* 25 15:9627-9639.

За необхідності, ген(и) може бути оптимізований для збільшення експресії в трансформованій клітині-хазяїні. Таким чином, гени можуть бути синтезовані, застосовуючи кодони, переважні для клітини-хазяїна, для поліпшеної експресії, або можуть бути синтезовані, застосовуючи кодони, переважні для хазяїна, використовуючи повтори. Звичайно, GC вміст гена буде збільшено. Див., наприклад, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 для обговорення кодонів, переважних для хазяїна. Способи синтезування генів, переважних для хазяїна, є широко відомими в галузі техніки. Див., наприклад, Патенти США №№ 6320100; 6075185; 5380831 і 5436391, Опубліковані заявки США №№ 20040005600 і 20010003849, і Murray et al. (1989) *Nucleic Acids Res.* 17:477-498, включені в даний опис за допомогою посилання.

В одному варіанті здійснення, нуклеїнові кислоти, що представляють інтерес, націлені для експресії в хлоропласт. У даному способі, якщо нуклеїнова кислота, що представляє інтерес, безпосередньо не вставлена в хлоропласт, касета для експресії буде додатково містити нуклеїнову кислоту, яка кодує транзитний пептид для напрямку генетичного продукту, що представляє інтерес, у хлоропласти. Такі транзитні пептиди відомі в галузі техніки. Див., наприклад, Von Heijne et al. (1991) *Plant Mol. Biol. Rep.* 9:104-126; Clark et al. (1989) *J. Biol. Chem.* 264:17544-17550; Della-Cioppa et al. (1987) *Plant Physiol.* 84:965-968; Romer et al. (1993) *Biochem. Biophys. Res. Commun.* 196:1414-1421; і Shah et al. (1986) *Science* 233:478-481.

Нуклеїнові кислоти, що представляють інтерес, які будуть націлені в хлоропласт, можуть бути оптимізовані для експресії в хлоропласті, беручи до уваги відмінності у використанні кодонів між ядром рослини й цією органелою. Таким чином, нуклеїнові кислоти, що представляють інтерес, можуть бути синтезовані, застосовуючи кодони, переважні для хлоропласта. Див., наприклад, Па-

тент США № 5380831, включений у даний опис за допомогою посилання.

Як правило, дана "касета для експресії в рослині" буде вставлена в "вектор для трансформації рослини". Під "вектором для трансформації" мають на увазі молекулу ДНК, що необхідна для ефективної трансформації клітини. Така молекула може складатися з однієї або декількох касет для експресії, і може бути організована в більше ніж одну "векторну" молекулу ДНК. Наприклад, подвійні вектори є векторами для трансформації рослини, які використовують два несповідні ДНК-вектори для кодування всіх необхідних цис- і транс-діючих функцій для трансформації рослинних клітин (Hellens and Mullineaux (2000) Trends in Plant Science 5:446-451). "Вектор" стосується конструкції нуклеїнової кислоти, сконструйованої для передачі між різними клітинами-хазяями. "Вектор для експресії" стосується вектора, який здатний вбудовуватися, інтегрувати й експресувати гетерологічні послідовності ДНК або фрагменти в чужорідній клітині.

Цей вектор для трансформації рослини може включати один або декілька ДНК-векторів, необхідних для досягнення трансформації рослини. Наприклад, звичайною практикою в галузі техніки є використання векторів для трансформації рослини, які складаються з більше ніж одного послідовного ДНК-сегмента. Ці вектори часто згадуються в галузі техніки як "подвійні вектори". Подвійні вектори так само, як і вектори з допоміжними плазмідами найчастіше застосовуються для *Agrobacterium*-опосередкованої трансформації, коли необхідні досить великий розмір і складність сегментів ДНК для досягнення ефективної трансформації, і вигідно розділити функції між окремими молекулами ДНК. Подвійні вектори, як правило, містять плазмідний вектор, що містить цис-діючі послідовності, необхідні для Т-ДНК переміщення (як, наприклад, ліву границю й праву границю), маркер селекції, що сконструйований так, щоб бути здатним до експресії в рослинній клітині, і ген, який представляє інтерес (ген, сконструйований так, щоб бути здатним до експресії в рослинній клітині, для якої бажане утворення трансгенних рослин). Також на цьому плазмідному векторі присутні послідовності, необхідні для реплікації в бактерії. Цис-діючі послідовності влаштовані таким чином, щоб забезпечувати ефективну передачу в рослинні клітини й експресію в них. Наприклад, ген маркера селекції й ген, що представляє інтерес, розташовані між лівою й правою границями. Часто другий плазмідний вектор містить транс-діючі фактори, які опосередковують Т-ДНК передачу від *Agrobacterium* до рослинних клітин. Ця плазміда часто містить фактори вірулентності (Vir-гени), які дозволяють інфікувати рослинні клітини за допомогою *Agrobacterium*, і перенесення ДНК за допомогою розрізування в послідовностях границі й ві-опосередкований перенос ДНК, як це розуміють в галузі техніки (Hellens and Mullineaux (2000) Trends in Plant Science, 5:446-451). Кілька типів штамів *Agrobacterium* (наприклад, LBA4404, GV3101, EHA101, EHA105, і т.д.), можуть бути використані для трансформації рослини. Другий плазмідний вектор не є необхідним для трансформації рослин

за допомогою інших способів, таких як мікропроекування, мікроін'єкція, електропорація, поліетиленгліколь, і т.д.

Трансформація рослини

Способи за винаходом стосуються введення нуклеотидної конструкції в рослину. Під "введенням" мають на увазі надання рослині нуклеотидної конструкції таким чином, що конструкція одержує доступ усередину клітини рослини. Способи за винаходом не вимагають використання конкретного способу введення рослині нуклеотидної конструкції, якщо нуклеотидна конструкція одержує доступ принаймні усередину однієї клітини рослини. Способи введення нуклеотидних конструкцій у рослини відомі в галузі техніки, включаючи, але не обмежуючись способами стабільної трансформації, способами тимчасової трансформації й вірус-опосередкованими способами.

Загалом, способи трансформації рослин стосуються перенесення гетерологічної ДНК у рослинні клітини-мішені (наприклад, незрілі або зрілі ембріони, суспензійні культури, недиференційований калюс, протопласти, і т.д.), після чого за допомогою застосування максимального граничного рівня підходящої селекції (залежно від гена маркера селекції й у цьому випадку "гліфосату") відновлюють трансформовані рослинні клітини із групи нетрансформованої клітинної маси. Експлантати, як правило, переносяться для свіжого доступу того ж самого середовища й культивуються звичайним способом. Згодом, трансформовані клітини диференціюються в пагони після переміщення в середовище для регенерації, доповнене максимальним граничним рівнем речовини селекції (наприклад, "гліфосат"). Пагони потім переносять у селективне середовище для вирощування рослин для відновлення корневих пагонів або саджанців. Трансгенний саджанець далі виростає в зрілі рослини й робить родючі насіння (наприклад, Hiei et al. (1994) The Plant Journal 6:271-282; Ishida et al. (1996) Nature Biotechnology 14:745-750). Експлантати, як правило, переносять свіжими в те ж саме середовище й культивують звичайним способом. Загальний опис методів і способів одержання трансгенних рослин знайдено в Ayres and Park (1994) Critical Reviews in Plant Science 13:219-239 і Bommineni and Jauhar (1997) Maydica 42:107-120. Оскільки трансформований матеріал містить багато клітин; і трансформовані й нетрансформовані клітини присутні в будь-якій частині обробленого калюсу-мішені або тканини, або групи клітин. Здатність знищувати нетрансформовані клітини й дозволяти трансформованим клітинам розмножуватися дає культури трансформованих рослин. Часто, здатність видалити нетрансформовані клітини є обмеженням для швидкого відновлення трансформованих рослинних клітин і успішного утворення трансгенних рослин. Для підтвердження присутності інтегрованого гетерологічного гена, що представляє інтерес, у геномі трансгенної рослини можуть бути застосовані молекулярні й біохімічні способи.

Одержання трансгенних рослин може бути виконане за допомогою одного з декількох способів, включаючи, але не обмежуючись введенням гетерологічної ДНК за допомогою *Agrobacterium* у рос-

линні клітини (*Agrobacterium*-опосередкована трансформація), бомбардування рослинних клітин гетерологічною чужорідною ДНК, прикріпленою до частинок, і різними іншими спрямованими способами без частинок (наприклад, Hiei et al. (1994) *The Plant Journal* 6:271-282; Ishida et al. (1996) *Nature Biotechnology* 14:745-750; Ayres and Park (1994) *Critical Reviews in Plant Science* 13:219-239; Bommineni and Jauhar (1997) *Maydica* 42:107-120), для перенесення ДНК.

Способи трансформації хлоропластів добре відомі в галузі техніки. Див., наприклад, Svab et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-606. Спосіб заснований на доставці за допомогою "рушниць" ДНК-частинок, що містять маркер селекції, що адресує ДНК до геному пластид за допомогою гомологічної рекомбінації. Крім цього, трансформація пластид може бути досягнута за допомогою трансактивації мовчазного трансгена в пластиді за допомогою тканино-переважної експресії РНК полімерази, кодованої в ядрі й спрямованої в пластиди. Така система була згадана в McBride et al. (1994) *Proc. Natl. Acad. Sci. USA* 90:5873-5877.

Клітини, які були трансформовані, можуть бути вирощені в рослині відповідно до загальноприйнятих способів. Див., наприклад, McCormick et al. (1986) *Plant Cell Reports* 5:81-84. Ці рослини потім можуть бути вирощені, і опилені або тим же самим трансформованим штамом, або відмінними штамми, і отриманий гібрид має конститутивну експресію бажаної ідентифікованої фенотипової особливості. Два або декілька поколінь можуть бути вирощені для впевненості в тому, що експресія бажаної фенотипової особливості стабільно підтримується й успадковується, а потім збирають насіння для підтвердження, що експресія бажаної фенотипової особливості була досягнута. Таким чином, даний винахід надає трансформовані насіння (які також називаються "трансгенне насіння"), які мають нуклеотидну конструкцію за винаходом, наприклад, касету для експресії за винаходом, стабільно включену в їх геном.

Вимірювання EPSPS-активності

В одному варіанті здійснення даного винаходу EPSPS-фермент стійкості до гліфосату має K_m для фосфоснолпірувату (PEP) між приблизно 1 і приблизно 150 мкМ, включаючи приблизно 2 мкМ, приблизно 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130 або приблизно 140 мкМ, і K_i (гліфосат)/ K_m (PEP) між приблизно 500 і приблизно 1000, приблизно 550, приблизно 600, 650, 700, 750, 800, 850, 900, 950 або до приблизно 1000. Як використовуються в даному описі, K_m і K_i виміряні в умовах, у яких фермент підкоряється кінетиці Міхаеліса-Ментена, рН приблизно 7. В одній необмежувальній методиці вимірювання використовуються фермент в очищеному вигляді в хлориді калію й буфері HEPES при рН 7 при кімнатній температурі й використовуються концентрація гліфосату від 0 до 10 мМ.

Кінетична активність EPSPS може бути виміряна, наприклад, за допомогою вимірювання вивільнення фосфату, який утворюється під час ката-

лізу субстрату EPSPS (наприклад, PEP і S3P) з одержанням наступного продукту реакції (наприклад, 5-снолпірувіл-3-фосфошикімової кислоти), використовуючи флуоресцентний тест, описаний Vazquez et al. (2003) *Anal. Biochem.* 320 (2):292-298. Цей тест заснований на окисленні нефлуоресцентної сполуки N-ацетил-3,7-дигідроксифеноксацину (Amplex® Red, Invitrogen, Carlsbad, CA) у флуоресцентну сполуку резорурфину за допомогою перекису водню (Zhou і Panchuk-Voloshina (1997) *Anal. Biochem.* 253:169-174). Реакція ґрунтується на використанні фосфату пурин-нуклеозидфосфорилази (PNP), ксантиноксидази (XOD) і пероксидази хрому (HRP). Вивільнення фосфату зв'язане з рівнем флуоресценції, що утворюється при перетворенні Amplex® Red у резорурфин. Флуоресценцію можна виміряти, наприклад, використовуючи флуориметровий фільтр, спектрофотометр для зчитування планшетів, спектрофлуориметр, спектрофотометр або подібні, застосовуючи способи, добре відомі в галузі техніки. Флуоресценція, яка утворилася за допомогою реакції, може бути виявлена, застосовуючи флуориметричну установку зі збудженням у діапазоні від приблизно 530 до приблизно 560 нм і емісією приблизно 590 нм. Поглинання можна вимірювати (наприклад, застосовуючи спектрофотометр або спектрофотометр для зчитування планшетів) приблизно при 565 нм.

В одному варіанті здійснення даний винахід охоплює зміну раніше вказаних умов тесту для розширення динамічного діапазону тесту для пристосування до більш широкого діапазону концентрацій субстрату. Зміна включає концентрацію XOD принаймні 1 О/мл, приблизно 1 до приблизно 1,25 О/мл, від приблизно 1,25 до приблизно 1,5 О/мл, від приблизно 1,5 до приблизно 2 О/мл, або більше ніж 2 О/мл; концентрацію PNP більше ніж 0,1 О/мл, від приблизно 0,1 до приблизно 0,5 О/мл, від приблизно 0,5 до приблизно 1 О/мл, від приблизно 1 до приблизно 1,5 О/мл, від приблизно 1,5 до приблизно 2 О/мл, або більше ніж 2 О/мл; і концентрацію Amplex® Red більше ніж 100 мкМ, від приблизно 100 до приблизно 200 мкМ, від приблизно 200 до приблизно 300 мкМ, від приблизно 300 до приблизно 400 мкМ, від приблизно 400 до приблизно 500 мкМ, від приблизно 500 до приблизно 600 мкМ, від приблизно 700 до приблизно 800 мкМ, від приблизно 800 до приблизно 900 мкМ, від приблизно 900 до приблизно 1000 мкМ, або більше ніж приблизно 1000 мкМ. Ця модифікація може бути застосована до тестів, що вимірюють кінетичну активність будь-якого ферменту, у якому фосфат вивільняється під час реакції, яка каталізується за допомогою ферменту.

Рослини

Даний винахід може бути використаний для трансформації будь-яких різновидів рослин, включаючи, але не обмежуючись однодольними й дводольними. Приклади рослин, що цікавлять, включають, але не обмежені зерновими (кукурудза), сорго, пшеницею, сояшником, томатом, хрестоцвітими, перцями, картоплею, бавовником, рисом, соєю, цукровим буряком, цукровим очеретом, тютюном, ячменем, і олійними культурами, Brassica sp., люцерною, житом, просом, сафлором, арахі-

сом, бататом, касавою, кавою, кокосовим горіхом, ананасом, цитрусовими деревами, какао, чаєм, бананом, авокадо, інжиром, гуавою, манго, маслинами, папайєю, горіхом кешью, маकाдамією, мигдалем, вівсом, овочами, декоративними рослинами й хвойними.

Овочі включають, але не обмежені, томатами, салатом, зеленими бобами, лімськими бобами, горохом, і членами роду *Cucumis*, такими як огірок, мускусна диня й диня мускусу. Декоративні рослини включають, але не обмежені азалією, гортензією, гібікусом, трояндами, тюльпанами, нарцисами, петуніями, гвоздиками, молочаєм і хризантемою. Переважно, рослинами за даним винаходом є хлібні злаки (наприклад, кукурудза, сорго, пшениця, соняшник, помідор, хрестоцвіт, перці, картопля, бавовник, рис, соя, цукровий буряк, цукровий очерет, тютюн, ячмінь, олійні культури, і т.д.).

Цей винахід є особливо підходящим для будь-якого члена однодольного сімейства рослин, включаючи, але не обмежуючись кукурудзою, рисом, ячменем, вівсом, пшеницею, сорго, житом, цукровим очеретом, ананасом, бататом, цибулею, бананом, кокосовим горіхом і фініками.

Оцінка трансформації рослини

Після введення гетерологічної чужорідної ДНК у рослинні клітини, трансформація або інтеграція гетерологічного гена в геном рослини підтверджується за допомогою різних способів, таких як аналіз нуклеїнових кислот, білків і метаболітів, зв'язаних із вбудованим геном.

ПЛР-аналіз є швидким методом для скринінгу трансформованих клітин, тканини або пагонів на наявність введеного гена на більш ранній стадії перед пересаджуванням у ґрунт (Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)). ПЛР виконується з використанням олігонуклеотидних праймерів, специфічних для гена, що представляє інтерес, або для вектора *Agrobacterium* як фон, і т.д.

Трансформація рослини може бути підтверджена за допомогою Саузерн-блот аналізу геномної ДНК (Sambrook and Russell (2001) вище). Загалом, загальна ДНК екстрагується із трансформанту, розщеплюється відповідними рестрикуючими ферментами, фракціонується в агарозному гелі й переноситься на нейлонову мембрану або нітроцелюлозу. Мембрана або "блот" можуть далі бути досліджені, наприклад, фрагментом ДНК-мішені, міченим радіоактивним ізотопом ^{32}P , для підтвердження інтеграції введеного гена в геном рослини відповідно до стандартних методик (Sambrook і Russell, 2001, вище).

При Нозерн-аналізі з певних тканин трансформанту виділяють РНК, фракціонують в агарозному формальдегідному гелі й переносять на нейлоновий фільтр відповідно до стандартних методик, які звичайно використовуються в галузі техніки (Sambrook and Russell (2001) вище). За допомогою гібридизації фільтра з радіоактивним зондом, отриманим з GDC за допомогою способів, відомих в галузі техніки, перевіряють експресію РНК, яка кодується *grg23* або *grg51* (Sambrook and Russell (2001) вище).

Для визначення наявності білка, кодованого геном стійкості до гербіциду за допомогою стандартних методик, на трансгенних рослинах можуть бути виконані вестерн-блотинг і біохімічні тести й т.п. (Sambrook and Russell (2001) вище), застосовуючи антитіла, які зв'язуються з одним або декількома епітопами, присутніми в білку стійкості до гербіциду.

Наступні приклади пропонуються як ілюстрація, а не як обмеження.

Експериментальна частина

Приклад 1: Одержання ATX213 0 8

ATX 21308 був виділений за допомогою висівання зразків ґрунту на Збагачене Мінімальне Живильне 3 середовище (EMM3), що містить фосфати й 50 мМ гліфосату. Оскільки EMM3 не містить ніяких ароматичних амінокислот, штам повинен бути стійким до гліфосату, щоб рости на цьому живильному середовищі.

Приблизно один грам ґрунту був розчинений приблизно в 10 мл води й перемішаний на вортексі протягом 5 секунд. 100 мкл цієї суспензії додані до 1 мл EMM3 з фосфатом, але без гліфосату. EMM3 містить (на літр, pH 7,0): 10 г сахарози, 1 г NH_4Cl , 0,2 г $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,01 г $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,007 г $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ та 10 мл розчину фосфату, що містить (на літр, pH 7,0) 210 г Na_2HPO_4 та 90 г NaH_2PO_4 . Культуру збовтують в барабані для тканинних культур, що крутиться, при 21°C протягом ночі та потім висівають на агар EMM3, що містить 50 мМ гліфосату. Через три дні ізолят висівають на агар Luria Bertani (LB) для підтвердження єдності морфології. Після шести днів одичну колонію переносять на агар EMM3, що містить 50 мМ гліфосату. Ізолят вирощують протягом ночі на чашках з 50 мМ гліфосатом. Один специфічний штам, позначений як ATX21308, був вибраний внаслідок його здатності до росту в присутності високих концентрацій гліфосату. Цей штам перевірений на здатність к росту в присутності гліфосату в рідкій культурі та здатний до вирощування при приблизно 300 мМ гліфосату при відпрацьованих умовах.

Приклад 2. Підготовка й скринування космідних бібліотек

Повна ДНК була одержана з культури ATX21308 з використанням способів, широко відомих в галузі техніки. ДНК була частково розщеплена ферментом рестрикції *Sau3A1* і лігована з векторним фрагментом SuperCos (Stratagene) відповідно до вказівок виробника. Продукти літування були упаковані у фагові частинки з використанням пакувального екстракту GigaPack XL (Stratagene), трансфіковані в клітини *E. coli* і висіяні на агарі LB, що містить 50 мкг/мл канаміцину для відбору колоній, що містять косміди.

Індивідуальні колонії були переколені в 384-лункові планшети, що містять бульйон LB і 50 мкг/мл канаміцину, і вирощувалися до насичення. Клітини цих культур були розведені 1:10, потім прикріплені на агарові чашки M63, що містять 50 мкг/мл канаміцину й або 0 мМ, 10 мМ, 20 мМ або 50 мМ гліфосату. [Агарове середовище M63 100 мМ KH_2PO_4 , 15 мМ $(\text{NH}_4)_2\text{SO}_4$, 50 мМ CaCl_2 , 1 мМ FeSO_4 , 50 мМ MgCl_2 , 55 мМ глюкози, 25 мг/літр L-проліну, 10 мг/л тіаміну HCl, достатньо NaOH для доведення pH до 7,0, і 15 г/л агару]. Трансформа-

нти з більш швидким ростом при більш високих концентраціях гліфосату, були ізольовані й розщеплені рестриктуємим ферментом *EcoRI* для ідентифікації космід з подібним малюнком рестрикції. Були ідентифіковані декілька клонів, які ростуть у присутності гліфосату й мають подібні *EcoRI* малюнки рестрикції. Один із цих космідних клонів, *рAX1924*, був вибраний для подальших експериментів.

Приклад 3. Ідентифікація *grg23* у косміді *рAX1924*

Для ідентифікації гена(ів), відповідального за стійкість до гліфосату, продемонстровану космідною *рAX1924*, ДНК із цього клону мутагенізували рухомими елементами. У цьому способі ідентифікували клони, які піддалися інсерціям транспозону й втратили здатність надавати стійкість до гліфосату. Місце розташування інсерцій транспозону встановлює відкриту рамку зчитування, відповідає лъну за фенотип стійкості до гліфосату.

Косміда *рAX1924 in vitro* була піддана мутагенезу транспозоном з використанням *EZ::TN Insertion Kit* (Epicentre, Madison, WI) відповідно до протоколу виробника. За цим способом фрагмент транспозону рандомізовано вставляється в космідну ДНК і в такий спосіб рандомізовано руйнує функцію генів у косміді. Цей транспозон містить ген, що кодує стійкості до триметоприму, таким чином, клони з інсерцією транспозону можуть бути вибрані за допомогою здатності до росту в присутності цього антибіотика. Місця розташування інсерцій транспозону можуть бути визначені за допомогою складання генетичної карти фрагмента рестрикції або за допомогою секвенування із праймерів, які відпалюються на транспозоні. Клон *рAX1924* із вставками транспозону висівали на живильне середовище M63, що містить гліфосат. Було ідентифіковано багато клонів, що містять транспозон, які загубили здатність до росту в присутності гліфосату, вказуючи, що транспозон зруйнував ген, відповідальний за стійкість.

Послідовність ДНК визначали для ділянки *рAX1924*, що містить інсерції транспозону, застосовуючи способи секвенування, добре відомі в галузі техніки. Використовуючи цю інформацію про послідовності, синтезували ДНК-праймери й використовували для визначення послідовності ДНК *рAX1924* ділянки, що вміщає інсерції транспозону. Аналіз отриманих послідовностей ДНК показує, що ця ділянка містить зв'язувальний ген. Цей ген позначений у даному описі як *grg23*. Аналіз *grg23* показує, що він здатний до утворення двох можливих білків у бактеріальних клітинах через наявність потенційних додаткових ділянок старту трансляції. Перший ORF (ORF1) починається зі старт-кодону GTG у положеннях 109-111 з SEQ ID NO:1, і закінчується стоп-кодом TAG біля нуклеотидів 1417-1419 з SEQ ID NO:1. Другий ORF (ORF2) починається зі старт-кодону ATG біля нуклеотидів 178-180 з SEQ ID NO:1, і закінчується стоп-кодомом TAG біля нуклеотидів 1417-1419 з SEQ ID NO:1. Трансляція ORF1 приводить до амінокислотної послідовності, наведеної в SEQ ID NO:2. Трансляція ORF2 приводить до амінокислотної послідовності, наведеної в SEQ ID NO:4.

Аналіз ділянки ДНК, що оточує *grg23*, припускає, що ORF2 передує ділянці зв'язування рибосоми, тоді як не очевидно, що ділянка зв'язування рибосоми передує початку трансляції ORF1. Крім того, вирівнювання обох відкритих рамок зчитування з репрезентативними EPSPS-ферментами показує, що деякі EPSPS-ферменти містять дане N-кінцеве подовження, кодоване в межах ORF1. Таким чином, функціональний ORF, кодований *grg23* у бактеріях, є ORF2. Тому, як використовуються в даному описі, GRG23 стосується того, що кодується ORF2 (нуклеотиди 178-1419 з SEQ ID NO:1). Проте, відомо в галузі техніки, що EPSPS-ферменти є досить толерантними до додаткових амінокислот на їх N-кінці. Тому, експресія ORF1 (нуклеотиди 109-1419 з SEQ ID NO:1) повинна також приводити до EPSPS, що має стійкість до гліфосату.

Для перевірки здатності ORF2 функціонувати як EPSPS і мати стійкість до гліфосату в клітині, ця відкрита рамка зчитування може бути субклонована й експресована в *E. coli*.

Приклад 4. Субклонування *grg23* у вектори для експресії в *E. coli*.

Ген, що кодує GRG23 ORF2 (який починається з ATG (положення 178-180 з SEQ ID NO:1), який експресує білок з 413 амінокислот), був субклонований у *рUC18* і *рRSF1b*, застосовуючи ту ж саму стратегію клонування, викладену загалом вище. Був синтезований ПЛП-праймер [5' CAGGGATCCGGCATGGAACTGATCGACTAGTG 3'], що додає сайт *BamHI*, за яким розташований GGC (5'-GGATCCGGC-3') безпосередньо з 5' від старт-ділянки. Був синтезований 2-ий праймер [5'ATTGGCGCGCCCTAGCCGGGGAGCGTAAG 3'], що додавав сайт *Ascl* безпосередньо з 3' від стоп-послідовності (5'-GGCGCGCC-3'). Кодувальну ділянку *grg23*, ампліфікували за допомогою ПЛП, застосовуючи PFUULTRA™ ДНК-полімеразу (Stratagene). Після ПЛП-ампліфікації *grg23*, застосовуючи дані праймери й розщеплення рестриктазами *BamHI/Ascl*, ПЛП-продукт був лігований у *рUC19* (розщеплений *BamHI* і *Ascl*) і *рRSF1b* (розщеплений *BamHI* і *Ascl*), і були отримані колонії, що містять вставку. Клон *рUC18-grg23* (позначений у даному описі як *рAX1927*) був підтверджений за допомогою розрізування рестриктазами й за допомогою секвенування ДНК.

Подібним чином, вектор для експресії *рAX1909* був розщеплений *BamHI* і *Ascl*, і вектор, що містить фрагмент, був виділений з гелю за допомогою способів, широко відомих в галузі техніки. *рAX1909* є похідним *PRSF-1b* (Novagen, San-Diego, CA), змінений так, щоб містити ділянку *BamHI* безпосередньо з 3' ділянки, що кодує збагачений гістидином "His-Tag". Таким чином, білки, клоновані в *рAX1909*, є гібридами в одній рамці, які містять додаткові амінокислоти MANHHHHHSGS. Вектори, такі як *рAX1909*, як правило, розроблені для експресії й очищення білків, і ці способи широко відомі в галузі техніки.

Розщеплений ПЛП-продукт, отриманий вище, був лігований у розщеплений вектор *рAX1909*, і були отримані колонії, що містять вставку. Клон *рAX1909-grg23* (позначений у даному описі як *рAX1926*) був підтверджений за допомогою розщеп-

лення рестриктазами й за допомогою секвенування ДНК. Властивість конструкції pAX1926 така, що передбачений продукт трансляції GRG23 містить амінокінцеве подовження, що складається з МАННННННН. Це N-кінцеве подовження включає 'гістидинову мітку' або мітку 'six-His', яка застосовується для полегшення очищення білка GRG23, що також широко відомо в галузі техніки.

Плазміда pAX1926, що містить grg23 ORF2, була депонована в Колекції культур служби сільськогосподарських досліджень (NRRL) 18 листопада 2005, і їй присвоєний номер доступу NRRL B-30888.

Приклад 5. grs23 надає стійкість до високих рівнів гліфосату

Конструкцію pUC18-Grg23 (pAX1927) трансформували в *E. coli* штам DH5 α і висівали на чашки з агаром LB, доповнені карбеніциліном (0,1 мг/мл). Були вибрані дві колонії, ресуспендовані в стерилізованій воді й висіяні на чашки M63, що містять або 0 мМ, 25 мМ, 50 мМ або 100 мМ гліфосату. Ізопропіл- β -D-тіогалактопіранозид (IPTG; 0,1 мМ), був також доданий на чашки. Як контроль, клітини, що містять тільки один вектор pUC18, були трансформовані й висіяні на чашки із гліфосатом. Після 2 днів росту, ці чашки були перевірені на ріст (Таблиця 1).

Таблиця 1

Конструкція	0 мМ гліфосату	25 мМ гліфосату	50 мМ гліфосату	100 мМ гліфосату
pUC18	+	-	-	-
pUC18-Grg23 (pAX1927)	+	++	++	++

Цей результат підтверджує, що експресія grg23 з утворенням GRG23-ORF2 надає стійкості до гліфосату в *E. coli* до принаймні 100 мМ. Додатково, ріст *E. coli*, що містять pAX1927, більше виражений у присутності гліфосату, ніж за відсутності гліфосату.

Приклад 6. Гомологія GRG23 з іншими білками
Виведена амінокислотна послідовність GRG23 має гомологію з EPSPS-ферментами, вказуючи, що grg23 кодує EPSPS.

Вивчення виведеної амінокислотної послідовності GRG-ORF2 (SEQ ID NO:4) показує, що в ній не міститься чотирьох доменів, типових для II Класу EPSPS-ферментів. У такий спосіб вона є новою, що не належить до класу II, EPSPS стійкості до гліфосату.

Пошук публічно доступних білкових баз даних, таких як SWISSPROT, показує, що в GRG23 є амінокислотна подібність із широким класом EPSPS-ферментів. Однак, жоден білок у всіх базах даних не має більше ніж 50%-ної гомології з амінокислотною послідовністю GRG23. Репрезентативне вирівнювання GRG23 з іншими EPSPS-ферментами показано на Фігурі 1.

Приклад 7. Очищення GRG23

Конструкцію pRSF1b-grg23 (pAX1926) експресували в *E. coli* після індукції IPTG і очищали в одну стадію, застосовуючи кобальтову хроматографічну колонку, як відомо в галузі техніки. Після елюції з колонки, очищений GRG23 був діалізований проти 50 мМ HEPES/100 мМ KCl, pH 7,0. Білок був більше ніж на 95% чистий за оцінкою за допомогою PAGE. Кількість GRG23 була визначена кількісно, застосовуючи спосіб Бредфорда, як це відомо в галузі техніки (Bradford (1976) Anal. Biochem. 72:248-254).

Приклад 8 Кінетичні тести на активність GRG23

Зразки очищених білків аналізували на активність EPSPS, використовуючи кінетичний тест, що включає інкубацію PEP (Sigma, St. Louis, MO) і S3P у буфері, що містить хлорид калію й HEPES при pH 7.0. Вивільнення фосфату детектували, використовуючи подвійний тест для флуоресцентної детекції фосфату на основі утворення Amplex Red,

як це відомо в галузі техніки (Vazquez et al., (2003) Anal. Biochem. 320:292-298). Оpubліковані умови аналізу можуть привести до насичення аналізованого в експериментах, у яких фосфат вивільняється дуже швидко. Це насичення трохи обмежує динамічний діапазон аналізованого, і вимагає певного діапазону концентрацій ферменту. Було визначено, що кінетичне обмеження флуоресцентного тесту для фосфату очевидно через комбінацію факторів, включаючи обмеження інозину й PNP. У даному винаході умови тесту були розроблені таким чином, що привели до значного поліпшення динамічного діапазону й дозволяють застосовувати більш широкий діапазон концентрацій ферменту й субстрату. Умови тесту, які були значно змінені, включають концентрації пуриннуклеозидфосфорилази (PNP), ксантиноксидаза (XOD), AMPLEX® Red і інозину, для кожного з яких були збільшені концентрації в тесті для пристосування до більш високої швидкості обороту фосфату. Цей тест був застосований для вимірювання активності EPSPS в 96-лунковому форматі з наступними поліпшеннями:

Таблиця 2

Поліпшений флуоресцентний тест

	Поліпшений тест	Vazquez et al., 2003	Одиниці
XOD	1	,4	МО/мл
PNP	2	,02	МО/мл
Інозин	2,25	1,5	мМ
HRP	1	1	МО/мл
Amp Red	1,100	50	мкМ
Hepes	26,25	-	мМ
KCl	26,25	-	мМ
pH	7	7,4	
Tris	-	50	мМ

Ферментативні тести проводилися в 96-лункових планшетах у загальному об'ємі 50 мкл. Реакції проводилися при кімнатній температурі при pH 7,0. Усе компоненти тесту крім PEP,

EPSPS і S3P були об'єднані в Master Mix і розфасовані у окремі пробірки в 96-лунковий планшет, застосовуючи багатоканальну піпетку. У кожен лунку потім були додані підходящі концентрації PEP. Свіжі розведення EPSPS були приготувані й додані у відповідні лунки. Кожний тест починали за допомогою додавання S3P.

Дані зі швидкості були нанесені на діаграму, і кінетичні параметри K_m і K_{cat} були визначені за допомогою застосування Рівняння Міхаеліса-Ментена, застосовуючи програму для обробки нелінійної кривої (KALEID AGRAPH®, Synergy Software). Значення K_i були визначені за допомогою вимірювання K_m (app) при множинних концентраціях гліфосату, і була складена діаграма K_m (app) як функція концентрації інгібітора.

Таблиця 3

Вплив гліфосату на K_m (app) GRG23

Концентрація гліфосату (мкМ)	K_m (app)
0	10,95
3000	18,89
5000	20,67
1000	25,23

За допомогою побудови графіка K_m (app) як функції концентрації гліфосату, може бути отримане лінійне представлення стійкості GRG23 до гліфосату. Точка перетину отриманої лінії з віссю X є $-K_i$. Побудова цієї лінії за даними з Таблиці 3 приводить до наступних значень:

Таблиця 4

Кінетичні значення для GRG23

Фермент	K_m (мкМ)	K_i (мкМ)	K_{cat} (s^{-1})	Відношення K_i/K_m
Grg23	10,95	9525	8,2	869

GRG23 дуже стійкий до гліфосату, з K_i більше ніж 9 мМ, і відношенням K_i/K_m більше ніж 800.

Приклад 9: Виділення ATX21308

Для штаму ATX 21313 приблизно один грам ґрунту суспендували в 10 мл води, і були використані 100 мкл інокуляції 1 мл культури в мінеральних живильних солей середовище А (MSMA) без гліфосату. MSMA містить (на 1 літр, рН 7,0) 1 г NH_4Cl , 10 г сахарози, 0,2 г $MgSO_4 \cdot 7H_2O$, 0,01 г $FeSO_4 \cdot 7H_2O$, 0,007 г $MnSO_4 \cdot H_2O$, доповнене фосфатами. Після нічної інкубації культура була висіяна на тверде середовище, що містить MSMA та 50 мМ гліфосату, інкубована протягом декількох днів, та інокульована на агарові чашки Luria Bertani для підтвердження єдиного типу колоній. Ріст в присутності 50 мМ гліфосату був рекомендований після повторного вирощування на MSMA, 50 мМ агарових чашках. Цей спосіб виділення привів до штаму ATX21313, який був здатний добре рости в таких умовах.

Приклад 10. Клонування стійких до гліфосату EPSP-синтаз

Геномна ДНК була отримана зі штаму ATX21313, і отримана ДНК частково була розщеплена ферментом рестрикції *Sau3A1* для одержання фрагментів ДНК із розміром приблизно 5 тисяч пар нуклеотидів. Ці молекули ДНК були відібрані за розміром на агарозному гелі, виділені й ліговані в "плечі" LAMBDA ZAP® вектора, який попередньо розщеплювали *Bam*HI. Ліговані "плечі" далі були упаковані у фагові частинки, і були визначені титри бактеріофага, як відомо в галузі техніки. Отримані бібліотеки ампліфікували за допомогою способів, відомих в галузі техніки, для одержання бібліотечного титру між 3×10^7 і 3×10^8 БУО/мл. Для кожної незалежної бібліотеки, *E. coli* (XL1 Blue MRF') була потім котрансфікована бактеріофагом з ампліфікованої бібліотеки так само, як і M13 фагом-помічником для забезпечення масового вирізування бібліотеки у вигляді інфекційної, кільцевої олдНК, як відомо в галузі техніки (Short et al. (1988) *Nucleic Acids Research* 16:7583-7600). Після центрифугування коінфікованих клітин, супернатант, що містить бактеріофаг, нагрівали до 65-70°C протягом 15-20 хвилин для інактивації залишкових часток фага лямбди. Розведення отриманої олдНК плазмідної бібліотеки трансфікували у свіжу культуру компетентних *E. coli* клітин XL1 Blue MRF', а також клітин XL-Blue MRF' (Δ aro) (XL1 Blue MRF'). Отримані трансфіковані клітини висівали на чашки M63, що містять канаміцин, 0,1 мМ IPTG і або 0 мМ, 20 мМ або 50 мМ гліфосату. Цей спосіб скринування дозволяє ідентифікувати клони, що містять толерантні до гліфосату EPSP-синтази, так само, як і клони, що несуть толерантність до гліфосату. Колонії, що ростуть на 20 мМ або 50 мМ гліфосату в штамі Δ aro або XL-Blue MRF', були переколені, і їх плазміді проаналізовані за допомогою рестриктивного розрізування для ідентифікації плазмід із загальним малюнком рестрикції. Окремі плазміді секвенували за допомогою способів, відомих в галузі техніки, з перевагою, відданою плазмідам, які проявляли стійкість до 50 мМ гліфосату.

Використовуючи цей підхід, який іноді модифікували для кожної бібліотеки, як відомо й цінується в галузі техніки, були пізнані клони бібліотеки, що містять гени EPSP-синтази.

Послідовності ділянок отриманих клонів були ідентифіковані як ділянки EPSP-синтази.

Приклад 11. ДНК і білкова послідовності EPSP-синтаз

Послідовність ДНК EPSP-синтази, стійкої до гліфосату, була визначена для рAX1967 за допомогою способів, відомих в галузі техніки. Послідовність ДНК grg51 представлена в даному описі як SEQ ID NO:5. Передбачений продукт трансляції grg51 (GRG51) представлений у даному описі як SEQ ID NO:6. GRG51 демонструє 97% амінокислотну ідентичність із GRG23 (SEQ ID NO:2).

Плазміда рAX1926, що містить grg23 ORF2, була депонована в Колекції культур служби сільськогосподарських досліджень (NRRL) 18 листопада 2005, і їй присвоєний номер доступу NRRL B-30949.

Таблиця 5 підсумує гомологію GRG23 і GRG51 з іншими EPSP-синтазними ферментами.

Таблиця 5

Амінокислотна ідентичність GRG23-ORF1 і GRG51 з репрезентативними Фермент-EPSPS-ферментам

EPSPS	% ідентичності з GRG23	% ідентичності з GRG51
GRG23	-	92%
GRG51	92%	-
B. Clausii	36%	35%
R. xylanophilus	39%	38%
E. coli	32%	32%
CP4	20%	21%
Zea_maze	32%	32%

Приклад 12. Клонування нових стійких до гліфосату EPSP-синтаз в E. coli вектор для експресії

Ген grg51, що міститься в рAX1967, був субклований в E. coli вектор для експресії pRSF1b (Invitrogen). Отримані клони були підтверджені за допомогою ДНК секвенування, і застосовані для індукції експресії grg51 в E. coli. Експресований His-теговий білок був потім очищений, як відомо в галузі техніки.

Приклад 13. Стійкість до гліфосату EPSP-синтаз

Клітини, що містять рAX1967, були висіяні на M63 + чашки, що містять антибіотик і або 0 мМ або 20 мМ гліфосату. Піст був оцінений після двох днів вирощування при 37°C. Спостерігали, що GRG51 надавав клітинам E. coli стійкість до 20 мМ гліфосату (Таблиця 6).

Таблиця 6

Гліфосатний скринінг

EPSPS	Плазмідний клон	Піст на 20 мМ гліфосату
Вектор	-	-
GRG51	рAX1967	++

Приклад 14. Конструювання syngrg23 і експресія

Послідовність нового гена, що кодує білок GRG23 (SEQ ID NO:2; Заявка на Патент США № 60/741166, зареєстрована 1 грудня 2005), була сконструйована й синтезована. Ця послідовність представлена як SEQ ID NO:12. Ця відкрита рамка зчитування, сконструйована в даному описі як "syngrg23", була клонована у вектор для експресії pRSF1b (Invitrogen) за допомогою способів, відомих в галузі техніки.

Ген Syngrg23, що кодує GRG23, був клонований у вектор pUC19 для створення рAX748. Для ампліфікації syngrg23 з рAX748 з використанням MUTAZYME® II системи (Stratagene) для введення випадкових мутацій в syngrg23, що кодує ділянку, були застосовані ПЛР-праймери, які фланкували syngrg23 у цьому векторі. Матриця була розведена 1:50 у підданий помилкам реакції ПЛР, і ампліфікація виконувалася протягом 30 циклів. Отриманий ПЛР-продукт був розщеплений ферментами рестрикції BstHI і SgsI, очищений від гелю й лігований

у вектор pRSF1b для створення бібліотеки мutowаних syngrg23.

Бібліотеки мutowаних syngrg23 були трансформовані в E. coli штам BL21*DE3 star (Invitrogen). Після трансформації окремі колонії були висіяні на середовище 1× M63, що містить 150 мМ гліфосату для відбору клонів, які зберегли ферментативну активність і стійкість росту.

Приклад 15. Скринування на стійкість до гліфосату на чашках

Бібліотечні літування були трансформовані в компетентні клітини E. coli BL21*DE3 (Invitrogen). Трансформації були виконані відповідно до інструкцій виробника з наступними модифікаціями. Після інкубації протягом 1 години при 37°C у середовищі SOC, клітини були осаджені центрифугуванням (5 хвилин, 1000×g, 4°C). Клітини промивали 1 мл M63+, знову центрифугували, і видаляли супернатант. Клітини промивали в другий раз 1 мл M63+ і повторно ресуспендували в 200 мкл M63+.

Для відбору мутантних ферментів GRG23, що надають стійкість до гліфосату в E. coli, клітини були висіяні на агарові чашки із середовищем M63+, що містять 150 мМ гліфосату, 0,05 мМ IPTG (ізопропіл-бета-D-тіогалактопіранозид) і 50 мкг/мл канаміцину. Середовище M63+ 100 мМ KH₂PO₄, 15 мМ (NH₄)₂SO₄, 50 мкМ CaCl₂, 1 мкМ FeSO₄, 50 мкМ MgCl₂, 55 мМ глюкози, 25 мг/літр L-проліну, 10 мг/л тіаміну HCl, достатньо NaOH для доведення pH до 7,0, і 15 г/л агару. Чашки інкубували протягом 36 годин при 37°C.

Окремі колонії були переколені, висіяні в 384-лункові планшети. Таким чином, були створені два 384-лункових планшети. Третій планшет з 384 клонами був зібраний з колоній, які росли на чашках без гліфосату.

Приклад 16. Ізоляція й аналіз варіантів GRG23, стійких до гліфосату

Клітини BL21*DE3, трансформовані мутагенізованими syngrg23 і/або варіантами grg23, були ідентифіковані за допомогою росту на чашках із гліфосатом. Екстракти мутагенізованих syngrg23 і варіантів grgr23 були приготовлені й протестовані на поліпшення ферментативної активності. Колонії, ідентифіковані на чашках із гліфосатом, були прикріплені в 96-лункові блоки, що містять середовище LB, і були вирощені до O.D. приблизно 0,6. Потім був доданий IPTG (0,5 мМ), і блоки інкубували протягом ночі при 20°C для індукції експресії білка. Екстракти білка були приготовлені з осадів клітин, застосовуючи реактив для культур POP (Novagen) і Lysonase (Novagen), і ферментативна активність у неочищених лізатах була виміряна після нагрівання екстрактів протягом 30 хвилин при 37°C. Екстракти з активністю, більшою, ніж два стандартні відхилення вище за середнє з ряду екстрактів, що містять відповідний контрольний білок (наприклад, GRG23), були вибрані для подальшого аналізу.

Клони, що показують підвищену активність після інкубації як неочищені екстракти, були вирощені в 250 мл культури LB, і експресія білка була викликана IPTG. Після індукції мутантний білок GRG23 був очищений з кожної культури за допомогою афінної хроматографії з використанням кобальтової смоли (Novagen). Далі очищені білки

були перевірені на ферментативну активність після нагрівання протягом 0, 2, 4 і приблизно 16 годин при 37°C.

Приклад 17. Поліпшені варіанти GRG23

З бібліотеки ДНК мутагенізованого *syngrg23* було ідентифіковано декілька клонів з поліпшеною активністю. Були визначені послідовності ДНК клонів, що відповідають цим екстрактам. У Таблиці 7 показані амінокислотні заміни, ідентифіковані в шести варіантах GRG23, які зберігали стійкість до гліфосату: *grg23* (L3P1.B20) (SEQ ID NO:26), що кодує амінокислотну послідовність GRG23 (L3P1.B20) (SEQ ID NO:27); *grg23* (L3P1.B3) (SEQ ID NO:28) що кодує амінокислотну послідовність GRG23 (L3P1.B3) (SEQ ID NO:29); *grg23* (L3P1.F18) (SEQ ID NO:30), що кодує амінокислотну послідовність GRG23 (L3P1.F18) (SEQ ID NO:31); і *grg23* (L3P1.023) SEQ ID NO:31 SEQ, що кодує амінокислотну послідовність GRG23 (L3P1.023) SEQ ID NO:32).

Таблиця 7

Мутації, ідентифіковані
в стійкі до гліфосату варіантах GRG23

Клон	Амінокислота (AA) в GRG23
L3P1B20	V206→I
L3P1B3	D75→H, E217→K
L3P1F18	T274→I
L3P1O23	R5→H

Клони були вирощені в 250 мл культури LB, і була викликана експресія білка, і білок був виділений, як описано вище. Очищені білки потім були протестовані на ферментативну активність після нагрівання протягом 0, 2 і 4 і приблизно 16 годин при 37°C. Виявилось, один із клонів, позначений "M5", мав збільшене співвідношення його ферментативної активності після тривалої інкубації при 37°C (Таблиця 8). Була визначена послідовність ДНК даного клону, і ген визначається в даному описі як *grg23* (*ace1*) (SEQ ID NO:14). Білок, експресований з *grg23* (*ace1*), позначений GRG23 (ACE1) (SEQ ID NO:15).

Таблиця 8

Час напівжиття GRG23 (ACE1)
у порівнянні з GRG23 при підвищеній температурі

Білок	Час напівжиття при 37°C (год)
GRG23	7
GRG23 (ACE1)	15,5

GRG23 (ACE1) містить 2 амінокислотні заміни в порівнянні з білком GRG23 дикого типу: A49→T і S276→T. Вектор pRSF1b, що містить цей ген, позначений pAX3801. На фігурі 1 представлена відносна стійкість GRG23 (ACE1) у порівнянні з GRG23 при підвищених температурах.

Приклад 18. Визначення EPSPS-активності варіантів GRG-23

Екстракти, що містять варіантні білки GRG23, були протестовані на синтазну активність EPSP, як

описано в Заявці на Патент США № 60/741166, зареєстрованої 1 грудня 2005, включеної в даний опис повністю за допомогою посилання. Тестування були проведені в сумарному об'ємі 50 мкл, що містить 0,5 мМ шикимат-3-фосфату, 200 мМ фосфоенолпірувату (PEP), 1 О/мл ксантиноксидази, 2 О/мл нуклеозидфосфорилази, 2,25 мМ інозину, 1 О/мл пероксидази хроні, 0-2 мМ гліфосату, 50 мМ HEPES/KOH pH 7,0, 100 мМ KCl і AMPLEX® Red (Invitrogen) відповідно до інструкцій виробника. Екстракти інкубували з усіма компонентами тесту крім шикимат-3-фосфату протягом 5 хвилин при кімнатній температурі, і тестування було почато за допомогою додавання шикимат-3-фосфату. EPSP-синтазна активність була виміряна з використанням флуоресцентного спектрометра Spectramax Gemini XPS (Molecular Dynamics, збудження: 555 нм; емісія: 590 нм).

Дотримуючись повного визначення, кінетичні параметри були виконані на очищеному білку, як описано раніше (Заявка на Патент США № 60/741166, зареєстрована 1 грудня 2005), доводячи кількість білка, визначену за допомогою методу Бредфорда, як відомо в галузі техніки. Для будь-якої однієї концентрації гліфосату, EPSP-синтазна активність була виміряна як функція широкого діапазону концентрацій PEP. Дані укладалися в рівняння Міхаеліса-Ментена з використанням програмного забезпечення KALEIDAGRAPH® (Synergy Software), для визначення K_m (гадана K_m) EPSP-синтази при тій концентрації гліфосату. Уявні значення K_m були визначені при не менше ніж 4 концентраціях гліфосату, і K_i EPSPS для гліфосату були обчислені за графіком уявних K_m у порівнянні з концентрацією гліфосату, використовуючи рівняння ($m_1 \cdot x / (m_2 + x)$; $m_1=1$; $m_2=1$), як відомо в галузі техніки.

Таблиця 9

Кінетика GRG23 (ACE1) у порівнянні з GRG23

	K_m (мкМ)	K_i (мкМ)	V_{max} (нМ/хв/мкг)
GRG23	12,2	13800	14,77
GRG23 (ACE1)	9,7	14620	13,73

Приклад 19. Ідентифікація *grg23* (*ace2*).

GRG23 (ACE1) містить дві амінокислотні заміни в порівнянні з GRG23. Щоб визначити, чи могли додаткові заміни в цих положеннях далі поліпшити активність, була отримана ДНК бібліотека, що привела до клонів, які експресують білки, які були переважно мутовані в обох положеннях 49 і 276 GRG23. Клони, що несуть стійкість до гліфосату, були вибрані за допомогою росту на чашках із гліфосатом, і вирощені й протестовані на кінетичні властивості, як було описано.

Дивно, один клон, позначений у даному описі *grg23* (*ace2*) (SEQ ID NO:16), що кодує GRG23 (ACE2) білок (SEQ ID NO:17), був ідентифікований як такий, що має поліпшену термостабільність. Послідовність ДНК *grg23* (*ace2*) показала, що GRG23 (ACE2) містить єдину амінокислотну заміну (залишок 276 GRG23 на аргінін).

Приклад 20. Порівняння GRG23 і GRG51, і мутагенез різних залишків

Для оцінки пермутацій амінокислотних послідовностей, імовірних з порівняння амінокислотних послідовностей GRG23 і GRG51, були отримані дві бібліотеки. Перша бібліотека вводила розмаїтість із амінокислотної послідовності GRG51 в кодувальну ділянку grg23 (ace2). Друга бібліотека вводила розмаїтість із амінокислотної послідовності GRG23 (ACE2) у кодувальну ділянку grg51.

Клони отриманих бібліотек були протестовані на (1) здатність надавати клітині стійкість до гліфосату, і (2) активність після тривалої інкубації при 37°C. У цілому десять клонів були секвеновані й проаналізовані більш докладно. Один конкретний клон, який визначається у даному описі як grg51.4 (SEQ ID NO:18), що кодує білок GRG51.4 (SEQ ID NO:19), містить декілька амінокислотних замін і відносно GRG23 (ACE2) і відносно GRG51. Амінокислотні заміни, які присутні в GRG51.4 відносно GRG23 (ACE2), були згодом введені в ген grg23 (ace2) для одержання grg23 (ace3) (SEQ ID NO:20), що кодує білок GRG23 (ACE3) (SEQ ID NO:21). GRG23 (ACE3) проявляє найвищу активність і термостабільність у порівнянні з GRG23 і GRG23 (ACE2).

GRG23 (ace1) був мутагенізований, і клони були протестовані для ідентифікації клонів, які експресують варіанти з поліпшеною термостабільністю та/або активністю. Один клон, grg23 (L5P2.J2) (SEQ ID NO:22), що кодує GRG23 (L5P2.J2) (SEQ ID NO:23), був ідентифікований на основі його поліпшених кінетичних властивостей. GRG23 (L5P2.J2), містить три амінокислотні заміни відносно GRG23 (ACE1), як показано в наступній Таблиці 10.

Таблиця 10

Амінокислотні заміни в GRG23 (L5P2.J2)

Амінокислота (AA) в GRG23 (L5P2.J2) відносно GRG23 (ACE1)
V101→F
A213→S
D284→N

Для створення клонів, які містять кожну ідентифіковану амінокислотну заміну GRG23 (L5P2.J2) в кодувальній ділянці grg23 (ace3) був застосований мутагенез олігонуклеотидами. Клон був ідентифікований як клон, що кодує білок, який має поліпшені кінетичні властивості в порівнянні з GRG23 (ACE3), і позначений як grg23 (ace4) (SEQ ID NO:24). Білок, кодований grg23 (ace4), позначений як GRG23 (ACE4) (SEQ ID NO:25), містить єдину амінокислотну заміну відносно GRG23 (ACE3) (Валін 101 на фенілаланін). На основі цих результатів, був виконаний окремий мутагенез олігонуклеотидами, щоб проаналізувати кінетику кожних можливих амінокислотних замін у положенні 101. Жодна з амінокислотних замін не привела до додаткового поліпшення кінетичних властивостей у порівнянні з GRG23 (ACE4).

Кінетика поліпшених варіантів

	K _m (мкМ)	K _i (мкМ)	V max (нМ/хв/мкг)
GRG23	14	10800	13
GRG51	15	21048	13
GRG23 (ACE1)	10	14620	14
GRG23 (ACE2)	11	18104	15
GRG51.4	19	26610	17
GRG23 (ACE3)	15	20000	17
GRG23 (L5P2.J2)	15	2500	23
GRG23 (ACE1)	14	5010	24

Приклад 21. Конструювання grg23 або grg51 для трансформації рослини

Відкрита рамка зчитування (ORF) grg23 або grg51 була ампліфікована за допомогою ПЛР із повнорозмірної кДНК матриці. Рестриктні сайти HindIII були додані до кожного кінця ORF під час ПЛР. Додатково, нуклеотидна послідовність АСС була додана відразу з 5' від старт-кодону гена для збільшення ефективності трансляції (Kozak (1987) Nucleic Acid Research 15:8125-8148; Joshi (1987) Nucleic Acid Research 15:6643-6653). Продукт ПЛР був клонований і секвенований, застосовуючи способи, відомі в галузі техніки, для підтвердження того, що ніякі мутації не введені під час ПЛР.

Плазміда, що містить grg23 або grg51 продукт ПЛР, була розщеплена HindIII, і фрагмент, що містить інтактну ORF, був ізольований. Цей фрагмент був клонований у рестриктні сайти HindIII плазмиди рAX200, вектор для експресії в рослинах, що містить промотор актину рису (McElroy et al. (1991) Molec. Gen. Genet. 231:150-160) і термінатор PinII (An et al. (1989) The Plant Cell 1:115-122). Фрагмент промотор-ген-термінатор із цієї проміжної плазмиди субклонували в плазмиду рSB11 (Japan Tobacco, Inc) для утворення підсумкової плазмиди, наприклад, рSB11GRG23. рSB11GRG23 була організована таким чином, що 3,91 т.п.н. фрагмент ДНК, що містить конструкцію промотор-grg23-термінатор може бути розщеплений за допомогою подвійного розщеплення KpnI і PmeI і використаний для трансформації рослин за допомогою аерозольної променевої ін'єкції. Структура рSB11GRG23 була перевірена за допомогою рестрикційного розрізування й електрофорезу в гелі, і за допомогою секвенування через різні зв'язки для клонування.

Плазміда була перенесена в Agrobacterium tumefaciens, штам LBA4404, що також несе плазмиду рSB1 (Japan Tobacco, Inc), використовуючи процедури трибатьківського схрещування, добре відомі в галузі техніки, і висіваючи в середовище, що містить спектиноміцин. Плазміда рSB11GRG23 несе стійкість до спектиноміцину, але є плазмідом вузького кола хазяїв і не може розмножуватися в Agrobacterium. Колонії, стійкі до спектиноміцину, з'являються, коли рSB11GRG23 вбудовується в плазмиду широкого кола хазяїв рSB1 за допомогою гомологічної рекомбінації. Продукт коінтеграції рSB1 і рSB11GRG23 перевірений за допомогою гібридизації за Саузерном. Штам Agrobacterium,

що несе коінтеграції, використовуються для трансформації кукурудзи за допомогою способу PureIntro (Japan Tobacco).

Приклад 22. Трансформація grs23 або grg51 у рослинні клітини

Качани кукурудзи збирали через 8-12 днів після запилення. З качанів виділяли ембріони, і дані ембріони, розміром 0,8-1,5 мм, застосовували для трансформації. Ембріони висівали на чашки стороною шийного щитка наверх у підходящому середовищі для інкубації, такому як середовище DN62A5S (3,98 г/л Солфі N6; 1 мл/л (1000х базового розчину) Вітамінів N6; 800 мг/л L-аспарагіну; 100 мг/л Муо-інозитулу 1,4 г/л L-проліну; 100 мг/л казамінових кислот; 50 г/л сахарози; 1 мл/л (1 мг/мл базового розчину) 2,4-D). Однак, середовища й солі, відмінні від DN62A5S, є підходящими й загальновідомі в галузі техніки. Ембріони інкубували протягом ночі при 25°C у темряві.

Отримані експлантати переносили у квадратні лунки (30-40 на чашці), переносили в осмотичні середовища протягом 30-45 хвилин, потім переносили на променеву чашку (див., наприклад, Публікацію PCT № WO/0138514 і Патент США № 5240842).

ДНК-конструкції, сконструйовані для експресії GRG23 у рослинних клітинах, переносили в ткани-

ні рослини з використанням каталізатора променю аерозолі, застосовуючи по суті умови, як описано в Публікації PCT № WO/0138514. Після опромінення ембріони інкубували протягом 30 хвилин в осмотичних середовищах і поміщали в інкубаційні середовища на ніч при 25°C у темряві. Для того, щоб уникнути надмірного ушкодження опромінених експлантатів, їх інкубували протягом принаймні 24 годин до переміщення в відновлювальне середовище. Ембріони потім розподіляли в середовище для відновлення протягом 5 днів при 25°C у темряві, потім переносили в селекційне середовище. Експлантати культивували в селективному середовищі аж до восьми тижнів, залежно від природи й особливостей конкретно застосовуваної селекції. Після періоду селекції отриманий каліус переносили на середовище для дозрівання ембріонів доти, поки не спостерігалось утворення зрілих соматичних ембріонів. Отримані зрілі соматичні ембріони далі поміщали під слабе освітлення, і запускали процес регенерації за допомогою способів, відомих в галузі техніки. Отриманим пагонам дозволяли вкоренитися в середовищі для вирощування рослин, і отримані рослини переносили в горщики розплідника й розмножували як трансгенні рослини.

Матеріали

Середовище DN62A5S

Компоненти	На літр	Джерело
Chus's N6 суміш основних солей (Prod. No. C 426)	3,98 г/л	Phytotechnology Labs
Chus's N6 розчин вітаміну (Prod. No. C 149)	1 мл/л (з 1000х стоку)	Phytotechnology Labs
L-Аспарагін	800 мг/л	Phytotechnology Labs
Муо-інозитол	100 мг/л	Sigma
L-пролін	1,4 г/л	Phytotechnology Labs
Казамінові кислоти	100 мг/л	Fisher Scientific
Сахароза	50 г/л	Phytotechnology Labs
2, 4-D (Prod. No. D-7299)	1 мл/л (з 1 мг/мл стоку)	Sigma

Доводять рН розчину до рН 5,8 IN KOH/1N KCl, додають Gelrite (Sigma) до 3 г/л і автоклавуєть. Після охолодження до 50°C, доводять до 2 мл/л 5 м/мл базовим розчином нітрату срібла (Phytotechnology Labs). Рецепт розрахований приблизно на 20 чашок.

Приклад 23. Трансформація grs23 або grg51 у рослинні клітини кукурудзи за допомогою Agrobacterium-опосередкованої трансформації.

Качани збирали через 8-12 днів після запилення. З качанів виділяли ембріони, і ці ембріони, розміром 0,8-1,5 мм, застосовували для трансформації. Ембріони висівали на чашки стороною шийного щитка наверх у підходящому середовищі для інкубації протягом ночі при 25°C у темряві. Однак, по суті не обов'язково культивувати ембріони протягом ночі. Ембріони, які контактують зі штамом Agrobacterium, що містить відповідні вектори для плазмиди Ti-опосередкованої передачі протягом 5-10 хвилин, потім висівають на чашки із середовищем для спільного культивування протягом 3 днів (при 25°C у темряві). Після спільного культивування експлантати переносили в середовище для періоду відновлення протягом п'яти днів (в 25°C у темряві). Експлантати інкубу-

вали на селективному середовищі аж до восьми тижнів залежно від природи й особливостей конкретного способу селекції.

Після періоду селекції, отриманий каліус переносили в середовище для дозрівання ембріона доти, поки не спостерігається утворення зрілих соматичних ембріонів. Отримані зрілі соматичні ембріони далі поміщали під слабе освітлення, і запускали процес регенерації за допомогою способів, відомих в галузі техніки. Отриманим пагонам дозволяли вкоренитися в середовищі для вирощування рослин, і отримані рослини переносили в горщики розплідника й розмножували як трансгенні рослини.

Всі публікації й заявки на патенти, згадані в описі, указують на рівень досвіду фахівців в галузі техніки, до якої належить даний винахід. Всі публікації й заявки на патенти включені в даний опис за допомогою посилання в тому ступені, як якщо конкретно й окремо вказано, що кожна окрема публікація або заявка на патент включені за допомогою посилання.

Хоча попередній винахід був описаний докладно за допомогою ілюстрації й прикладу для переважного розуміння, очевидно, що певні зміни

й модифікації можуть бути здійснені в рамках

доданої формули винаходу.

,

Перелік послідовностей

<110> Cheryl Peters
Jill Burdette
Philip E. Hammer
Brian Vande Berg
Laura Cooper Schouten
Brian Carr

<120> Гени GRG23 та GRG51, які надають стійкість до гербіцидів

<130> 45600/320129

<150> 60/741,166

<151> 2005-12-01

<150> 60/817,799

<151> 2006-06-30

<160> 33

<170> FastSEQ для Windows версія 4,0

<210> 1

<211> 1892

<212> ДНК

<213> *Arthrobacter globiformis*

<220>

<221> misc_feature

<222> (0)...(0)

<223> Strain ATX21308

<221> CDS

<222> (109)...(1419)

<221> misc_feature

<222> 1801

<223> n = A,T,C або G

<221> misc_feature

<222> 1801

<223> n = A,T,C або G

<400> 1

gggaccacat gctgctcctg atttcagggc tgctgccggt atggaccagg gtttagagag 60
ggacggcacg catccgggcc cttatcggac caacgccaac agcggctcg gtg gcc ttg 117
Met Ala Leu
1

gag cgg ggc cag cac ggc cga tca cgt aga ctc ttt gga gct tcg ctc 165
Glu Arg Gly Gln His Gly Arg Ser Arg Arg Leu Phe Gly Ala Ser Leu
5 10 15

gaa agg atc acc atg gaa act gat cga cta gtg atc cca gga tcg aaa 213
Glu Arg Ile Thr Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys
20 25 30 35

agc atc acc aac cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg 261
Ser Ile Thr Asn Arg Ala Leu Leu Leu Ala Ala Ala Ala Lys Gly Thr
40 45 50

43	92622	44	
tcg gtc ctg gtg aga cca ttg gtc agc gcc gat acc tca gca ttc aaa			309
Ser Val Leu Val Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys			
55	60	65	
act gca att cag gcc ctc ggt gcc aac gtc tca gcc gac ggt gac aat			357
Thr Ala Ile Gln Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn			
70	75	80	
tgg gtc gtt gaa ggc ctg ggt cag gca ccc cac ctc gac gcc gac atc			405
Trp Val Val Glu Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile			
85	90	95	
tgg tgc gag gat gca ggt acc gtg gcc cgg ttc ctc cct cca ttc gtc			453
Trp Cys Glu Asp Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val			
100	105	110	115
gcc gca gga cag ggg aag ttc acc gtc gac gga agc gag cag ctg cgg			501
Ala Ala Gly Gln Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg			
120	125	130	
cgg cgc ccg ctt cgg ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc			549
Arg Arg Pro Leu Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala			
135	140	145	
cgc gtc tcc tcc gag cag ctg ccc cta aca att gaa gcg agc ggg ctg			597
Arg Val Ser Ser Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu			
150	155	160	
gca ggc ggg gag tac gaa att gaa gcc cat cag agc agc cag ttc gcc			645
Ala Gly Gly Glu Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala			
165	170	175	
tcc ggc ctg atc atg gcc gcc ccg tac gcg cga caa ggc ctg cgt gtg			693
Ser Gly Leu Ile Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val			
180	185	190	195
cgg ata cca aat ccc gtg agc cag ccc tac ctc acg atg aca ctg cgg			741
Arg Ile Pro Asn Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg			
200	205	210	
atg atg agg gac ttc ggc ctt gag acc agc acc gac gga gcc acc gtc			789
Met Met Arg Asp Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val			
215	220	225	
agc gtc cct ccc ggg cgc tac aca gcc cgg cgg tat gaa att gaa ccg			837
Ser Val Pro Pro Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro			
230	235	240	
gac gcg tca act gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc			885
Asp Ala Ser Thr Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly			
245	250	255	
cga agc ttc gaa ttc cag ggc ctt ggc aca gac agc atc caa ggc gac			933
Arg Ser Phe Glu Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp			
260	265	270	275
acg tca ttc ttc aat gta ctt ggg cgg ctc ggt gca gag gtc cac tgg			981
Thr Ser Phe Phe Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp			
280	285	290	
gca ccc aac tcg gtc acc ata tcc gga ccg gaa agg ctg aac ggc gac			1029
Ala Pro Asn Ser Val Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp			
295	300	305	

45	92622	46	
att gaa gtg gat atg ggc gag ata tgc gac acc ttc atg aca ctc gcg			1077
Ile Glu Val Asp Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala			
310	315	320	
gcg att gcc cct cta gcc gat gga ccc atc acg ata acc aac att gcc			1125
Ala Ile Ala Pro Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly			
325	330	335	
cat gca cgg ttg aag gaa tcc gac cgc atc tgc gcg atg gaa acc aac			1173
His Ala Arg Leu Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn			
340	345	350	355
ctg cga acg ctc ggt gta caa acc gac gtc gga cac gac tgg atg cga			1221
Leu Arg Thr Leu Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg			
360	365	370	
atc tac ccc tct acc ccg cac ggc ggc aga gtc aat tgc cac cgg gac			1269
Ile Tyr Pro Ser Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp			
375	380	385	
cac agg atc gcc atg gcg ttt tca atc ctg gga ctg cga gtg gac ggg			1317
His Arg Ile Ala Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly			
390	395	400	
att acc ctc gac gac cct caa tgt gtc ggg aag acc ttt cct ggc ttc			1365
Ile Thr Leu Asp Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe			
405	410	415	
ttc gac tac ctt gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc			1413
Phe Asp Tyr Leu Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro			
420	425	430	435
ggc tag tgacttcctc tccggcggac gctaggcatc ggaaaacgaa tcctgacatg			1469
Gly *			

accgacctcc tcgcgtcacg gcgtgtctgc cggtacccaa gcattctgcc ttagccgctt 1529
ccgcggcccc ttatgcttcc tggttgtcca gattttcatc cgggatgttg cctgaccttg 1589
agcagggcaa tcagctgttc agcactgtca atggtgtggg cctgaaggc ggcttcgatg 1649
gctgccacgt cggcggctct catcgctgtc acgacacgca gatgcgcttc ataggcacgt 1709
tcaggatccg ccctcgtcgc ctgatcctga gccaaaggcaa tagttaagt tagcctcogtt 1769
ggcggccaga gccgaagcaa taaggagttt tncgaggcca ccagattcc ccgggtggaa 1829
ggcgatatgg gcttcatgct gaactatggg gtccggatgg aagtgaacttt tcaactctgc 1889
cca 1892

<210> 2

<211> 436

<212> PRT

<213> Arthrobacter globiformis

<400> 2

Met	Ala	Leu	Glu	Arg	Gly	Gln	His	Gly	Arg	Ser	Arg	Arg	Leu	Phe	Gly
1				5				10						15	
Ala	Ser	Leu	Glu	Arg	Ile	Thr	Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro
		20						25				30			
Gly	Ser	Lys	Ser	Ile	Thr	Asn	Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala
		35				40					45				
Lys	Gly	Thr	Ser	Val	Leu	Val	Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser
		50				55					60				
Ala	Phe	Lys	Thr	Ala	Ile	Gln	Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp
65				70					75					80	
Gly	Asp	Asn	Trp	Val	Val	Glu	Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp
		85					90						95		
Ala	Asp	Ile	Trp	Cys	Glu	Asp	Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro

47										92622										48																													
100										105										110																													
Pro	Phe	Val	Ala	Ala	Gly	Gln	Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Pro	Phe	Val	Ala	Ala	Gly	Gln	Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Pro	Phe	Val	Ala	Ala	Gly	Gln	Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu		
115										120										125																													
Gln	Leu	Arg	Arg	Arg	Pro	Leu	Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Gln	Leu	Arg	Arg	Arg	Pro	Leu	Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Gln	Leu	Arg	Arg	Arg	Pro	Leu	Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His		
130										135										140																													
Leu	Gly	Ala	Arg	Val	Ser	Ser	Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Leu	Gly	Ala	Arg	Val	Ser	Ser	Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Leu	Gly	Ala	Arg	Val	Ser	Ser	Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala		
145										150										155																													
Ser	Gly	Leu	Ala	Gly	Gly	Glu	Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Ser	Gly	Leu	Ala	Gly	Gly	Glu	Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Ser	Gly	Leu	Ala	Gly	Gly	Glu	Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser		
165										170										175																													
Gln	Phe	Ala	Ser	Gly	Leu	Ile	Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Gln	Phe	Ala	Ser	Gly	Leu	Ile	Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Gln	Phe	Ala	Ser	Gly	Leu	Ile	Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly		
180										185										190																													
Leu	Arg	Val	Arg	Ile	Pro	Asn	Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Leu	Arg	Val	Arg	Ile	Pro	Asn	Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Leu	Arg	Val	Arg	Ile	Pro	Asn	Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met		
195										200										205																													
Thr	Leu	Arg	Met	Met	Arg	Asp	Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Thr	Leu	Arg	Met	Met	Arg	Asp	Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Thr	Leu	Arg	Met	Met	Arg	Asp	Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly		
210										215										220																													
Ala	Thr	Val	Ser	Val	Pro	Pro	Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ala	Thr	Val	Ser	Val	Pro	Pro	Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ala	Thr	Val	Ser	Val	Pro	Pro	Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu		
225										230										235																													
Ile	Glu	Pro	Asp	Ala	Ser	Thr	Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Ile	Glu	Pro	Asp	Ala	Ser	Thr	Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Ile	Glu	Pro	Asp	Ala	Ser	Thr	Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala		
245										250										255																													
Val	Ser	Gly	Arg	Ser	Phe	Glu	Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Val	Ser	Gly	Arg	Ser	Phe	Glu	Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Val	Ser	Gly	Arg	Ser	Phe	Glu	Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile		
260										265										270																													
Gln	Gly	Asp	Thr	Ser	Phe	Phe	Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Gln	Gly	Asp	Thr	Ser	Phe	Phe	Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Gln	Gly	Asp	Thr	Ser	Phe	Phe	Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu		
275										280										285																													
Val	His	Trp	Ala	Pro	Asn	Ser	Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Val	His	Trp	Ala	Pro	Asn	Ser	Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Val	His	Trp	Ala	Pro	Asn	Ser	Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu		
290										295										300																													
Asn	Gly	Asp	Ile	Glu	Val	Asp	Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Asn	Gly	Asp	Ile	Glu	Val	Asp	Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Asn	Gly	Asp	Ile	Glu	Val	Asp	Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met		
305										310										315																													
Thr	Leu	Ala	Ala	Ile	Ala	Pro	Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Thr	Leu	Ala	Ala	Ile	Ala	Pro	Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Thr	Leu	Ala	Ala	Ile	Ala	Pro	Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr		
325										330										335																													
Asn	Ile	Gly	His	Ala	Arg	Leu	Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Asn	Ile	Gly	His	Ala	Arg	Leu	Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Asn	Ile	Gly	His	Ala	Arg	Leu	Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met		
340										345										350																													
Glu	Thr	Asn	Leu	Arg	Thr	Leu	Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Glu	Thr	Asn	Leu	Arg	Thr	Leu	Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Glu	Thr	Asn	Leu	Arg	Thr	Leu	Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp		
355										360										365																													
Trp	Met	Arg	Ile	Tyr	Pro	Ser	Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	Trp	Met	Arg	Ile	Tyr	Pro	Ser	Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	Trp	Met	Arg	Ile	Tyr	Pro	Ser	Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys		
370										375										380																													
His	Arg	Asp	His	Arg	Ile	Ala	Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	His	Arg	Asp	His	Arg	Ile	Ala	Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	His	Arg	Asp	His	Arg	Ile	Ala	Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg		
385										390										395																													
Val	Asp	Gly	Ile	Thr	Leu	Asp	Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Val	Asp	Gly	Ile	Thr	Leu	Asp	Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Val	Asp	Gly	Ile	Thr	Leu	Asp	Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe		
405										410										415																													
Pro	Gly	Phe	Phe	Asp	Tyr	Leu	Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Pro	Gly	Phe	Phe	Asp	Tyr	Leu	Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Pro	Gly	Phe	Phe	Asp	Tyr	Leu	Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu		
420										425										430																													
Thr	Leu	Pro	Gly													Thr	Leu	Pro	Gly														Thr	Leu	Pro	Gly													
435																																																	

<210> 3
 <211> 1892
 <212> ДНК
 <213> Arthrobacter globiformis

<220>
 <221> CDS
 <222> (178)...(1419)

<221> misc_feature
 <222> 1801
 <223> n = A,T,C a60 G

<400> 3
 gggaccacat gctgctcctg atttcagggc tgctgccggt atggaccagg gtttagagag 60
 ggacggcagc catccgggcc ottatcggac caacgccaac agcggtcggt ggccttgag 120
 cggggccagc acggccgac acgtagactc ttggagctt cgctcgaaag gatcacc atg 180
 Met

49	92622	50	
gaa act gat cga cta gtg atc cca gga tgc aaa agc atc acc aac cgg			228
Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn Arg			
5	10	15	
gct ttg ctt ttg gct gcc gca gcg aag ggc acg tgc gtc ctg gtg aga			276
Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val Arg			
20	25	30	
cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca att cag gcc			324
Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln Ala			
35	40	45	
ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa ggc			372
Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu Gly			
50	55	60	65
ctg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gat gca			420
Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp Ala			
70	75	80	
ggt acc gtg gcc cgg ttc ctc cct cca ttc gtc gcc gca gga cag ggg			468
Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln Gly			
85	90	95	
aag ttc acc gtc gac gga agc gag cag ctg cgg cgg cgc ccg ctt cgg			516
Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu Arg			
100	105	110	
ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc gag			564
Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser Glu			
115	120	125	
cag ctg ccc cta aca att gaa gcg agc ggg ctg gca ggc ggg gag tac			612
Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu Tyr			
130	135	140	145
gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc atg			660
Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile Met			
150	155	160	
gcc gcc ccg tac gcg cga caa ggc ctg cgt gtg cgg ata cca aat ccc			708
Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn Pro			
165	170	175	
gtg agc cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac ttc			756
Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp Phe			
180	185	190	
ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct ccc ggg			804
Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro Gly			
195	200	205	
cgc tac aca gcc cgg cgg tat gaa att gaa ccg gac gcg tca act gcg			852
Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr Ala			
210	215	220	225
tgc tac ttc gcc gcc gct tcc gcc gtc tct ggc cga agc ttc gaa ttc			900
Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu Phe			
230	235	240	
cag ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc aat			948
Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe Asn			
245	250	255	

51	92622	52	
gta ctt ggg cgg ctc ggt gca gag gtc cac tgg gca ccc aac tcg gtc			996
Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser Val			
260	265	270	
acc ata tcc gga ccg gaa agg ctg aac ggc gac att gaa gtg gat atg			1044
Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp Met			
275	280	285	
ggc gag ata tcg gac acc ttc atg aca ctc gcg gcg att gcc cct cta			1092
Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro Leu			
290	295	300	305
gcc gat gga ccc atc acg ata acc aac att ggc cat gca cgg ttg aag			1140
Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu Lys			
	310	315	320
gaa tcc gac cgc atc tcg gcg atg gaa acc aac ctg cga acg ctc ggt			1188
Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu Gly			
	325	330	335
gta caa acc gac gtc gga cac gac tgg atg cga atc tac ccc tct acc			1236
Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser Thr			
	340	345	350
ccg cac ggc ggc aga gtc aat tgc cac cgg gac cac agg atc gcc atg			1284
Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala Met			
	355	360	365
gcg ttt tca atc ctg gga ctg cga gtg gac ggg att acc ctc gac gac			1332
Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp Asp			
370	375	380	385
cct caa tgt gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt gga			1380
Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu Gly			
	390	395	400
cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag tgacttcctc			1429
Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			
	405	410	
tccggcggac gctagggcatc ggaaaacgaa tcctgacatg accgacctcc tcgcgtcacg			1489
gcgtgtctgc cggtacccaa gcattctgcc ttagccgctt ccgcggcccc ttatgctttc			1549
tggttggtcca gattttcatc cgggatgttg cctgacctg agcagggcaa tcagctgttc			1609
agcactgtca atgggtgtggg ccctgaaggc ggcttcgatg gctgccacgt cggcggctct			1669
catcgctgtc acgacacgca gatgcgcttc ataggcacgt tcaggatccg ccctcgctgc			1729
ctgatcctga gccaaaggcaa tagttagatg tgcctccgtt ggccggccaga gccgaagcaa			1789
taaggagttt tncgaggcca cccagattcc ccgggtggaa ggcgatatgg gcttcatgct			1849
gaactatggg gtccggatgg aagtgacttt tcaactctgc cca			1892
<210> 4			
<211> 413			
<212> PRT			
<213> Arthrobacter globiformis			
<400> 4			
Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn			
1 5 10 15			
Arg Ala Leu Leu Leu Ala Ala Ala Ala Lys Gly Thr Ser Val Leu Val			
20 25 30			
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln			
35 40 45			
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu			
50 55 60			

53										92622										54										
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp															
65					70					75				80																
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln															
				85					90					95																
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu															
			100					105						110																
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser															
		115					120						125																	
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu															
	130					135					140																			
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile															
145					150					155				160																
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn															
				165					170					175																
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp															
			180					185					190																	
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro															
		195					200						205																	
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr															
	210					215					220																			
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu															
225					230					235				240																
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe															
				245					250					255																
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser															
			260					265					270																	
Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp															
		275					280					285																		
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro															
	290					295					300																			
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu															
305					310					315				320																
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu															
				325					330					335																
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser															
			340					345					350																	
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala															
		355					360					365																		
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp															
	370					375					380																			
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu															
385					390					395				400																
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																		
				405					410																					

<210> 5

<211> 1242

<212> ДНК

<213> Невідомо

<220>

<223> Виділено із ґрунтового зразка

<221> CDS

<222> (1)...(1242)

<400> 5

atg gaa act gat cga cta gtg atc cca gga tcg aaa agc atc acc aac 48

Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn

1

5

10

15

cgg gct ttg ctt ttg gct gcc gca gcg aag ggc gcg tcg gtc ctg gtg 96

Arg Ala Leu Leu Leu Ala Ala Ala Ala Lys Gly Ala Ser Val Leu Val

55	92622	56	
20	25	30	
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca att cag Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln 35 40 45			144
gcc ctc ggt gcc aac gtc tca gcg gac ggt gat gat tgg gtc gtt gaa Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asp Trp Val Val Glu 50 55 60			192
ggc ctg ggc cag gca ccc aac ctc gac gcc gac atc tgg tgc gag gat Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala Asp Ile Trp Cys Glu Asp 65 70 75 80			240
gcc ggt acc gtg gcc cgg ttc ctc cct cca ttc gtc gcc gca gga cag Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln 85 90 95			288
ggg aag ttc acc gtc gac gga agc gag cag ctg cgg cgg cgc ccg ctt Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu 100 105 110			336
cgg ccc gtg gtc gac gcc atc cgc cac ctg gcc gcc cgc gtc tcc tcc Arg Pro Val Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser 115 120 125			384
gag cag ctg ccc cta acg att gaa gcg agc ggg ctg gca gcc ggg gag Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu 130 135 140			432
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggt ctg atc Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile 145 150 155 160			480
atg gcc gcc ccg tac gcg cga caa gcc ctg cgt gtt cgg ata cca aat Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn 165 170 175			528
ccc gtg agc cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp 180 185 190			576
ttc gcc att gag acc agc acc gac gga gcg acc gtc agc gtt cct ccc Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro 195 200 205			624
ggg cgc tac aca gcg cgg cgg tat gag att gaa ccg gac gcg tca act Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr 210 215 220			672
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct gcc cgg cgc ttc gaa Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Arg Phe Glu 225 230 235 240			720
ttc cag gcc ctt gcc aca gac agc atc caa gcc gac acg tca ttc ttc Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe 245 250 255			768
aat gta ctt ggg cgg ctc gcc gca gag gtc cac tgg gca tcc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Ser Asn Ser 260 265 270			816
gtc acc ata tcc gga ccg gaa agg ctg acc gcc gac att gaa gtg gat Val Thr Ile Ser Gly Pro Glu Arg Leu Thr Gly Asp Ile Glu Val Asp			864

57	92622	58
275	280	285
atg ggc gag ata tcg gac acc ttc atg aca ctg gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro 290 295 300		912
cta gcc gat gga ccc atc acg ata aca aac att ggc cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu 305 310 315 320		960
aag gaa tcc gac cgc atc tcg gcg atg gaa agc aac ctt cga atg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Ser Asn Leu Arg Met Leu 325 330 335		1008
ggt gta caa acc gac gtc gga cac gac tgg atg cga atc tac ccc tct Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350		1056
acc ccg cac ggc ggc aga gtc aat tgc cac cgg gac cac agg atc gcc Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365		1104
atg gcg ttt tca atc ctg gga ctg cga gtg gac ggg att acc ctc gac Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 370 375 380		1152
gac cct caa tgt gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu 385 390 395 400		1200
gga cgc ctt ttc ccg gaa aag gcg ctt acg ctc ccc ggc tag Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *		1242

<210> 6

<211> 413

<212> PRT

<213> Невідомо

<220>

<223> Виділено із ґрунтового зразка

<400> 6

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

59					92622					60				
145					150				155				160	
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	
				165					170				175	
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	
			180					185				190	Arg	
Phe	Gly	Ile	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	
		195					200				205	Pro	Pro	
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	
	210					215				220		Ser	Thr	
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Arg	
225					230					235		Phe	Glu	
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	
			245					250				255	Phe	
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Ser	
		260						265				270	Asn	
Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Thr	Gly	Asp	Ile	Glu	
		275				280						285	Val	
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	
	290					295				300		Ala	Pro	
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	
305					310					315		Arg	Leu	
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Ser	Asn	Leu	Arg	
			325						330			335	Met	
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	
		340					345					350	Pro	
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	
		355				360					365	Ile	Ala	
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	
	370				375					380		Leu	Asp	
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	
385					390					395		Tyr	Leu	
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly		
			405					410						

<210> 7
 <211> 446
 <212> PRT
 <213> *Bacillus clausii*

<400> 7
 Met Val Gln Phe Asp Ser Gln Ala Arg Ser Pro Trp Thr Pro Leu Ala
 1 5 10 15
 Gly Val Glu Arg Leu Arg Leu Thr Pro Ser Gln Lys Arg Ile Asn Ala
 20 25 30
 Thr Leu Glu Val Pro Gly Ser Lys Ser Ala Thr Asn Arg Ala Leu Leu
 35 40 45
 Leu Ala Ala Val Ala Ser Gly Thr Ser Thr Leu Arg Asn Ala Leu Lys
 50 55 60
 Ser Asp Asp Thr Tyr Trp Cys Ile Glu Ala Leu Lys Lys Thr Gly Val
 65 70 75 80
 Glu Ile Ala Val Asp Gly Ser Asn Val Thr Val Tyr Gly Arg Gly Gly
 85 90 95
 Val Phe His Ser Gly Ser Leu Tyr Ile Gly Ser Ala Gly Thr Ala Gly
 100 105 110
 Arg Phe Leu Pro Gly Met Leu Ala Ala Thr Gly Asn Trp His Val
 115 120 125
 Glu Ala Ser His Ser Met Asn Lys Arg Pro Ile Ala Pro Leu Val Lys
 130 135 140
 Thr Leu Gln Ala Leu Gly Ala Asn Ile Gln Tyr Gly Ser Arg Arg Gly
 145 150 155 160
 His Tyr Pro Leu Ser Ile Ser Gly Glu Gly Leu Asn Gly Gly Lys Val
 165 170 175
 Asn Met Ser Gly Gln Leu Ser Ser Gln Phe Ile Ser Gly Cys Leu Leu

61										92622										62									
180										185										190									
Ala	Ala	Pro	Leu	Ala	Lys	Asn	Pro	Val	Ser	Ile	Thr	Val	Lys	Asp	Gly														
		195						200						205															
Ile	Val	Gln	Gln	Ala	Tyr	Val	Arg	Ile	Thr	Ile	Asp	Leu	Met	Ala	Ala														
		210					215							220															
Phe	Gly	Val	Glu	Val	Lys	Ala	Ala	Pro	Asp	Trp	Ser	Leu	Leu	Glu	Val														
225					230				235					240															
Asn	Pro	Ser	Pro	Tyr	Val	Ala	Asn	Asp	Ile	Ala	Ile	Glu	Ala	Asp	Ala														
				245					250					255															
Ser	Thr	Ala	Cys	Tyr	Phe	Leu	Ala	Leu	Ala	Ala	Ile	Thr	Ala	Gly	Lys														
			260					265						270															
Ile	Arg	Ile	Arg	His	Phe	Ser	Thr	Lys	Thr	Ser	Gln	Pro	Asp	Ile	Leu														
		275					280							285															
Phe	Val	Ser	Ile	Leu	Lys	Arg	Met	Gly	Cys	Asn	Phe	Glu	Ile	Gly	Pro														
		290				295								300															
Ser	Phe	Val	Glu	Gly	Glu	Gly	Pro	Thr	Arg	Leu	Arg	Gly	Gly	Phe	Thr														
305					310									320															
Val	Asn	Met	Asn	Glu	Leu	Ser	Asp	Gln	Ala	Leu	Thr	Leu	Ala	Ala	Ile														
				325					330					335															
Ser	Pro	Phe	Ala	Asp	Gly	Pro	Ile	Ala	Ile	Glu	Gly	Val	Gly	His	Ile														
			340					345						350															
Arg	His	His	Glu	Cys	Asp	Arg	Ile	Arg	Ala	Ile	Cys	Thr	Glu	Leu	Ser														
		355					360							365															
Arg	Leu	Gly	Ile	Arg	Val	Glu	Glu	Arg	His	Asp	Gly	Leu	Thr	Val	Tyr														
		370				375								380															
Pro	Gly	Gln	Pro	Lys	Pro	Thr	Val	Val	Asn	Thr	Tyr	Asp	Asp	His	Arg														
385					390									400															
Met	Ala	Met	Ala	Leu	Ala	Leu	Ile	Gly	Ala	Lys	Val	Asp	Gly	Ile	Glu														
				405					410					415															
Leu	Asp	Asp	Pro	Gly	Cys	Val	Ala	Lys	Thr	Cys	Pro	Ser	Tyr	Phe	Ser														
			420					425						430															
Met	Leu	Ala	Gln	Thr	Gly	Ile	Gly	Val	Lys	Ala	Val	Ser	Pro																
		435					440							445															

<210> 8
 <211> 447
 <212> PRT
 <213> Rubrobacter xylophilus

<400> 8
 Met Ser Gly Val Ser Gly Val Pro Gly Val Asp Phe Gly Ile Glu Glu
 1 5 10 15
 Val Arg Gly Ser Phe Pro Glu Glu Met Glu Val Ala Pro Leu Glu Arg
 20 25 30
 Pro Pro Asp Ala Thr Val Arg Leu Pro Gly Ser Lys Ser Ile Thr Asn
 35 40 45
 Arg Ala Leu Leu Val Ala Ala Leu Ala Gly Gly Thr Ser Arg Ile Glu
 50 55 60
 Asn Pro Leu Leu Ala Asp Asp Pro Phe Trp Leu Met Asn Ala Leu Val
 65 70 75 80
 Gly Leu Gly Phe Gly Val Arg Val Gly Glu Gly Ala Val Glu Val
 85 90 95
 Ala Gly Gly Gly Gly Gly Ile Pro Ala Pro Ser Ala Asp Val Phe Val
 100 105 110
 Gly Asn Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Ala Leu Ala Leu
 115 120 125
 Gly Ser Gly Pro Tyr Arg Val Asp Gly Thr Pro Arg Met Arg Glu Arg
 130 135 140
 Pro Val Ala Glu Leu Val Glu Ala Leu Arg Ala Leu Gly Ala Arg Val
 145 150 155 160
 Glu Cys Glu Glu Arg Glu Gly His Leu Pro Leu Val Val Arg Gly Gly
 165 170 175
 Ala Arg Gly Gly Gly Glu Ile Ser Val Ser Gly Glu Arg Ser Ser Gln

63										92622										64									
			180										185										190						
Phe	Leu	Ser	Gly	Leu	Leu	Ile	Ser	Ala	Pro	Cys	Leu	Pro	Gly	Gly	Leu														
		195					200					205																	
Thr	Val	Arg	Pro	Arg	Gly	Ala	Leu	Val	Ser	Arg	Pro	Tyr	Val	Asp	Ile														
	210					215						220																	
Thr	Val	Arg	Val	Met	Arg	Ser	Phe	Gly	Ala	Ser	Val	Glu	Glu	Glu	Pro														
225					230							235			240														
Ser	Gly	Ala	Ala	Phe	Arg	Val	Ala	Pro	Gly	Ala	Tyr	Arg	Ala	Thr	Ala														
				245					250					255															
Tyr	Arg	Val	Glu	Pro	Asp	Ala	Ser	Ala	Ala	Ser	Tyr	Phe	Leu	Ala	Ala														
			260					265					270																
Ala	Ala	Leu	Thr	Ala	Gly	Arg	Val	Ile	Pro	Gly	Leu	Gly	Arg	Ser															
		275				280						285																	
Ser	Leu	Gln	Gly	Asp	Val	Ala	Phe	Ala	Gly	Ile	Leu	Arg	Arg	Met	Gly														
	290					295						300																	
Cys	Arg	Val	Ser	Leu	Ser	Glu	Asp	Arg	Ile	Glu	Leu	Ala	Gly	Pro	Pro														
305					310							315			320														
Arg	Leu	Arg	Gly	Val	Glu	Ala	Asp	Met	Asn	Ala	Ile	Ser	Asp	Thr	Met														
				325					330					335															
Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro	Phe	Ala	Ser	Ser	Pro	Thr	Leu	Ile														
			340					345					350																
Lys	Asn	Val	Ala	His	Thr	Arg	Leu	Gln	Glu	Thr	Asp	Arg	Leu	Ala	Ala														
		355				360						365																	
Val	Ala	Ala	Glu	Leu	Ser	Arg	Leu	Gly	Val	Arg	Val	His	Glu	Thr	Pro														
	370					375						380																	
Asp	Ser	Leu	Arg	Ile	Ile	Pro	Gly	Lys	Val	Arg	Pro	Ala	Ala	Ile	Arg														
385					390							395			400														
Thr	Tyr	Gly	Asp	His	Arg	Met	Ala	Met	Ala	Phe	Ser	Leu	Val	Gly	Leu														
				405					410					415															
Arg	Val	Arg	Gly	Val	Arg	Ile	Leu	Asp	Pro	Gly	Cys	Val	Thr	Lys	Thr														
			420					425					430																
Leu	Pro	Gly	Tyr	Phe	Arg	Leu	Leu	Glu	Gly	Leu	Arg	Arg	Gly	Gly															
		435					440						445																

<210> 9
 <211> 444
 <212> PRT
 <213> Zea mays

<400> 9
 Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile Ser Gly
 1 5 10 15
 Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile Leu Leu
 20 25 30
 Leu Ala Ala Leu Ser Glu Gly Thr Val Val Asp Asn Leu Leu Asn
 35 40 45
 Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu Gly Leu
 50 55 60
 Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val Gly Cys
 65 70 75 80
 Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln Leu Phe
 85 90 95
 Leu Gly Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala Val Thr
 100 105 110
 Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro Arg Met
 115 120 125
 Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln Leu Gly
 130 135 140
 Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val Arg Val
 145 150 155 160
 Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser Gly Ser
 165 170 175
 Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro Leu Ala

65										92622										66									
				180										185										190					
-Leu	Gly	Asp	Val	Glu	Ile	Glu	Ile	Ile	Asp	Lys	Leu	Ile	Ser	Ile	Pro														
		195						200						205															
Tyr	Val	Glu	Met	Thr	Leu	Arg	Leu	Met	Glu	Arg	Phe	Gly	Val	Lys	Ala														
	210						215							220															
Glu	His	Ser	Asp	Ser	Trp	Asp	Arg	Phe	Tyr	Ile	Lys	Gly	Gly	Gln	Lys														
225					230					235				240															
Tyr	Lys	Ser	Pro	Lys	Asn	Ala	Tyr	Val	Glu	Gly	Asp	Ala	Ser	Ser	Ala														
				245						250				255															
Ser	Tyr	Phe	Leu	Ala	Gly	Ala	Ala	Ile	Thr	Gly	Gly	Thr	Val	Thr	Val														
			260							265				270															
Glu	Gly	Cys	Gly	Thr	Thr	Ser	Leu	Gln	Gly	Asp	Val	Lys	Phe	Ala	Glu														
		275					280							285															
Val	Leu	Glu	Met	Met	Gly	Ala	Lys	Val	Thr	Trp	Thr	Glu	Thr	Ser	Val														
	290						295							300															
Thr	Val	Thr	Gly	Pro	Pro	Arg	Glu	Pro	Phe	Gly	Arg	Lys	His	Leu	Lys														
305					310					315				320															
Ala	Ile	Asp	Val	Asn	Met	Asn	Lys	Met	Pro	Asp	Val	Ala	Met	Thr	Leu														
				325						330				335															
Ala	Val	Val	Ala	Leu	Phe	Ala	Asp	Gly	Pro	Thr	Ala	Ile	Arg	Asp	Val														
			340							345				350															
Ala	Ser	Trp	Arg	Val	Lys	Glu	Thr	Glu	Arg	Met	Val	Ala	Ile	Arg	Thr														
		355					360							365															
Glu	Leu	Thr	Lys	Leu	Gly	Ala	Ser	Val	Glu	Glu	Gly	Pro	Asp	Tyr	Cys														
	370						375							380															
Ile	Ile	Thr	Pro	Pro	Glu	Lys	Leu	Asn	Val	Thr	Ala	Ile	Asp	Thr	Tyr														
385					390					395				400															
Asp	Asp	His	Arg	Met	Ala	Met	Ala	Phe	Ser	Leu	Ala	Ala	Cys	Ala	Glu														
				405						410				415															
Val	Pro	Val	Thr	Ile	Arg	Asp	Pro	Gly	Cys	Thr	Arg	Lys	Thr	Phe	Pro														
			420							425				430															
Asp	Tyr	Phe	Asp	Val	Leu	Ser	Thr	Phe	Val	Lys	Asn																		
		435					440																						

<210> 10
 <211> 455
 <212> PRT
 <213> Agrobacterium sp.

<400> 10
 Met Ser His Gly Ala Ser Ser Arg Pro Ala Thr Ala Arg Lys Ser Ser
 1 5 10 15
 Gly Leu Ser Gly Thr Val Arg Ile Pro Gly Asp Lys Ser Ile Ser His
 20 25 30
 Arg Ser Phe Met Phe Gly Gly Leu Ala Ser Gly Glu Thr Arg Ile Thr
 35 40 45
 Gly Leu Leu Glu Gly Glu Asp Val Ile Asn Thr Gly Lys Ala Met Gln
 50 55 60
 Ala Met Gly Ala Arg Ile Arg Lys Glu Gly Asp Thr Trp Ile Ile Asp
 65 70 75 80
 Gly Val Gly Asn Gly Gly Leu Leu Ala Pro Glu Ala Pro Leu Asp Phe
 85 90 95
 Gly Asn Ala Ala Thr Gly Cys Arg Leu Thr Met Gly Leu Val Gly Val
 100 105 110
 Tyr Asp Phe Asp Ser Thr Phe Ile Gly Asp Ala Ser Leu Thr Lys Arg
 115 120 125
 Pro Met Gly Arg Val Leu Asn Pro Leu Arg Glu Met Gly Val Gln Val
 130 135 140
 Lys Ser Glu Asp Gly Asp Arg Leu Pro Val Thr Leu Arg Gly Pro Lys
 145 150 155 160
 Thr Pro Thr Pro Ile Thr Tyr Arg Val Pro Met Ala Ser Ala Gln Val
 165 170 175
 Lys Ser Ala Val Leu Leu Ala Gly Leu Asn Thr Pro Gly Ile Thr Thr

67										92622										68									
			180										185										190						
Val	Ile	Glu	Pro	Ile	Met	Thr	Arg	Asp	His	Thr	Glu	Lys	Met	Leu	Gln														
		195						200				205																	
Gly	Phe	Gly	Ala	Asn	Leu	Thr	Val	Glu	Thr	Asp	Ala	Asp	Gly	Val	Arg														
	210						215					220																	
Thr	Ile	Arg	Leu	Glu	Gly	Arg	Gly	Lys	Leu	Thr	Gly	Gln	Val	Ile	Asp														
225					230						235				240														
Val	Pro	Gly	Asp	Pro	Ser	Ser	Thr	Ala	Phe	Pro	Leu	Val	Ala	Ala	Leu														
			245						250					255															
Leu	Val	Pro	Gly	Ser	Asp	Val	Thr	Ile	Leu	Asn	Val	Leu	Met	Asn	Pro														
		260						265				270																	
Thr	Arg	Thr	Gly	Leu	Ile	Leu	Thr	Leu	Gln	Glu	Met	Gly	Ala	Asp	Ile														
		275					280					285																	
Glu	Val	Ile	Asn	Pro	Arg	Leu	Ala	Gly	Gly	Glu	Asp	Val	Ala	Asp	Leu														
	290						295				300																		
Arg	Val	Arg	Ser	Ser	Thr	Leu	Lys	Gly	Val	Thr	Val	Pro	Glu	Asp	Arg														
305					310						315				320														
Ala	Pro	Ser	Met	Ile	Asp	Glu	Tyr	Pro	Ile	Leu	Ala	Val	Ala	Ala	Ala														
			325					330						335															
Phe	Ala	Glu	Gly	Ala	Thr	Val	Met	Asn	Gly	Leu	Glu	Glu	Leu	Arg	Val														
		340					345							350															
Lys	Glu	Ser	Asp	Arg	Leu	Ser	Ala	Val	Ala	Asn	Gly	Leu	Lys	Leu	Asn														
		355					360					365																	
Gly	Val	Asp	Cys	Asp	Glu	Gly	Glu	Thr	Ser	Leu	Val	Val	Arg	Gly	Arg														
	370					375					380																		
Pro	Asp	Gly	Lys	Gly	Leu	Gly	Asn	Ala	Ser	Gly	Ala	Ala	Val	Ala	Thr														
385					390						395				400														
His	Leu	Asp	His	Arg	Ile	Ala	Met	Ser	Phe	Leu	Val	Met	Gly	Leu	Val														
			405					410						415															
Ser	Glu	Asn	Pro	Val	Thr	Val	Asp	Asp	Ala	Thr	Met	Ile	Ala	Thr	Ser														
		420					425							430															
Phe	Pro	Glu	Phe	Met	Asp	Leu	Met	Ala	Gly	Leu	Gly	Ala	Lys	Ile	Glu														
		435					440					445																	
Leu	Ser	Asp	Thr	Lys	Ala	Ala																							
450						455																							

<210> 11
 <211> 427
 <212> PRT
 <213> E. coli

<400> 11
 Met Glu Ser Leu Thr Leu Gln Pro Ile Ala Arg Val Asp Gly Thr Ile
 1 5 10 15
 Asn Leu Pro Gly Ser Lys Thr Val Ser Asn Arg Ala Leu Leu Ala
 20 25 30
 Ala Leu Ala His Gly Lys Thr Val Leu Thr Asn Leu Leu Asp Ser Asp
 35 40 45
 Asp Val Arg His Met Leu Asn Ala Leu Thr Ala Leu Gly Val Ser Tyr
 50 55 60
 Thr Leu Ser Ala Asp Arg Thr Arg Cys Glu Ile Ile Gly Asn Gly Gly
 65 70 75 80
 Pro Leu His Ala Glu Gly Ala Leu Glu Leu Phe Leu Gly Asn Ala Gly
 85 90 95
 Thr Ala Met Arg Pro Leu Ala Ala Ala Leu Cys Leu Gly Ser Asn Asp
 100 105 110
 Ile Val Leu Thr Gly Glu Pro Arg Met Lys Glu Arg Pro Ile Gly His
 115 120 125
 Leu Val Asp Ala Leu Arg Leu Gly Gly Ala Lys Ile Thr Tyr Leu Glu
 130 135 140
 Gln Glu Asn Tyr Pro Pro Leu Arg Leu Gln Gly Gly Phe Thr Gly Gly
 145 150 155 160
 Asn Val Asp Val Asp Gly Ser Val Ser Ser Gln Phe Leu Thr Ala Leu

69										92622										70									
165										170										175									
Leu	Met	Thr	Ala	Pro	Leu	Ala	Pro	Glu	Asp	Thr	Val	Ile	Arg	Ile	Lys														
180										185										190									
Gly	Asp	Leu	Val	Ser	Lys	Pro	Tyr	Ile	Asp	Ile	Thr	Leu	Asn	Leu	Met														
195										200										205									
Lys	Thr	Phe	Gly	Val	Glu	Ile	Glu	Asn	Gln	His	Tyr	Gln	Gln	Phe	Val														
210										215										220									
Val	Lys	Gly	Gly	Gln	Ser	Tyr	Gln	Ser	Pro	Gly	Thr	Tyr	Leu	Val	Glu														
225										230										235									
Gly	Asp	Ala	Ser	Ser	Ala	Ser	Tyr	Phe	Leu	Ala	Ala	Ala	Ala	Ile	Lys														
245										250										255									
Gly	Gly	Thr	Val	Lys	Val	Thr	Gly	Ile	Gly	Arg	Asn	Ser	Met	Gln	Gly														
260										265										270									
Asp	Ile	Arg	Phe	Ala	Asp	Val	Leu	Glu	Lys	Met	Gly	Ala	Thr	Ile	Cys														
275										280										285									
Trp	Gly	Asp	Asp	Tyr	Ile	Ser	Cys	Thr	Arg	Gly	Glu	Leu	Asn	Ala	Ile														
290										295										300									
Asp	Met	Asp	Met	Asn	His	Ile	Pro	Asp	Ala	Ala	Met	Thr	Ile	Ala	Thr														
305										310										315									
Ala	Ala	Leu	Phe	Ala	Lys	Gly	Thr	Thr	Arg	Leu	Arg	Asn	Ile	Tyr	Asn														
325										330										335									
Trp	Arg	Val	Lys	Glu	Thr	Asp	Arg	Leu	Phe	Ala	Met	Ala	Thr	Glu	Leu														
340										345										350									
Arg	Lys	Val	Gly	Ala	Glu	Val	Glu	Glu	Gly	His	Asp	Tyr	Ile	Arg	Ile														
355										360										365									
Thr	Pro	Pro	Glu	Lys	Leu	Asn	Phe	Ala	Glu	Ile	Ala	Thr	Tyr	Asn	Asp														
370										375										380									
His	Arg	Met	Ala	Met	Cys	Phe	Ser	Leu	Val	Ala	Leu	Ser	Asp	Thr	Pro														
385										390										395									
Val	Thr	Ile	Leu	Asp	Pro	Lys	Cys	Thr	Ala	Lys	Thr	Phe	Pro	Asp	Tyr														
405										410										415									
Phe	Glu	Gln	Leu	Ala	Arg	Ile	Ser	Gln	Ala	Ala																			
420										425																			

<210> 12
 <211> 1242
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> синтетичний grg23

<221> CDS
 <222> (1)...(1242)

<400> 12		
atg gaa act gat cgc ctt gtg atc cca gga tcg aaa agc atc acc aac	48	
Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn		
1 5 10 15		
cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg		96
Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val		
20 25 30		
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag		144
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln		
35 40 45		
gcc ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa		192
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu		
50 55 60		
ggc ctg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gac		240

71										92622					72				
Gly 65	Leu	Gly	Gln	Ala	Pro 70	His	Leu	Asp	Ala	Asp 75	Ile	Trp	Cys	Glu	Asp 80				
gca Ala	ggt Gly	act Thr	gtg Val	gcc Ala 85	cgg Arg	ttc Phe	ctc Leu	cct Pro	cca Pro 90	ttc Phe	gta Val	gcc Ala	gca Ala	ggt Gly 95	cag Gln	288			
ggg Gly	aag Lys	ttc Phe 100	acc Thr	gtc Val	gac Asp	gga Gly	tca Ser	gag Glu 105	cag Gln	ctg Leu	cgg Arg	cgc Arg 110	ccg Pro	ctt Leu		336			
cgg Arg	ccc Pro	ctg Leu 115	gtc Val	gac Asp	ggc Gly	atc Ile	cgc Arg 120	cac His	ctg Leu	ggc Gly	gcc Ala	cgc Arg 125	gtc Val	tcc Ser	tcc Ser	384			
gag Glu 130	cag Gln	ctg Leu	ccc Pro	ctt Leu	aca Thr	att Ile 135	gaa Glu	gcg Ala	agc Ser	ggg Gly	ctg Leu 140	gca Ala	ggc Gly	ggg Gly	gag Glu	432			
tac Tyr 145	gaa Glu	att Ile	gaa Glu	gcc Ala	cat His 150	cag Gln	agc Ser	agc Ser	cag Gln 155	ttc Phe	gcc Ala	tcc Ser	ggc Gly	ctg Leu 160	atc Ile	480			
atg Met	gcc Ala	gcc Ala	ccg Pro	tac Tyr 165	gcg Ala	aga Arg	caa Gln	ggc Gly	ctg Leu 170	cgt Arg	gtg Val	cgg Arg	ata Ile	cca Pro 175	aat Asn	528			
ccc Pro	gtg Val	tca Ser 180	cag Gln	ccc Pro	tac Tyr	ctc Leu	acg Thr	atg Met 185	aca Thr	ctg Leu	cgg Arg	atg Met 190	atg Met	agg Arg	gac Asp	576			
ttc Phe	ggc Gly	ctt Leu 195	gag Glu	acc Thr	agc Ser	acc Thr	gac Asp 200	gga Gly	gcc Ala	acc Thr	gtc Val	agc Ser 205	gtc Val	cct Pro	cca Pro	624			
ggg Gly 210	cgc Arg	tac Tyr	aca Thr	gcc Ala	cgg Arg	cgg Arg 215	tat Tyr	gaa Glu	ata Ile	gaa Glu 220	ccg Pro	gat Asp	gcg Ala	tca Ser	act Thr	672			
gcg Ala 225	tcg Ser	tac Tyr	ttc Phe	gcc Ala	gcc Ala 230	gct Ala	tcc Ser	gcc Ala	gtc Val	tct Ser 235	ggc Gly	agg Arg	agc Ser	ttc Phe 240	gaa Glu	720			
ttt Phe	caa Gln	ggc Gly	ctt Leu	ggc Gly 245	aca Thr	gac Asp	agc Ser	atc Ile	caa Gln 250	ggc Gly	gac Asp	acg Thr	tca Ser	ttc Phe 255	ttc Phe	768			
aat Asn	gta Val	ctt Leu 260	ggg Gly	cgg Arg	ctc Leu	ggt Gly	gcg Ala	gag Glu 265	gtc Val	cac His	tgg Trp	gca Ala	ccc Pro 270	aac Asn	tcg Ser	816			
gtc Val	acc Thr 275	ata Ile	tct Ser	gga Gly	ccg Pro	gaa Glu	agg Arg 280	ctg Leu	aac Asn	ggc Gly	gac Asp	att Ile 285	gaa Glu	gtg Val	gat Asp	864			
atg Met 290	ggc Gly	gag Glu	att Ile	tcg Ser	gac Asp	acc Thr 295	ttc Phe	atg Met	aca Thr	ctc Leu 300	gcg Ala	gcg Ala	att Ile	gcc Ala	cct Pro	912			
ttg Leu 305	gcc Ala	gat Asp	gga Gly	ccc Pro	atc Ile 310	acg Thr	ata Ile	acc Thr	aac Asn	att Ile 315	ggt Gly	cat His	gca Ala	cgg Arg	ttg Leu 320	960			
aag	gaa	tcc	gac	cgc	atc	tca	gcg	atg	gaa	acc	aac	ctg	cgc	acg	ctc	1008			

73										92622										74									
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu														
325										330										335									
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct 1056 Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350																													
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct 1104 Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365																													
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac 1152 Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 370 375 380																													
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt 1200 Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu 385 390 395 400																													
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag 1242 Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly * 405 410																													

<210> 13

<211> 413

<212> PRT

<213> Штучна послідовність

<220>

<223> синтетичний grg23

<400> 13

Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn				
1				5					10					15					
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val				
		20						25					30						
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln				
		35					40				45								
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu				
	50				55						60								
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp				
	65				70					75				80					
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln				
			85						90					95					
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu				
		100					105						110						
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser				
		115					120					125							
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu				
	130					135					140								
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile				
	145				150					155					160				
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn				
			165						170					175					
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp				
		180						185					190						
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro				
		195					200					205							
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr				
	210					215					220								
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu				
	225				230					235					240				

75								92622								76			
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe				
				245					250					255					
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser				
			260						265					270					
Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp				
			275				280						285						
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro				
	290					295					300								
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu				
305					310					315					320				
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu				
				325					330					335					
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser				
			340						345				350						
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala				
		355					360					365							
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp				
	370					375				380									
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu				
385					390					395					400				
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly							
				405					410										

77	92622	78	
cgg ccc ctg gtc gac ggc atc	cgc cac ctg ggc gcc cgc gtc tcc tcc		384
Arg Pro Leu Val Asp Gly Ile	Arg His Leu Gly Ala Arg Val Ser Ser		
115	120	125	
gag cag ctg ccc ctt aca att	gaa gcg agc ggg ctg gca ggc ggg gag		432
Glu Gln Leu Pro Leu Thr	Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu		
130	135	140	
tac gaa att gaa gcc cat cag	agc agc cag ttc gcc tcc ggc ctg atc		480
Tyr Glu Ile Glu Ala His	Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile		
145	150	155	160
atg gcc gcc ccg tac gcg aga	caa ggc ctg cgt gtg cgg ata cca aat		528
Met Ala Ala Pro Tyr Ala Arg	Gln Gly Leu Arg Val Arg Ile Pro Asn		
165	170	175	
ccc gtg tca cag ccc tac ctc	acg atg aca ctg cgg atg atg agg gac		576
Pro Val Ser Gln Pro Tyr Leu	Thr Met Thr Leu Arg Met Met Arg Asp		
180	185	190	
ttc ggc ctt gag acc agc acc	gac gga gcc acc gtc agc gtc cct cca		624
Phe Gly Leu Glu Thr Ser Thr	Asp Gly Ala Thr Val Ser Val Pro Pro		
195	200	205	
ggg cgc tac aca gcc ccg ccg	tat gaa ata gaa ccg gat gcg tca act		672
Gly Arg Tyr Thr Ala Arg Arg	Tyr Glu Ile Glu Pro Asp Ala Ser Thr		
210	215	220	
gcg tcg tac ttc gcc gcc gct	tcc gcc gtc tct ggc agg agc ttc gaa		720
Ala Ser Tyr Phe Ala Ala Ala	Ser Ala Val Ser Gly Arg Ser Phe Glu		
225	230	235	240
ttt caa ggc ctt ggc aca gac	agc atc caa ggc gac acg tca ttc ttc		768
Phe Gln Gly Leu Gly Thr Asp	Ser Ile Gln Gly Asp Thr Ser Phe Phe		
245	250	255	
aat gta ctt ggg ccg ctc ggt	gcg gag gtc cac tgg gca ccc aac tcg		816
Asn Val Leu Gly Arg Leu Gly	Ala Glu Val His Trp Ala Pro Asn Ser		
260	265	270	
gtc acc ata act gga ccg gaa	agg ctg aac ggc gac att gaa gtg gat		864
Val Thr Ile Thr Gly Pro Glu	Arg Leu Asn Gly Asp Ile Glu Val Asp		
275	280	285	
atg ggc gag att tcg gac acc	ttc atg aca ctc gcg gcg att gcc cct		912
Met Gly Glu Ile Ser Asp Thr	Phe Met Thr Leu Ala Ala Ile Ala Pro		
290	295	300	
ttg gcc gat gga ccc atc acg	ata acc aac att ggt cat gca cgg ttg		960
Leu Ala Asp Gly Pro Ile Thr	Ile Thr Asn Ile Gly His Ala Arg Leu		
305	310	315	320
aag gaa tcc gac cgc atc tca	gcg atg gaa acc aac ctg cgc acg ctc		1008
Lys Glu Ser Asp Arg Ile Ser	Ala Met Glu Thr Asn Leu Arg Thr Leu		
325	330	335	
ggt gta caa acc gac gtc gga	cac gac tgg atg aga atc tac ccc tct		1056
Gly Val Gln Thr Asp Val Gly	His Asp Trp Met Arg Ile Tyr Pro Ser		
340	345	350	
acc ccg cac ggc ggt aga gtg	aat tgc cac cgg gac cac agg atc gct		1104
Thr Pro His Gly Gly Arg Val	Asn Cys His Arg Asp His Arg Ile Ala		
355	360	365	

79										92622										80										
atg	gcg	ttt	tca	atc	ctg	gga	ctg	aga	gtg	gac	ggg	att	acc	ctc	gac	1152														
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp															
370					375					380																				
gac	cct	caa	tgc	gtc	ggg	aag	acc	ttt	cct	ggc	ttc	ttc	gac	tac	ctt	1200														
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu															
385					390					395					400															
gga	cgc	ctt	ttc	ccc	gaa	aag	gcg	ctt	acg	ctc	ccc	ggc	tag			1242														
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly	*																	
405					410																									

<210> 15
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg23(ace1)

<400> 15

Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn
1				5					10					15	
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val
			20					25					30		
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln
		35					40					45			
Thr	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu
	50					55					60				
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp
65					70					75				80	
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln
				85					90					95	
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu
			100					105					110		
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser
		115					120					125			
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu
	130					135					140				
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile
145					150					155					160
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn
				165					170					175	
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp
			180					185					190		
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro
		195					200					205			
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr
	210					215					220				
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu
225					230					235					240
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe
				245					250					255	
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser
			260					265					270		
Val	Thr	Ile	Thr	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp
		275					280					285			
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro
	290					295					300				
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu
305					310					315					320
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu

81										92622										82									
				325										330										335					
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser														
			340					345					350																
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala														
		355					360						365																
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp														
	370					375						380																	
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu														
385					390							395			400														
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																	
			405					410																					

<210> 16
 <211> 1242
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg23(ace2)

<221> CDS
 <222> (1)...(1242)

<400> 16	
atg gaa act gat cgc ctt gtg atc cca gga tcg aaa agc atc acc aac	48
Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn	
1 5 10 15	
cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg	96
Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val	
20 25 30	
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag	144
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln	
35 40 45	
gcc ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa	192
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu	
50 55 60	
ggc ctg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gac	240
Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp	
65 70 75 80	
gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag	288
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln	
85 90 95	
ggg aag ttc acc gtc gac gga tca gag cag ctg cgg cgg cgc ccg ctt	336
Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu	
100 105 110	
cgg ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc	384
Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser	
115 120 125	
gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca ggc ggg gag	432
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu	
130 135 140	
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc	480
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile	
145 150 155 160	

83	92622	84	
- atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn 165 170 175			528
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp 180 185 190			576
ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct cca Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro 195 200 205			624
ggg cgc tac aca gcc cgg cgg tat gaa ata gaa ccg gat gcg tca act Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr 210 215 220			672
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu 225 230 235 240			720
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe 245 250 255			768
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser 260 265 270			816
gtc acc ata cgg gga ccg gaa agg ctg aac ggc gac att gaa gtg gat Val Thr Ile Arg Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp 275 280 285			864
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro 290 295 300			912
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu 305 310 315 320			960
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu 325 330 335			1008
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350			1056
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365			1104
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 370 375 380			1152
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu 385 390 395 400			1200
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			1242
405 410			

<210> 17
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg23(ace2)

<400> 17
 Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn
 1 5 10 15
 Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val
 20 25 30
 Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln
 35 40 45
 Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu
 50 55 60
 Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp
 65 70 75 80
 Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln
 85 90 95
 Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu
 100 105 110
 Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser
 115 120 125
 Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu
 130 135 140
 Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile
 145 150 155 160
 Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn
 165 170 175
 Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp
 180 185 190
 Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro
 195 200 205
 Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr
 210 215 220
 Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu
 225 230 235 240
 Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe
 245 250 255
 Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser
 260 265 270
 Val Thr Ile Arg Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp
 275 280 285
 Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro
 290 295 300
 Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu
 305 310 315 320
 Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu
 325 330 335
 Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser
 340 345 350
 Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala
 355 360 365
 Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp
 370 375 380
 Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu
 385 390 395 400
 Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly
 405 410

87

92622

88

<210> 18
 <211> 1242
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg51.4

<221> CDS
 <222> (1)...(1242)

<400> 18
 atg gaa act gat cga cta gtg atc cca gga tcg aaa agc atc acc aac 48
 Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn
 1 5 10 15

egg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg 96
 Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val
 20 25 30

aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca att cag 144
 Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln
 35 40 45

gcc ctc ggt gcc aac gtc tca gcg gac ggt gat gat tgg gtc gtt gaa 192
 Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asp Trp Val Val Glu
 50 55 60

ggc ctg ggc cag gca ccc aac ctc gac gcc gac atc tgg tgc gag gat 240
 Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala Asp Ile Trp Cys Glu Asp
 65 70 75 80

gcc ggt acc gtg gcc cgg ttc ctc cct cca ttc gtc gcc gca gga cag 288
 Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln
 85 90 95

ggg aag ttc acc gtc gac gga agc gag cag ctg cgg cgg cgc ccg ctt 336
 Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu
 100 105 110

egg ccc gtg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc 384
 Arg Pro Val Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser
 115 120 125

gag cag ctg ccc cta acg att gaa gcg agc ggg ctg gca ggc ggg gag 432
 Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu
 130 135 140

tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggt ctg atc 480
 Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile
 145 150 155 160

atg gcc gcc ccg tac gcg cga caa ggc ctg cgt gtt cgg ata cca aat 528
 Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn
 165 170 175

ccc gtg agc cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac 576
 Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp
 180 185 190

ttc ggc att gag acc agc acc gac gga gcg acc gtc agc gtt cct ccc 624
 Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro

89	92622	90
195	200	205
ggg cgc tac aca gcg cgg cgg tat gag att gaa ccg gac gcg tca act Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr 210 215 220		672
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc cgg cgc ttc gaa Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Arg Phe Glu 225 230 235 240		720
ttc cag ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe 245 250 255		768
aat gta ctt ggg cgg ctc ggc gca gag gtc cac tgg gca tcc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Ser Asn Ser 260 265 270		816
gtc acc ata cgc gga ccg gaa agg ctg acc ggc gac att gaa gtg gat Val Thr Ile Arg Gly Pro Glu Arg Leu Thr Gly Asp Ile Glu Val Asp 275 280 285		864
atg ggc gag ata tcg gac acc ttc atg aca ctg gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro 290 295 300		912
cta gcc gat gga ccc atc acg ata aca aac att ggc cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu 305 310 315 320		960
aag gaa tcc gac cgc atc tcg gcg atg gaa agc aac ctt cga acg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Ser Asn Leu Arg Thr Leu 325 330 335		1008
ggt gta caa acc gac gtc gga cac gac tgg atg cga atc tac ccc tct Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350		1056
acc ccg cac ggc gcc aga gtc aat tgc cac cgg gac cac agg atc gcc Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365		1104
atg gcg ttt tca atc ctg gga ctg cga gtg gac ggg att acc ctc gac Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 370 375 380		1152
gac cct caa tgt gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu 385 390 395 400		1200
gga cgc ctt ttc ccg gaa aag gcg ctt acg ctc ccc ggc tag Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly * 405 410		1242

<210> 19
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg51.4

91

92622

92

<400> 19

```

Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn
 1      5      10      15
Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val
      20      25      30
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln
      35      40      45
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asp Trp Val Val Glu
      50      55      60
Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala Asp Ile Trp Cys Glu Asp
      65      70      75      80
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln
      85      90      95
Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu
      100      105      110
Arg Pro Val Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser
      115      120      125
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu
      130      135      140
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile
      145      150      155      160
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn
      165      170      175
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp
      180      185      190
Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro
      195      200      205
Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr
      210      215      220
Ala Ser Tyr Phe Ala Ala Ser Ala Val Ser Gly Arg Arg Phe Glu
      225      230      235      240
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe
      245      250      255
Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Ser Asn Ser
      260      265      270
Val Thr Ile Arg Gly Pro Glu Arg Leu Thr Gly Asp Ile Glu Val Asp
      275      280      285
Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro
      290      295      300
Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu
      305      310      315      320
Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Ser Asn Leu Arg Thr Leu
      325      330      335
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser
      340      345      350
Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala
      355      360      365
Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp
      370      375      380
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu
      385      390      395      400
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly
      405      410

```

<210> 20

<211> 1242

<212> ДНК

<213> Штучна послідовність

<220>

<223> grg23(ace3)

<221> CDS

93

92622

94

<222> (1)...(1242)

<400> 20

atg gaa act gat cgc ctt gtg atc cca gga tcg aaa agc atc acc aac	48
Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn	
1 5 10 15	
cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg	96
Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val	
20 25 30	
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag	144
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln	
35 40 45	
gcc ctc ggt gcc aac gtc tca gcc gac ggt gac gat tgg gtc gtt gaa	192
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Trp Val Val Glu	
50 55 60	
ggc ctg ggt cag gca ccc aac ctc gac gcc gac atc tgg tgc gag gac	240
Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala Asp Ile Trp Cys Glu Asp	
65 70 75 80	
gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag	288
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln	
85 90 95	
ggg aag ttc acc gtc gac gga tca gag cag ctg cgg cgg cgc ccg ctt	336
Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu	
100 105 110	
cgg ccc gtg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc	384
Arg Pro Val Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser	
115 120 125	
gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca ggc ggg gag	432
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu	
130 135 140	
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc	480
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile	
145 150 155 160	
atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat	528
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn	
165 170 175	
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac	576
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp	
180 185 190	
ttc ggc att gag acc agc acc gac gga gcc acc gtc agc gtc cct cca	624
Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro	
195 200 205	
ggg cgc tac aca gcc cgg cgg tat gaa ata gaa ccg gat gcg tca act	672
Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr	
210 215 220	
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg cgc ttc gaa	720
Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Arg Phe Glu	
225 230 235 240	
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc	768

95										92622										96																	
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe																						
				245					250						255																						
aat	gta	ctt	ggg	cgg	ctc	ggt	gcg	gag	gtc	cac	tgg	gca	tcc	aac	tcg																						
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Ser	Asn	Ser																						
			260				265						270																								
gtc	acc	ata	cgt	gga	ccg	gaa	agg	ctg	acc	ggc	gac	att	gaa	gtg	gat																						
Val	Thr	Ile	Arg	Gly	Pro	Glu	Arg	Leu	Thr	Gly	Asp	Ile	Glu	Val	Asp																						
			275				280					285																									
atg	ggc	gag	att	tcg	gac	acc	ttc	atg	aca	ctc	gcg	gcg	att	gcc	cct																						
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro																						
			290				295					300																									
ttg	gcc	gat	gga	ccc	atc	acg	ata	acc	aac	att	ggt	cat	gca	cgg	ttg																						
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu																						
305					310				315					320																							
aag	gaa	tcc	gac	cgc	atc	tca	gcg	atg	gaa	agc	aac	ctg	cgc	acg	ctc																						
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Ser	Asn	Leu	Arg	Thr	Leu																						
				325				330					335																								
ggt	gta	caa	acc	gac	gtc	gga	cac	gac	tgg	atg	aga	atc	tac	ccc	tct																						
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser																						
			340				345					350																									
acc	ccg	cac	ggc	ggt	aga	gtg	aat	tgc	cac	cgg	gac	cac	agg	atc	gct																						
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala																						
			355				360					365																									
atg	gcg	ttt	tca	atc	ctg	gga	ctg	aga	gtg	gac	ggg	att	acc	ctc	gac																						
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp																						
			370			375				380																											
gac	cct	caa	tgc	gtc	ggg	aag	acc	ttt	cct	ggc	ttc	ttc	gac	tac	ctt																						
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu																						
385				390					395					400																							
gga	cgc	ctt	ttc	ccc	gaa	aag	gcg	ctt	acg	ctc	ccc	ggc	tag																								
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly	*																								
			405				410																														

<210> 21
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg23(ace3)

<400> 21
 Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn
 1 5 10 15
 Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val
 20 25 30
 Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln
 35 40 45
 Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asp Trp Val Val Glu
 50 55 60
 Gly Leu Gly Gln Ala Pro Asn Leu Asp Ala Asp Ile Trp Cys Glu Asp
 65 70 75 80

97										92622										98									
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln														
				85					90					95															
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu														
			100					105					110																
Arg	Pro	Val	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser														
			115					120					125																
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu														
			130					135					140																
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile														
145					150					155					160														
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn														
				165					170					175															
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp														
			180					185					190																
Phe	Gly	Ile	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro														
			195				200						205																
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr														
			210				215					220																	
Ala	Ser	Tyr	Phe	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Arg	Phe	Glu															
225					230					235				240															
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe														
				245					250					255															
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Ser	Asn	Ser														
			260					265					270																
Val	Thr	Ile	Arg	Gly	Pro	Glu	Arg	Leu	Thr	Gly	Asp	Ile	Glu	Val	Asp														
			275					280					285																
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro														
			290				295					300																	
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu														
305					310					315				320															
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Ser	Asn	Leu	Arg	Thr	Leu														
				325					330					335															
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser														
			340					345					350																
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala														
			355				360						365																
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp														
			370				375					380																	
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu														
385					390					395				400															
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																	
				405					410																				

<210> 22

<211> 1242

<212> ДНК

<213> Штучна послідовність

<220>

<223> grg23(L5P2.J2)

<221> CDS

<222> (1)...(1242)

<400> 22

atg	gaa	act	gat	cgc	ctt	gtg	atc	cca	gga	tcg	aaa	agc	atc	acc	aac	48
Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn	
1				5					10					15		

cgg	gct	ttg	ctt	ttg	gct	gcc	gca	gcg	aag	ggc	acg	tcg	gtc	ctg	gtg	96
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val	
				20				25					30			

99	92622	100	
aga cca ttg gtc agc gcc gat	acc tca gca ttc aaa act gca atc cag	144	
Arg Pro Leu Val Ser Ala Asp	Thr Ser Ala Phe Lys Thr Ala Ile Gln		
35	40 45		
acc ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa	192		
Thr Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu			
50 55 60			
ggc ctg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gac	240		
Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp			
65 70 75 80			
gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag	288		
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln			
85 90 95			
ggg aag ttc acc ttc gac gga tca gag cag ctg cgg cgg cgc ccg ctt	336		
Gly Lys Phe Thr Phe Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu			
100 105 110			
cgg ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc	384		
Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser			
115 120 125			
gag cag ctg ccc ctt aca att gaa gcg agt ggg ctg gca ggc ggg gag	432		
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu			
130 135 140			
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc	480		
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile			
145 150 155 160			
atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat	528		
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn			
165 170 175			
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac	576		
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp			
180 185 190			
ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct cca	624		
Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro			
195 200 205			
ggg cgc tac aca tcc cgg cgg tat gaa ata gaa ccg gat gcg tca act	672		
Gly Arg Tyr Thr Ser Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr			
210 215 220			
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa	720		
Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu			
225 230 235 240			
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc	768		
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe			
245 250 255			
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg	816		
Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser			
260 265 270			
gtc acc ata act gga ccg gaa agg ctg aac ggc aac att gaa gtg gat	864		
Val Thr Ile Thr Gly Pro Glu Arg Leu Asn Gly Asn Ile Glu Val Asp			
275 280 285			

101	92622	102	
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct			912
Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro			
290	295	300	
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg			960
Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu			
305	310	315	320
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc			1008
Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu			
	325	330	335
ggg gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct			1056
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser			
	340	345	350
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct			1104
Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala			
	355	360	365
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac			1152
Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp			
370	375	380	
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt			1200
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu			
385	390	395	400
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag			1242
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			
	405	410	

<210> 23

<211> 413

<212> PRT

<213> Штучна послідовність

<220>

<223> grg23(L5P2.J2)

<400> 23

Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn
1				5					10				15		
Arg	Ala	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val	
		20					25					30			
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln
		35				40					45				
Thr	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu
	50				55					60					
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp
65				70					75				80		
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln
			85			90						95			
Gly	Lys	Phe	Thr	Phe	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu
		100				105						110			
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser
		115				120					125				
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu
	130				135					140					
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile
145				150					155					160	
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn

103										92622										104									
165										170										175									
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp	Met	Thr	Leu	Arg	Met	Met	Arg	Asp						
180										185										190									
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro														
195										200										205									
Gly	Arg	Tyr	Thr	Ser	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr														
210										215										220									
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu														
225										230										235									
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe														
245										250										255									
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser														
260										265										270									
Val	Thr	Ile	Thr	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asn	Ile	Glu	Val	Asp														
275										280										285									
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro														
290										295										300									
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu														
305										310										315									
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu														
325										330										335									
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser														
340										345										350									
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala														
355										360										365									
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp														
370										375										380									
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu														
385										390										395									
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																	
405										410																			

<210> 24
 <211> 1242
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg23(ace4)

<221> CDS
 <222> (1)...(1242)

<400> 24																								
atg	gaa	act	gat	cgc	ctt	gtg	atc	cca	gga	tcg	aaa	agc	atc	acc	aac	48								
Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn									
1				5				10					15											
cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg 96																								
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val									
				20				25					30											
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag 144																								
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln									
				35				40					45											
gcc ctc ggt gcc aac gtc tca gcc gac ggt gac gat tgg gtc gtt gaa 192																								
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asp	Trp	Val	Val	Glu									
				50				55				60												
ggc ctg ggt cag gca ccc aac ctc gac gcc gac atc tgg tgc gag gac 240																								
Gly	Leu	Gly	Gln	Ala	Pro	Asn	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp									
				65				70				75												

105	92622	106	
gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln	85	90	95
ggg aag ttc acc ttc gac gga tca gag cag ctg cgg cgg cgc ccg ctt Gly Lys Phe Thr Phe Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu	100	105	110
cgg ccc gtg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc Arg Pro Val Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser	115	120	125
gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca ggc ggg gag Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu	130	135	140
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile	145	150	155
atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn	165	170	175
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp	180	185	190
ttc ggc att gag acc agc acc gac gga gcc acc gtc agc gtc cct cca Phe Gly Ile Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro	195	200	205
ggg cgc tac aca gcc cgg cgg tat gaa ata gaa ccg gat gcg tca act Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr	210	215	220
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg cgc ttc gaa Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Arg Phe Glu	225	230	240
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe	245	250	255
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca tcc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Ser Asn Ser	260	265	270
gtc acc ata cgg gga ccg gaa agg ctg acc ggc gac att gaa gtg gat Val Thr Ile Arg Gly Pro Glu Arg Leu Thr Gly Asp Ile Glu Val Asp	275	280	285
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro	290	295	300
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu	305	310	315
aag gaa tcc gac cgc atc tca gcg atg gaa agc aac ctg cgc acg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Ser Asn Leu Arg Thr Leu	325	330	335

107	92622	108	
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct			1056
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser			
340	345	350	
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct			1104
Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala			
355	360	365	
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac			1152
Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp			
370	375	380	
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt			1200
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu			
385	390	395	400
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag			1242
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			
405	410		

<210> 25
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg23(ace4)

<400> 25

Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn
1				5					10					15	
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val
		20						25					30		
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln
		35					40					45			
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asp	Trp	Val	Val	Glu
	50				55					60					
Gly	Leu	Gly	Gln	Ala	Pro	Asn	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp
65					70				75					80	
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln
			85					90					95		
Gly	Lys	Phe	Thr	Phe	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu
			100				105					110			
Arg	Pro	Val	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser
	115					120					125				
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu
130					135					140					
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile
145				150					155					160	
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn
			165					170					175		
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp
			180				185					190			
Phe	Gly	Ile	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro
	195						200				205				
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr
210					215					220					
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Arg	Phe	Glu
225				230					235					240	
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe
			245					250						255	

111	92622	112
115	120	125
gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca ggc ggg gag Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu 130 135 140		432
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile 145 150 155 160		480
atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn 165 170 175		528
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp 180 185 190		576
ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc atc cct cca Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Ile Pro Pro 195 200 205		624
ggg cgc tac aca gcc ccg ccg tat gaa ata gaa ccg gat gcg tca act Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr 210 215 220		672
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu 225 230 235 240		720
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe 245 250 255		768
aat gta ctt ggg ccg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser 260 265 270		816
gtc acc ata tct gga ccg gaa agg ctg aac ggc gac att gaa gtg gat Val Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp 275 280 285		864
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro 290 295 300		912
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu 305 310 315 320		960
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu 325 330 335		1008
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350		1056
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365		1104
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 1152		

113	92622	114
370	375	380
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt		1200
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu		
385	390	400
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag		1242
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *		
405	410	
gg		1244
<210> 27		
<211> 413		
<212> PRT		
<213> Штучна послідовність		
<220>		
<223> grg23(L3P1.B20)		
<400> 27		
Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn		
1	5	10
Arg Ala Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val		15
20	25	30
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln		
35	40	45
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu		
50	55	60
Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp		
65	70	75
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln		
85	90	95
Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu		
100	105	110
Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser		
115	120	125
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu		
130	135	140
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile		
145	150	155
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn		
165	170	175
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp		
180	185	190
Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Ile Pro Pro		
195	200	205
Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr		
210	215	220
Ala Ser Tyr Phe Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu		
225	230	235
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe		
245	250	255
Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser		
260	265	270
Val Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp		
275	280	285
Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro		
290	295	300
Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu		
305	310	315
Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu		
325	330	335
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser		

115										92622										116									
			340									345										350							
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala														
		355					360					365										365							
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp														
	370					375					380																		
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu														
385					390					395					400														
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																	
			405						410																				

<210> 28
 <211> 1244
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg23(L3P1.B3)

<221> CDS
 <222> (1)...(1242)

<400> 28

atg	gaa	act	gat	cgc	ctt	gtg	atc	cca	gga	tcg	aaa	agc	atc	acc	aac		48
Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn		
1				5					10					15			
cgg	gct	ttg	ctt	ttg	gct	gcc	gca	gcg	aag	ggc	acg	tcg	gtc	ctg	gtg		96
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val		
			20					25					30				
aga	cca	ttg	gtc	agc	gcc	gat	acc	tca	gca	ttc	aaa	act	gca	atc	cag		144
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln		
			35				40						45				
gcc	ctc	ggt	gcc	aac	gtc	tca	gcc	gac	ggt	gac	aat	tgg	gtc	gtt	gaa		192
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu		
	50					55					60						
ggc	ctg	ggt	cag	gca	ccc	cac	ctc	gac	gcc	cac	atc	tgg	tgc	gag	gac		240
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	His	Ile	Trp	Cys	Glu	Asp		
	65					70				75					80		
gca	ggt	act	gtg	gcc	cgg	ttc	ctc	cct	cca	ttc	gta	gcc	gca	ggt	cag		288
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln		
				85					90					95			
ggg	aag	ttc	acc	gtc	gac	gga	tca	gag	cag	ctg	cgg	cgg	cgc	ccg	ctt		336
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu		
			100					105					110				
cgg	ccc	ctg	gtc	gac	ggc	atc	cgc	cac	ctg	ggc	gcc	cgc	gtc	tcc	tcc		384
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser		
			115				120					125					
gag	cag	ctg	ccc	ctt	aca	att	gaa	gcg	agc	ggg	ctg	gca	ggc	ggg	gag		432
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu		
			130				135					140					
tac	gaa	att	gaa	gcc	cat	cag	agc	agc	cag	ttc	gcc	tcc	ggc	ctg	atc		480
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile		
	145					150				155					160		
atg	gcc	gcc	ccg	tac	gcg	aga	caa	ggc	ctg	cgt	gtg	cgg	ata	cca	aat		528

117	92622	118	
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn			
165	170	175	
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac			576
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp			
180	185	190	
ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct cca			624
Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro			
195	200	205	
ggg cgc tac aca gcc cgg cgg tat aaa ata gaa ccg gat gcg tca act			672
Gly Arg Tyr Thr Ala Arg Arg Tyr Lys Ile Glu Pro Asp Ala Ser Thr			
210	215	220	
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa			720
Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu			
225	230	235	240
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc			768
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe			
245	250	255	
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg			816
Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser			
260	265	270	
gtc acc ata tct gga ccg gaa agg ctg aac ggc gac att gaa gtg gat			864
Val Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp			
275	280	285	
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct			912
Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro			
290	295	300	
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg			960
Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu			
305	310	315	320
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc			1008
Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu			
325	330	335	
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct			1056
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser			
340	345	350	
acc ccg cac ggc ggt aga gtg aat tgc cac cga gac cac agg atc gct			1104
Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala			
355	360	365	
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac			1152
Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp			
370	375	380	
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt			1200
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu			
385	390	395	400
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag			1242
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			
405	410		
gg			1244

119

92622

120

<210> 29

<211> 413

<212> PRT

<213> Штучна послідовність

<220>

<223> grg23(L3P1.B3)

<400> 29

Met	Glu	Thr	Asp	Arg	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn
1				5					10				15		
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val
			20				25						30		
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln
		35				40					45				
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu
	50				55					60					
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	His	Ile	Trp	Cys	Glu	Asp
65				70					75					80	
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln
			85					90					95		
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu
			100				105					110			
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser
		115				120					125				
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu
	130				135						140				
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile
145				150					155					160	
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn
			165					170					175		
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp
		180					185					190			
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro
	195					200					205				
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Lys	Ile	Glu	Pro	Asp	Ala	Ser	Thr
	210				215						220				
Ala	Ser	Tyr	Phe	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu	
225				230					235					240	
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe
			245					250					255		
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser
	260					265						270			
Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp
	275					280					285				
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro
	290				295					300					
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu
305				310					315					320	
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu
			325						330				335		
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser
		340				345						350			
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala
	355					360					365				
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp
	370				375						380				
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu
385				390					395						400
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly			
			405						410						

<210> 30
 <211> 1244
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg23(L3P1.F18)

<221> CDS
 <222> (1)...(1242)

<400> 30
 atg gaa act gat cgc ctt gtg atc cca gga tcg aaa agc atc acc aac 48
 Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn
 1 5 10 15

cgg gct ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg 96
 Arg Ala Leu Leu Leu Ala Ala Ala Lys Gly Thr Ser Val Leu Val
 20 25 30

aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag 144
 Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln
 35 40 45

gcc ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa 192
 Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu
 50 55 60

ggc ctg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gac 240
 Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp
 65 70 75 80

gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag 288
 Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln
 85 90 95

ggg aag ttc acc gtc gac gga tca gag cag ctg cgg cgg cgc ccg ctt 336
 Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu
 100 105 110

cgg ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc 384
 Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser
 115 120 125

gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca ggc ggg gag 432
 Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu
 130 135 140

tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc 480
 Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile
 145 150 155 160

atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat 528
 Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn
 165 170 175

ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac 576
 Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp
 180 185 190

ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct cca 624
 Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro
 195 200 205

123	92622	124	
ggg cgc tac aca gcc cgg cgg tat gaa ata gaa ccg gat gcg tca act			672
Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr			
210	215	220	
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa			720
Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu			
225	230	235	240
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttc ttc			768
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe			
	245	250	255
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg			816
Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser			
	260	265	270
gtc atc ata tct gga ccg gaa agg ctg aac ggc gac att gaa gtg gat			864
Val Ile Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp			
	275	280	285
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct			912
Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro			
	290	295	300
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg			960
Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu			
305	310	315	320
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc			1008
Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu			
	325	330	335
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct			1056
Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser			
	340	345	350
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct			1104
Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala			
	355	360	365
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac			1152
Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp			
	370	375	380
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt			1200
Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu			
385	390	395	400
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctc ccc ggc tag			1242
Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly *			
	405	410	

gg 1244

<210> 31
 <211> 413
 <212> PRT
 <213> Штучна послідовність

<220>
 <223> grg23(L3P1.F18)

<400> 31
 Met Glu Thr Asp Arg Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn

125										92622										126									
1				5						10										15									
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val													
			20						25																				
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln														
		35					40																						
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu														
	50					55																							
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp														
65				70					75																				
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln														
				85					90																				
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu														
			100					105																					
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser														
		115				120																							
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu														
	130					135																							
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile														
145				150					155																				
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn														
			165					170																					
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp														
			180					185																					
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro														
		195				200																							
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr														
	210					215																							
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu														
225				230					235																				
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe														
			245					250																					
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser														
		260						265																					
Val	Ile	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp														
		275				280																							
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro														
	290					295																							
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu														
305				310					315																				
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu														
			325						330																				
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser														
			340					345																					
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala														
		355				360																							
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp														
	370					375																							
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu														
385				390					395																				
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly																	
				405					410																				

<210> 32
 <211> 1244
 <212> ДНК
 <213> Штучна послідовність

<220>
 <223> grg23(L3P1.O23)

<221> CDS
 <222> (1)...(1242)

127	92622	128
<400> 32		
atg gaa act gat cac ctt gtg atc cca gga tcg aaa agc atc acc aac		48
Met Glu Thr Asp His Leu Val Ile Pro Gly Ser Lys Ser Ile Thr Asn		
1 5 10 15		
cgg gcg ttg ctt ttg gct gcc gca gcg aag ggc acg tcg gtc ctg gtg		96
Arg Ala Leu Leu Leu Ala Ala Ala Ala Lys Gly Thr Ser Val Leu Val		
20 25 30		
aga cca ttg gtc agc gcc gat acc tca gca ttc aaa act gca atc cag		144
Arg Pro Leu Val Ser Ala Asp Thr Ser Ala Phe Lys Thr Ala Ile Gln		
35 40 45		
gcc ctc ggt gcc aac gtc tca gcc gac ggt gac aat tgg gtc gtt gaa		192
Ala Leu Gly Ala Asn Val Ser Ala Asp Gly Asp Asn Trp Val Val Glu		
50 55 60		
ggc ttg ggt cag gca ccc cac ctc gac gcc gac atc tgg tgc gag gac		240
Gly Leu Gly Gln Ala Pro His Leu Asp Ala Asp Ile Trp Cys Glu Asp		
65 70 75 80		
gca ggt act gtg gcc cgg ttc ctc cct cca ttc gta gcc gca ggt cag		288
Ala Gly Thr Val Ala Arg Phe Leu Pro Pro Phe Val Ala Ala Gly Gln		
85 90 95		
ggg aag ttc acc gtc gac gga tca gag cag ctg cgg cgg cgc ccg ctt		336
Gly Lys Phe Thr Val Asp Gly Ser Glu Gln Leu Arg Arg Arg Pro Leu		
100 105 110		
cgg ccc ctg gtc gac ggc atc cgc cac ctg ggc gcc cgc gtc tcc tcc		384
Arg Pro Leu Val Asp Gly Ile Arg His Leu Gly Ala Arg Val Ser Ser		
115 120 125		
gag cag ctg ccc ctt aca att gaa gcg agc ggg ctg gca gcc ggg gag		432
Glu Gln Leu Pro Leu Thr Ile Glu Ala Ser Gly Leu Ala Gly Gly Glu		
130 135 140		
tac gaa att gaa gcc cat cag agc agc cag ttc gcc tcc ggc ctg atc		480
Tyr Glu Ile Glu Ala His Gln Ser Ser Gln Phe Ala Ser Gly Leu Ile		
145 150 155 160		
atg gcc gcc ccg tac gcg aga caa ggc ctg cgt gtg cgg ata cca aat		528
Met Ala Ala Pro Tyr Ala Arg Gln Gly Leu Arg Val Arg Ile Pro Asn		
165 170 175		
ccc gtg tca cag ccc tac ctc acg atg aca ctg cgg atg atg agg gac		576
Pro Val Ser Gln Pro Tyr Leu Thr Met Thr Leu Arg Met Met Arg Asp		
180 185 190		
ttc ggc ctt gag acc agc acc gac gga gcc acc gtc agc gtc cct cca		624
Phe Gly Leu Glu Thr Ser Thr Asp Gly Ala Thr Val Ser Val Pro Pro		
195 200 205		
ggg cgc tac aca gcc cgg cgg tat gaa ata gaa ccg gat gcg tca act		672
Gly Arg Tyr Thr Ala Arg Arg Tyr Glu Ile Glu Pro Asp Ala Ser Thr		
210 215 220		
gcg tcg tac ttc gcc gcc gct tcc gcc gtc tct ggc agg agc ttc gaa		720
Ala Ser Tyr Phe Ala Ala Ala Ser Ala Val Ser Gly Arg Ser Phe Glu		
225 230 235 240		
ttt caa ggc ctt ggc aca gac agc atc caa ggc gac acg tca ttt ttc		768
Phe Gln Gly Leu Gly Thr Asp Ser Ile Gln Gly Asp Thr Ser Phe Phe		
245 250 255		

129	92622	130	
aat gta ctt ggg cgg ctc ggt gcg gag gtc cac tgg gca ccc aac tcg Asn Val Leu Gly Arg Leu Gly Ala Glu Val His Trp Ala Pro Asn Ser 260 265 270			816
gtc acc ata tct gga ccg gaa agg ctg aac ggc gac att gaa gtg gat Val Thr Ile Ser Gly Pro Glu Arg Leu Asn Gly Asp Ile Glu Val Asp 275 280 285			864
atg ggc gag att tcg gac acc ttc atg aca ctc gcg gcg att gcc cct Met Gly Glu Ile Ser Asp Thr Phe Met Thr Leu Ala Ala Ile Ala Pro 290 295 300			912
ttg gcc gat gga ccc atc acg ata acc aac att ggt cat gca cgg ttg Leu Ala Asp Gly Pro Ile Thr Ile Thr Asn Ile Gly His Ala Arg Leu 305 310 315 320			960
aag gaa tcc gac cgc atc tca gcg atg gaa acc aac ctg cgc acg ctc Lys Glu Ser Asp Arg Ile Ser Ala Met Glu Thr Asn Leu Arg Thr Leu 325 330 335			1008
ggt gta caa acc gac gtc gga cac gac tgg atg aga atc tac ccc tct Gly Val Gln Thr Asp Val Gly His Asp Trp Met Arg Ile Tyr Pro Ser 340 345 350			1056
acc ccg cac ggc ggt aga gtg aat tgc cac cgg gac cac agg atc gct Thr Pro His Gly Gly Arg Val Asn Cys His Arg Asp His Arg Ile Ala 355 360 365			1104
atg gcg ttt tca atc ctg gga ctg aga gtg gac ggg att acc ctc gac Met Ala Phe Ser Ile Leu Gly Leu Arg Val Asp Gly Ile Thr Leu Asp 370 375 380			1152
gac cct caa tgc gtc ggg aag acc ttt cct ggc ttc ttc gac tac ctt Asp Pro Gln Cys Val Gly Lys Thr Phe Pro Gly Phe Phe Asp Tyr Leu 385 390 395 400			1200
gga cgc ctt ttc ccc gaa aag gcg ctt acg ctt ccc ggc tag Gly Arg Leu Phe Pro Glu Lys Ala Leu Thr Leu Pro Gly * 405 410			1242

gg 1244

<210> 33

<211> 413

<212> PRT

<213> Штучна послідовність

<220>

<223> grg23(L3P1.O23)

<400> 33

Met	Glu	Thr	Asp	His	Leu	Val	Ile	Pro	Gly	Ser	Lys	Ser	Ile	Thr	Asn
1				5					10					15	
Arg	Ala	Leu	Leu	Leu	Ala	Ala	Ala	Ala	Lys	Gly	Thr	Ser	Val	Leu	Val
			20					25					30		
Arg	Pro	Leu	Val	Ser	Ala	Asp	Thr	Ser	Ala	Phe	Lys	Thr	Ala	Ile	Gln
			35				40					45			
Ala	Leu	Gly	Ala	Asn	Val	Ser	Ala	Asp	Gly	Asp	Asn	Trp	Val	Val	Glu
	50					55				60					
Gly	Leu	Gly	Gln	Ala	Pro	His	Leu	Asp	Ala	Asp	Ile	Trp	Cys	Glu	Asp
65					70				75					80	
Ala	Gly	Thr	Val	Ala	Arg	Phe	Leu	Pro	Pro	Phe	Val	Ala	Ala	Gly	Gln
				85				90						95	

131						92622						132					
Gly	Lys	Phe	Thr	Val	Asp	Gly	Ser	Glu	Gln	Leu	Arg	Arg	Arg	Pro	Leu		
			100					105						110			
Arg	Pro	Leu	Val	Asp	Gly	Ile	Arg	His	Leu	Gly	Ala	Arg	Val	Ser	Ser		
			115					120						125			
Glu	Gln	Leu	Pro	Leu	Thr	Ile	Glu	Ala	Ser	Gly	Leu	Ala	Gly	Gly	Glu		
			130					135						140			
Tyr	Glu	Ile	Glu	Ala	His	Gln	Ser	Ser	Gln	Phe	Ala	Ser	Gly	Leu	Ile		
145					150					155					160		
Met	Ala	Ala	Pro	Tyr	Ala	Arg	Gln	Gly	Leu	Arg	Val	Arg	Ile	Pro	Asn		
				165					170					175			
Pro	Val	Ser	Gln	Pro	Tyr	Leu	Thr	Met	Thr	Leu	Arg	Met	Met	Arg	Asp		
			180					185						190			
Phe	Gly	Leu	Glu	Thr	Ser	Thr	Asp	Gly	Ala	Thr	Val	Ser	Val	Pro	Pro		
			195					200						205			
Gly	Arg	Tyr	Thr	Ala	Arg	Arg	Tyr	Glu	Ile	Glu	Pro	Asp	Ala	Ser	Thr		
								215			220						
Ala	Ser	Tyr	Phe	Ala	Ala	Ala	Ser	Ala	Val	Ser	Gly	Arg	Ser	Phe	Glu		
225					230					235					240		
Phe	Gln	Gly	Leu	Gly	Thr	Asp	Ser	Ile	Gln	Gly	Asp	Thr	Ser	Phe	Phe		
				245					250					255			
Asn	Val	Leu	Gly	Arg	Leu	Gly	Ala	Glu	Val	His	Trp	Ala	Pro	Asn	Ser		
			260					265						270			
Val	Thr	Ile	Ser	Gly	Pro	Glu	Arg	Leu	Asn	Gly	Asp	Ile	Glu	Val	Asp		
			275					280						285			
Met	Gly	Glu	Ile	Ser	Asp	Thr	Phe	Met	Thr	Leu	Ala	Ala	Ile	Ala	Pro		
								295			300						
Leu	Ala	Asp	Gly	Pro	Ile	Thr	Ile	Thr	Asn	Ile	Gly	His	Ala	Arg	Leu		
305					310					315					320		
Lys	Glu	Ser	Asp	Arg	Ile	Ser	Ala	Met	Glu	Thr	Asn	Leu	Arg	Thr	Leu		
				325					330					335			
Gly	Val	Gln	Thr	Asp	Val	Gly	His	Asp	Trp	Met	Arg	Ile	Tyr	Pro	Ser		
			340					345						350			
Thr	Pro	His	Gly	Gly	Arg	Val	Asn	Cys	His	Arg	Asp	His	Arg	Ile	Ala		
		355						360						365			
Met	Ala	Phe	Ser	Ile	Leu	Gly	Leu	Arg	Val	Asp	Gly	Ile	Thr	Leu	Asp		
								375			380						
Asp	Pro	Gln	Cys	Val	Gly	Lys	Thr	Phe	Pro	Gly	Phe	Phe	Asp	Tyr	Leu		
385					390					395					400		
Gly	Arg	Leu	Phe	Pro	Glu	Lys	Ala	Leu	Thr	Leu	Pro	Gly					
				405					410								


```

GRG23      ---MALERGQHGRSRRLFGASLERITMETDR-----LVIPGSKSITNRALLLAAAAGKT 51
GRG51      -----METDR-----LVIPGSKSITNRALLLAAAAGKA 28
B_Clausii  ----MVQFDSQARSPTWPLAGVERLRLTPSQKRINATLEVPGSKSATNRALLLAAVASGT 56
R_xylanophilus MSGVSGVPGVDFGIEEVRCFSPEEMEVAPLERPPDATVRLPGSKSITNRALLVAALAGGT 60
E_coli      -----MESLTLQPIARVDG--TINLPGSKTVSNRALLLAAALAHGK 38
CP4        -----MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGE 44
Zea_maize   -----AGAEIIVLQPIKEISG--TVKLPGSKSLSNRILLALLALSEGT 40
           : :*:.*: :*: :... : *

GRG23      SVLVRPLVSADTSAFKTAIQAL--GANVSADGDNWVVEGLGQAPHLAD-----IWCEDA 104
GRG51      SVLVRPLVSADTSAFKTAIQAL--GANVSADGDDWVVEGLGQAPNLDAD-----IWCEDA 81
B_Clausii  STLRLNALKSDDTYWCIEALKKT--GVEIAVDGNSNVTYVYGRGGVFHSGS-----LYIGSA 108
R_xylanophilus SRIENPLLADDPFWLMNALVGLGFGVRVGEAGAVEVAGGGGGIPAPSAD-----VFGVNA 115
E_coli      TVLTNLLDSDDVHRHMLNALTALGVSYTLSADRTRCEIIGNGGPLHAEGA---LELFLGNA 95
CP4        TRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVGNGLLAPEAP-----LDFGNA 99
Zea_maize   TVVDNLLNSEDVHYMLGALRTLGLSVEADKAAKRAVVVGCGGKFPVEDAKEEVQLFLGNA 100
           : : * . * * : * * : . *

GRG23      GTVARFLPPFVAAG--QGKFTVDGSEQLRRLRPLRPLVDGIRHLGARVSSE----QLPLTI 158
GRG51      GTVARFLPPFVAAG--QGKFTVDGSEQLRRLRPLRPLVDGIRHLGARVSSE----QLPLTI 135
B_Clausii  GTAGRFLPGMLAAA--TGNWHVEASHSMNKRPIAPLVKTLQALGANIQYGSRRGHYPLSI 166
R_xylanophilus GTVARFLPPALALG--SGPYRVDGTPRMRERPVAELVEALRALGARVECEEREHGLPLVV 173
E_coli      GTAMRPLAAALCLG--SNDIVLTGEPRMKERPIGHLVDALRLGGAKITYLEQENYPLRL 153
CP4        ATGCRLTMGLVG VY--DFDSTFIGDASLTRKPMGRVLNPLREMGVQVKSSEDG--DRLPVTL 156
Zea_maize   GTAMRPLTAAVTAAGGNATYVLDGVPMRERPIGDLVVGLKQLGADVDCFLGTDCPPVRV 160
           . * * : . . : .*: : : : * . : * : :


```

Φirypa 1A

```

GRG23      EA-SGLAGGEYEIEAHQSSQFASGLIMAAPYARQGLRVIRPN-PVSQPYLTMTLRMMRDF 216
GRG51      EA-SGLAGGEYEIEAHQSSQFASGLIMAAPYARQGLRVIRPN-PVSQPYLTMTLRMMRDF 193
B_Clausii  SG-EGLNGGKVNMSGQLSSQFISGCLLAAPLAKNPVSITVKDGIVQQAYVRITIDLMAAF 225
R_xylanophilus RG-GARGGGEISVSGERSQQFLSGLLISAPCLPGGLTVRPRGALVSRPYVDITVRVMRSF 232
E_coli      Q--GGFTGDNVDVDSVSSQFLTALLMTAPLAPEDTVIRIKGDLVSKPYIDITLNLMTF 211
CP4        RG--PKTPTPITYRVPMASAVKSAVLLAGLNTPGITTVIEPIMTRDTEKMLQGFGANL 214
Zea_maize   NGIGGLPGGKVKLSGSISSQYLSALLMAAPLALGDVEIEIIDKLISIPYVEMTLRLMERF 220
           : * . . : : * : : . :

GRG23      GLET--STDGATVSVPGR-YTAR-RYEIEPDASTASYFAAASAVSGRSFEFQGLGTDSI 272
GRG51      GIET--STDGATVSVPGR-YTAR-RYEIEPDASTASYFAAASAVSGRRFEFQGLGTDSI 249
B_Clausii  GVEVKAAPDWSLLEVNPS-PYAN-DIAIEADASTACYFLALAAITAGKIRIRHFSTKTS 283
R_xylanophilus GASVEEPPSGAARVAPCA-YRAT-AYRVEPDASAASYFLAAAALTAGRVVIPGLGRSSL 290
E_coli      GVEIEN-QHYQQFVVKGGQSYQSPGYTLVEGDASSASYFLAAAIAKGGTVKVTGIGRNSM 270
CP4        TVETDADGVRTIRLEGRGK--LTGQVIDVPGDPSSTAFPLVAALLVPG--SDVTILNLVLMN 271
Zea_maize   GVKAHSDSWDRFYIKGGQKYKSPKNAYVEGDASSASYFLAGAAITGGTVTVTEGCGTTSI 280
           . . : : *.*::: . : : . .

GRG23      QGDTSFNNVLGRLGAEVHWPNSVTISGP-----ERLNGDIEVDMG--EISDTFMTL 322
GRG51      QGDTSFNNVLGRLGAEVHWPNSVTISGP-----ERLTGDIEVDMG--EISDTFMTL 299
B_Clausii  QPDILFVSILKRMGCNFEIGPSFVEGEGP-----TRLRGGFTVNMN---ELSDQALTL 333
R_xylanophilus QGDVAFAGILRRMGCRVSLSEDRIELAGP-----PRLRG-VEADMN---AISDTMMTL 339
E_coli      QGDIRFADVLEKMGATICWGDDYISCT-----RGELNAIDMDMN---HIPDAAMTI 318
CP4        PTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPIL 331
Zea_maize   QGDVKFAEVLEMMGAKVTWTETSVTVTGPPREP-FGRKHLKAIDVNMN---KMPDVAMTL 336
           : * :*. . : . . : * :


```

Φirypa 1B

GRG23 AAIAPLADGPITITNIGHARLKESDRISAMETNLRITLGVQTDVGHDMRIYP-----S 375
 GRG51 AAIAPLADGPITITNIGHARLKESDRISAMESNLRMLGVQTDVGHDMRIYP-----S 352
 B_Clausii AAISPFADGPIAIEGVGHIRHHECDRIRAICTELSRIGIRVEERHDGLTVYP-----G 386
 R_xylanophilus AAIAPFASSPTLIKVAHTRLQETDRLAAVAELSRLGVRVHETPDSLRIIP-----G 392
 E_coli ATAALFAKGTTRLRNIYNWRVKETDRLFAMATELRKVGAEEVEEGHDYIRITPP-----E 372
 CP4 AVAAAFAGATVMNGLEELRVKESDRLSAVANGLKLVGDCDEGETSLVVRGRPDGKGLG 391
 Zea_maize AVVALFADGPTAIRDVASWRVKETERMVAIRTELTKLGASVEEGPDYCIITPP-----E 390

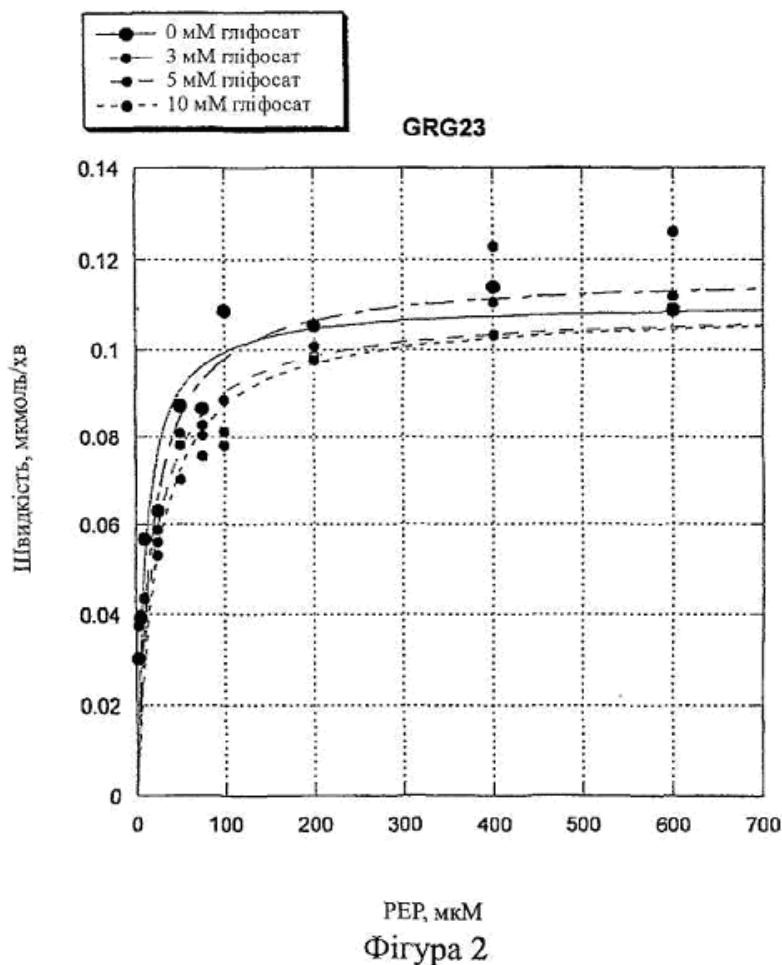
. : :. . . : . : * : * : * : * * . :

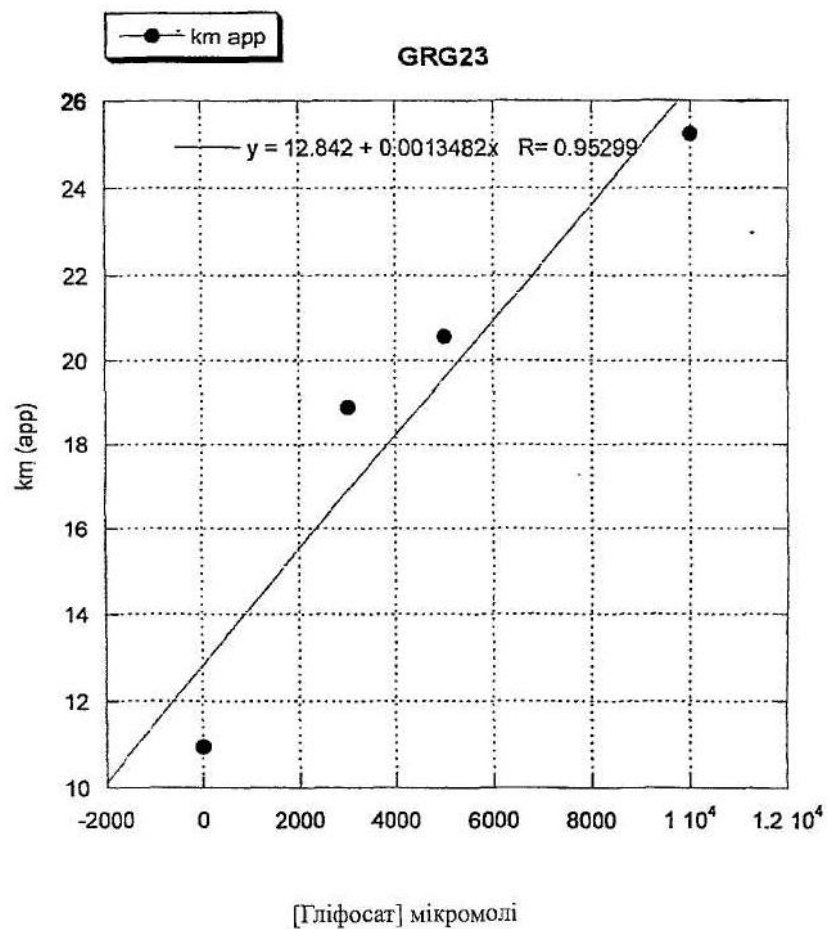
GRG23 TPHGGRVNCNRDHRIMAFSILGLR-VDGITLDDPQCVCVGTFFPGFFDYLGRLFPEKALT 434
 GRG51 TPHGGRVNCNRDHRIMAFSILGLR-VDGITLDDPQCVCVGTFFPGFFDYLGRLFPEKALT 411
 B_Clausii QPKPTVVNTYDDHRMAMALALIGAK-VDGIELDDPGCVAKTCCPSYFSMLAQTGIGVKA 445
 R_xylanophilus KVRPAAIRTYGDHRMAMAFSLVGLR-VRGVRILDPGCVTTKTLPGYFRLLLEGLRRGG--- 447
 E_coli KLNFAEITATYNDHRMAMCFSLVALS-DTPVTILDPKCTAKTFPDYFEQLARISQAA--- 427
 CP4 NASGAAVATHLDHRIMAFSLVGLSENPTVDDATMIATSFPEFMDLMAGLGAKIELSD 451
 Zea_maize KLVNTAIDTYDDHRMAMAFSLAACA-EVPVTIRDPGCTRKTFPDYFDVLSTFVKN--- 444

: : ***:**.: : . : : * . . : : * : :

GRG23 PG-- 436
 GRG51 PG-- 413
 B_Clausii P-- 446
 R_xylanophilus ---
 E_coli ---
 CP4 TKAA 455
 Zea_maize ---

Фигура 1С





Фігура 3