



УКРАЇНА

(19) UA

(11) 120590

(13) C2

(51) МПК

C12N 15/82 (2006.01)

C12N 15/113 (2010.01)

C07K 14/415 (2006.01)

МІНІСТЕРСТВО РОЗВИТКУ  
ЕКОНОМІКИ, ТОРГІВЛІ ТА  
СІЛЬСЬКОГО ГОСПОДАРСТВА  
УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА ВІНАХІД

(21) Номер заявки: а 2015 12799  
(22) Дата подання заявки: 06.06.2014  
(24) Дата, з якої є чинними права на винахід: 10.01.2020  
(31) Номер попередньої заявки відповідно до Паризької конвенції: 10 2013 010 026.7  
(32) Дата подання попередньої заявки відповідно до Паризької конвенції: 17.06.2013  
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку: DE  
(41) Публікація відомостей про заявку: 25.05.2016, Бюл.№ 10  
(46) Публікація відомостей про видачу патенту: 10.01.2020, Бюл.№ 1  
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ: PCT/DE2014/000310, 06.06.2014  
(72) Винахідник(и):  
Торйек Отто (DE),  
Боршардт Дітріх (DE),  
Мешелке Вольфганг (DE),  
Лейн Йєнс Крістоф (DE)  
(73) Власник(и):  
КВС СААТ СЕ,  
Grimsehlstr. 31, 37555 Einbeck, Germany (DE)

(74) Представник:  
Бенатов Даніель Емілович, реєстр. №224  
(56) Перелік документів, взятих до уваги експертизою:  
GIDNER SARA ET AL. QTL mapping of BNYVV resistance from the WB41 source in sugar beet. GENOME. 04.2005, vol. 48cl, № 2, P. 279-285  
LEIN JENS CHRISTOPH ET AL. Resistance gene analogues are clustered on chromosome 3 of sugar beet and cosegregate with QTL for rhizomania resistance. GENOME. 01.2007, vol. 50, № 1, P. 61-71  
TIAN YANYAN ET AL. The absence of TIR-type resistance gene analogues in the sugar beet (Beta vulgaris L.) genome. JOURNAL OF MOLECULAR EVOLUTION. 01.2004, vol. 58, № 1, P. 40-53  
GRIMMER M K ET AL. An anchored linkage map for sugar beet based on AFLP, SNP and RAPD markers and QTL mapping of a new source of resistance to Beet necrotic yellow vein virus. THEORETICAL AND APPLIED GENETICS; INTERNATIONAL JOURNAL OF PLANT BREEDING RESEARCH, SPRINGER, BERLIN, DE. 09.02.2007, vol. 114, № 7, P. 1151-1160  
BUTORINA A K ET AL. Molecular genetic investigation of sugar beet (L.). RUSSIAN JOURNAL OF GENETICS, NAUKA/INTERPERIODICA, MO. 08.10.2011, vol. 47, № 10, P. 1141-1150  
AMIRI R ET AL. A new RAPD marker for beet necrotic yellow vein virus resistance gene in Beta vulgaris. BIOLOGIA PLANTARUM, KLUWER ACADEMIC PUBLISHERS, DO. 21.03.2009, vol. 53, № 1, P. 112-119  
PAVLI OURANIA I ET AL. Achievements and prospects in breeding for rhizomania resistance in sugar beet. FIELD CROPS RESEARCH. 06.2011, vol. 122, № 3, P. 165-172

(54) МОЛЕКУЛА НУКЛЕІНОВОЇ КИСЛОТИ, ЯКА ЗУМОВЛЮЄ РЕЗИСТЕНТНІСТЬ ДО ВІРУСУ НЕКРОТИЧНОГО ПОЖОВТІННЯ ЖИЛОК БУРЯКА

(57) Реферат:

UA 120590 C2

Винахід стосується молекули нуклеїнової кислоти, що кодує поліпептид, здатний зумовлювати резистентність до вірусу некротичного пожовтіння жилок буряка у рослини роду буряків, де молекула нуклеїнової кислоти містить нуклеотидну послідовність до SEQ ID NO: 1, яка кодує амінокислотну послідовність SEQ ID NO: 2 або SEQ ID NO: 3, вектора, клітини-хазяїна, поліпептиду, трансгенної рослини, які містять вказану молекули нуклеїнової кислоти, способу одержання та способу ідентифікації вказаної молекули нуклеїнової кислоти.

Область техніки, до якої належить винахід.

Винахід належить до молекули нуклеїнової кислоти, яка кодує поліпептид, здатний зумовлювати стійкість до патогену, а саме до "вірусу некротичного пожовтіння жилок буряка" (BNYVV) у рослин, а саме у рослин з роду буряків, де експресується цей поліпептид. Цей винахід також належить до самого поліпептиду, який здатний зумовлювати стійкість до патогену, а саме до BNYVV у рослин роду буряків, де експресується цей поліпептид, що кодується молекулою нуклеїнової кислоти, до якої належить цей винахід. Також винахід належить до трансгенної рослини, рослинної клітини, органу, тканини, частини або насіння, що містить молекулу нуклеїнової кислоти або її частини, а також до способів одержання трансгенних рослин або рослинних клітин цього типу. Винахід також включає способи для виявлення молекули нуклеїнової кислоти, яка зумовлює стійкість, та способи селекції рослин або рослинних клітин, що міститимуть молекулу нуклеїнової кислоти, яка зумовлює стійкість.

Передумови для винаходу

Ризоманія є найбільш серйозним захворюванням цукрових буряків у світі, яке впливає на прибутковість та може спричинити втрати доходів від врожайності у розмірі 50 % та більше. Захворювання, що також описується як "розростання кореня", спричиняється "вірусом некротичного пожовтіння жилок буряка" (BNYVV) та передається найпростішими виду *Polymyxa betae* через ґрунт. Інфекція проявляється підвищеною проліферацією тонких та вторинних коренів, а також формуванням значно зменшеного тіла основного кореня зі зниженим вмістом цукру. Інфіковані рослини слабо поглинають воду і, таким чином, стають більш схильними до висихання. Розповсюдження інфекції по всій рослині спричиняє пожовтіння жилок листків, появу на них некротичних вогнищ та жовтих плям. Оскільки повне вилікування рослини неможливе, так само як і у випадку інших вірусних захворювань, пошкоджень можна уникнути лише культивуючи стійкі види. Фактично, на теперішній момент вивчаються три основних гени проти ризоманії: RZ-1 (інша назва "Holly"), RZ-2 та RZ-3. Також в літературі описані і інші гени, що зумовлюють стійкість до ризоманії, проте, вони мають менше значення. Наразі ген резистентності RZ-1 вже включено до більшості гібридних ліній (батьківське насіння та/або компоненти опилувача батьківських форм). Однак, було виявлено, що стійкість, зумовлена геном RZ-1, є недостатньою у сильно заражених регіонах або в регіонах, де наявні різноманітні типи BNYVV (наприклад, роботи Sohi та Maleki, 2004). Тому, деякий час назад було запропоновано комбінувати RZ-1, наприклад, з RZ-2 або RZ-3. RZ-2 та RZ-3 мають походження з виду Буряк звичайний, підвиду приморський (*Beta vulgaris* subsp. *maritima*) (WB42, WB41), та знаходяться в тій же частині хромосоми 3 на генетичній мапі геному цукрового буряка, що і RZ-1, котрий, однак, знаходиться нижче від RZ-2 та RZ-3. Scholten та співавт. (1999) визначили відстань між головними RZ-генами RZ-1 та RZ-2, яка склала 20-25 сантиморганід (cM). Gidner та співавт. (2005) отримали коротшу відстань, а саме 5 cM між RZ-1 та RZ-2, а також не підтвердили, що RZ-2 та RZ-3 розташовані в одному й тому ж локусі. Schmidlin та співавт. (2008) за допомогою аналізу експресії у інфікованих буряках визначали різно-індуковані гени, які, однак, не відповідали генам RZ-2 та RZ-3. У дослідженні Larson та співавт. (2008) у цукрових буряках за допомогою методу MALDI-TOF-MS були виявлені деякі білки, індуковані BNYVV, однак, ці білки, що кодувалися RZ-1, RZ-2 або RZ-3, вчені ідентифікувати не змогли. Крім того, ця послідовність навколо гену резистентності є повторюваною, що робить розробку діагностичних маркерів особливо складною. До теперішнього часу, специфічно для генів стійкості до ризоманії не було ні високороздільних маркованих генетичних карт, ні верифікованих генів-кандидатів. Також, дотепер не була повністю відома функціональна основа, тобто генетична структура цих генів стійкості.

Для ефективної культивації сортів, стійких до ризоманії, здатних протистояти ризику появи окремих особин, в яких відсутня стійкість до BNYVV, необхідно постійно виявляти нові гени стійкості та інтегрувати їх до пулу генів у врожайних культурах, таких як цукровий буряк.

Суть винаходу

Даний винахід створено на основі описаного вище рівня техніки. Задачею цього винаходу є виявлення молекули нуклеїнової кислоти та/або поліпептиду, здатного зумовлювати стійкість рослини до ризоманії. Також задачею винаходу є одержання трансгенної рослини, стійкої до ризоманії, та розробка способу її розведення. Ще однією задачею цього винаходу є розробка та використання молекулярних маркерів, що дозволяють ефективно культивувати стійкі до ризоманії рослини та виводити нові стійкі лінії рослин.

Варіанти реалізації цього винаходу, які вирішують поставлену в основу винаходу задачу, базуються на точному генетичному картуванні, ідентифікації, ізоляції та описі характеристик гену, що походить із донорського виду, а саме виду *Beta vulgaris* subsp. *maritima*, та кодує

поліпептид або білок, здатний зумовлювати стійкість до патогену у рослини, в якій експресується цей поліпептид.

Деякі терміни, використані в цій заявці, будуть вперше пояснені та деталізовані далі:

Термін "приблизно" у поєднанні з визначенням довжини нуклеотидної послідовності означає відхилення до  $\pm 200$  пар азотистих основ, бажано до  $\pm 100$  пар азотистих основ та найбільш бажано до  $\pm 50$  пар азотистих основ.

"Рослина роду Буряк" відноситься до родини амарантових (Amaranthaceae). Ці рослини включають рослини видів Буряк великоплідний (*Beta macrocarpa*), Буряк звичайний (*Beta vulgaris*), *Beta lomatogona*, *Beta macrohiza*, Буряк вінчиоцвітний (*Beta corolliflora*), Буряк трьохматочковий (*Beta trigyna*) та Буряк карликовий (*Beta nana*). Рослини роду Буряк звичайний (*Beta vulgaris*) - це здебільшого рослини підвиду приморський (*Beta vulgaris* subsp. *maritima* (Seemangold)) або підвиду столовий (*Beta vulgaris* subsp. *vulgaris*). Вони, включають, наприклад, *Beta vulgaris* subsp. *vulgaris* var. *altissima* (цукровий буряк у більш вузькому розумінні), *Beta vulgaris* ssp. *vulgaris* var. *vulgaris* (Мангольд), *Beta vulgaris* ssp. *vulgaris* var. *conditiva* (червоний буряк), *Beta vulgaris* ssp. *vulgaris* var. *crassa/alba* (буряк кормовий).

Під терміном "гібридизувати" або "гібридизація" мається на увазі процес, в якому одониткова молекула нуклеїнової кислоти прикріплюється до іншої нитки нуклеїнової кислоти, що є у максимально можливій мірі комплементарною, тобто формує пари азотистих основ. Стандартні методи гібридизації описані, наприклад, у Sambrook та співавт., 2001. Це означає, що бажано, щоб не менше 60 % азотистих основ у молекулі нуклеїнової кислоти формували пари з основами нитки нуклеїнової кислоти, комплементарної у максимально можливій мірі, більш бажано - не менше 65 %, 70 %, 75 %, 80 % або 85 %, найбільш бажано 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % або 99 % азотистих основ. Вірогідність такого збігу залежить від того, наскільки жорсткі умови гібридизації дотримуються.

Термін "жорсткість" відноситься до умов гібридизації. Так, високий ступінь жорсткості означає, що парування азотистих основ є контрольованим, а низький ступінь жорсткості означає, що парування азотистих основ є послабленим. Жорсткість умов гібридизації, зокрема, залежить від концентрації солі, сили іонів чи температури. Загалом, ступінь жорсткості може бути збільшений шляхом підвищення температури та/або зниження вмісту солі. "Жорсткі умови гібридизації" - це умови, при яких гібридизація відбувається переважно лише між гомологічними молекулами нуклеїнових кислот. Термін "умови гібридизації" в цьому випадку стосується не лише умов, які переважають під час фактичного приєднання нуклеїнових кислот, але також стосується умов, що переважають під час наступних стадій відмивання. Жорсткими є умови гібридизації, при яких, наприклад, гібридизуються переважно лише ті молекули нуклеїнових кислот, що мають не менше 70 %, бажано не менше 75 %, 80 %, 85 % або 90 %, особливо бажано не менше 95 %, 96 %, 97 %, 98 % або 99 % ідентичних послідовностей. Приклади жорстких умов гібридизації: гібридизація у 4 x SSC при 65 °C та наступне багаторазове промивання в 0.1 x SSC при 65 °C протягом приблизно 1 години. Термін "жорсткі умови гібридизації" в цьому випадку може також означати наступне: гібридизація при 68 °C в 0.25 M фосфату натрію, pH 7.2, 7 % додецил сульфату натрію, 1 mM EDTA та 1 % BSA протягом 16 годин та наступне дворазове промивання у 2 x SSC та 0,1 % додецил сульфату натрію при 68 °C. Гібридизація переважно проводиться при жорстких умовах.

"Ізольована молекула нуклеїнової кислоти" - це молекула нуклеїнової кислоти, виділена з її природного або первинного середовища. Цей термін також включає молекули нуклеїнових кислот, що створені синтетично.

"Ізольований поліпептид" - це поліпептид, що виділений зі свого природного або первинного середовища. Цей термін також включає синтетично створені поліпептиди.

"Молекулярний маркер" - це нуклеїнова кислота, що є поліморфною у популяції рослин. Цей маркер, таким чином, може виявляти та диференціювати різні алельні стани (алелі). Відомі аналітичні методи, що використовуються для цієї мети, наприклад: RFLP, AFLP, SNP, SSR або KASP. Термін "молекулярний маркер" відноситься до послідовності нуклеотидів, що є комплементарними, принаймні у максимально можливій мірі комплементарними або гомологічними до геномних послідовностей, наприклад, нуклеїнові кислоти, що використовуються як зонди або праймери. Маркери, за допомогою яких можна описати поліморфізм, можуть бути визначені за допомогою напрацьованих методів. Такі методи включають наприклад, ПЛР-ампліфікацію, специфічну до послідовності, визначення поліморфізму довжини рестрикційних фрагментів (RFLPs), визначення поліморфізму полінуклеотидів за допомогою алель-специфічної гібридизації (ASH), визначення ампліфікованої варіаційної послідовності геному рослини, визначення самореплікації послідовності, визначення простих повторів послідовності (SSRs), визначення одиночних



нуклеотидних поліморфізмів (SNPs), або визначення ампліфікованих поліморфізмів довжини фрагменту (AFLPs). Крім того, відомі методи визначення маркерів експресованої послідовності (ESTs) та маркерів простих повторів послідовності (SSR), що є похідними від методів визначення маркерів експресованої послідовності (EST) та випадково ампліфікованої поліморфної ДНК (RAPD).

"Промотор" - це регуляторна нетрансльована послідовність ДНК, типово розташована в області кодування згори, складається зі зв'язуючої точки для РНК-полімерази та ініціює транскрипцію ДНК.

"Патоген" - це організм, що при взаємодії з рослиною призводить до виникнення симптомів захворювання в одному або більше органів рослини. Патогени включають в себе, наприклад, тваринні, грибі, бактеріальні, вірусні організми або ооміцети.

"Інфекція, спричинена патогеном" - це найбільш ранній момент, коли патоген вступає у взаємодію з тканинами рослини-хазяїна. Наприклад, у випадку вірусного патогену BNYVV, хвороба передається найпростішим Polymyxa betae. Polymyxa формує спори, що можуть виживати у ґрунті протягом багатьох місяців. Вірус також живе у цих спорах. Коли ці тимчасово неактивні спори проростають та формують рухомі зооспори, вірус отримує змогу перейти через спори у клітини тканин рослини-хазяїна, та взаємодіяти з ними. (Esser 2000).

"Органи" рослини - це, наприклад, листя, вісь росту, стебло, коріння, гіпокотиль (підсім'ядольне коліно), вегетативні бруньки, меристеми, зав'язі, пильники, сім'я та плоди.

"Частини" рослини - це сукупність кількох органів, наприклад, квітки та насінини, або частина органа, наприклад, зріз стебла.

"Тканини" рослини - це, наприклад, калюс, паренхіма, меристемна тканина, тканина листка, стебла, кореня, пухлинна тканина рослини або репродуктивна тканина.

"Клітини" рослини - це, наприклад, ізольовані клітини з клітинною стінкою, їх агрегати або протопласти.

Термін "стійкість" має широке значення, що охоплює сукупність засобів захисту від затримки пригнічення розвитку хвороби. Важливим прикладом є патоген вірусу некротичного пожовтіння жилок буряка. Згідно цього винаходу, стійкі клітини або рослини переважно стають стійкими до BNYVV. Сстійкість проти патогену прирівнюється до стійкості проти хвороби, спричиненої цим патогеном, наприклад, стійкість проти BNYVV прирівнюється до стійкості проти ризоманії.

Термін "Трансгенна рослина" в цьому випадку стосується рослини, у геномі якої є принаймні одна інтегрована нуклеїнова кислота. В цьому випадку це може бути гетерологічна нуклеїнова кислота. Бажано, щоб нуклеїнова кислота була інтегрована стабільно. Це означає, що інтегрована нуклеїнова кислота стабільно утримується у рослині, може експресуватися, а також може стабільно передаватися нащадкам.

Цей винахід описує молекулу нуклеїнової кислоти, що кодує поліпептид, здатний зумовлювати стійкість до патогену у рослині, в якій експресується цей поліпептид. Молекула нуклеїнової кислоти включає нуклеотидну послідовність, відібрану за наступними параметрами:

а) нуклеотидна послідовність, що кодує поліпептид, який містить амінокислотну послідовність згідно SEQ ID NO: 2 або SEQ ID NO: 3,

б) нуклеотидна послідовність, що містить кодуючу послідовність ДНК згідно SEQ ID NO: 1,

в) нуклеотидна послідовність, що гібридується з комплементарною послідовністю нуклеотидів згідно а) чи б) в жорстких умовах.,

г) нуклеотидна послідовність, яка кодує поліпептид, що виходить після заміщення, делеції та /або приєднання однієї чи більше амінокислот у послідовності, кодованої послідовністю нуклеотидів згідно а) або б), з поліпептиду, кодованого послідовністю нуклеотидів згідно а) чи б),

д) нуклеотидна послідовність, що кодує поліпептид, який має послідовність амінокислот не менш ніж на 60 % ідентичну послідовності амінокислот, що кодується послідовністю нуклеотидів згідно а) чи б), або

е) послідовність нуклеотидів, що кодує принаймні один нуклеотид-зв'язуючий домен (NBS або NB-ARC) відповідно до позицій амінокислот 168-227 у SEQ ID NO: 2 або 182-241 у SEQ ID NO: 3, щонайменше один лейцин-багатий домен (LRR) відповідно до позицій амінокислот 591-613 у SEQ ID NO: 2 або 605-627 у SEQ ID NO: 3 та/або принаймні один внутрішній повторюваний домен (IR) відповідно до позицій амінокислот 1013-1072 у SEQ ID NO: 2 або 1027-1086 у SEQ ID NO: 3.

Молекула нуклеїнової кислоти може бути ізольованою. Бажано, щоб це була ДНК, особливо бажано - кДНК (кодуюча ДНК). Поліпептид, що кодується молекулою нуклеїнової кислоти, описаної в цьому винаході, переважно зумовлює стійкість до вірусного патогену, а саме до "Вірусу некротичного пожовтіння жилок буряка" (BNYVV), який зумовлює захворювання рослини

на ризоманію. Крім того, поліпептид, що кодується молекулою нуклеїнової кислоти, як описано в цьому винаході, особливо у рослин роду буряків, зумовлює стійкість до патогену. Описані рослини - це, як правило, рослини видів *Beta vulgaris*, а саме рослини підвидів *Beta vulgaris subsp. maritima* або *Beta vulgaris subsp. Vulgaris*; вони включають, зокрема, врожайні типи, наприклад, цукровий буряк, червоний буряк, кормовий буряк, листовий буряк, буряк Мангольд.

Відповідно до цього винаходу, в одному з варіантів реалізації молекули нуклеїнової кислоти вона включає в себе послідовність нуклеотидів відповідно до пункту а). Послідовність амінокислот у поліпептиді, що кодується за послідовністю SEQ ID NO: 2, та/або поліпептиді, що кодується за послідовністю SEQ ID NO: 3, утворює білок резистентності гену RZ-3. В цьому випадку це ген, що кодує білок резистентності типу NBS-LRR, який характеризується певною структурною послідовністю. Загальна структура таких білків резистентності у рослин була добре вивчена (Martin та співавт.2003). Проте, розкриття принципу структурної будови саме домену, відомого як LRR (лейцин-збагачений домен), який потенційно є доменом для ідентифікації здебільшого невідомих патогенетичних чинників, у найближчому майбутньому не передбачається. Таким чином, ідентифікація гену, що кодує стійкість до BNYVV або відповідного білка на основі тільки наявних знань, неможлива. Ідентифікація гену резистентності RZ-3 відбувалася шляхом клонування на основі генетичних карт, що вимагало ретельного генетичного картування та точного картування таргетної області, де спочатку передбачалося розміщення гену резистентності RZ-3. Процес розробки буде описано детальніше далі в цьому документі.

Ідентифікований білок резистентності належить до типу NBS-LRR та має нуклеотид-зв'язуючий домен (NBS, також відомий як NB-ARC) (адапторами для зв'язування нуклеотидів є фактор активації протеаз апоптозу-1 (APAF-1)), R білки та CED-4)), що відповідає позиціям амінокислот 168-227 у послідовності SEQ ID NO: 2 або 182-241 у послідовності SEQ ID NO: 3, лейцин-багатий домен (LRR), що відповідає позиціям амінокислот 591-613 у SEQ ID NO: 2 або 605-627 у SEQ ID NO: 3 та/або принаймні один внутрішній повторюваний домен (IR; внутрішній домен-повтор), що відповідає позиціям амінокислот 1013-1072 у SEQ ID NO: 2 або 1027-1086 у SEQ ID NO: 3. Домен NBS кодується нуклеотидами 2019-2882 у SEQ ID NO: 1, домен LRR кодується нуклеотидами 3288-3356 у SEQ ID NO: 1, а домен IR кодується нуклеотидами 4554-4871 у SEQ ID NO: 1. Домен NB-ARC є центральним доменом, що зв'язує нуклеотиди. Скоріше за все це функціональний домен АТФази, що ймовірно регулює активність білка резистентності. Домен NB-ARC складається з трьох субдоменів: NB, ARC1 та ARC2. Характерною ознакою домену NB-ARC є наявність фактора активації протеаз апоптозу-1 (APAF-1), що ймовірно відповідає за реакцію гіперчутливості, а також мотиви hhGRExE, Walker-A- або P-loop, Walker-B, GxP, RNBS- від A до D та MHD (Ooijen et al., 2008). Деякі із вказаних мотивів на сьогоднішній день вже можуть бути ідентифіковані. У подальших варіантах реалізації молекули нуклеїнової кислоти, як показано в цьому винаході, молекула нуклеїнової кислоти містить послідовність нуклеотидів згідно до пункту б). Послідовність нуклеотидів включає кодує послідовності ДНК відповідно до SEQ ID NO: 1, що кодує послідовності амінокислот відповідно до SEQ ID NO: 2 та SEQ ID NO: 3.

У наступному варіанті здійснення молекули нуклеїнової кислоти, молекула нуклеїнової кислоти містить послідовність нуклеотидів відповідно до пункту г). Ця послідовність нуклеотидів кодує поліпептид, що є похідним поліпептиду, який кодований послідовністю нуклеотидів відповідно до а) або б). Цей похідний поліпептид представляє собою отриману послідовність амінокислот, що містить принаймні одну заміну, делецію або приєднання однієї або більше амінокислот у функціонально значимій частині кодованого поліпептиду/білка. Якщо відбувається заміна однієї амінокислоти іншою амінокислотою, що має ті самі, еквівалентні чи подібні хіміко-фізичні властивості - це називається "консервативна заміна" або "напівконсервативна заміна". Приклади фізико-хімічних властивостей амінокислоти - це, наприклад, гідрофобність або заряд. Спеціалісту в цій галузі відомо, яка заміна амінокислоти є консервативною, а яка - напівконсервативною. Загальне знання галузі дає змогу спеціалісту визначити, ідентифікувати та виявити делеції амінокислот та їх приєднання, що є шкідливими для функціональності білка резистентності RZ-3, а також ті позиції, в яких це можливо. Спеціалісту в цій галузі відомо, що в разі модифікацій у послідовності амінокислот білка NBS-LRR (заміни, делеції або приєднання однієї чи більше амінокислот), функціональність вищевказаного збереженого домена повинна залишитися незмінною, і тому в цих доменах можлива лише обмежена кількість модифікацій вищезгаданого типу. Послідовність нуклеотидів для цих варіантів реалізації винаходу кодує похідні або отримані послідовності амінокислот, коли послідовність нуклеотидів гомологічна або ідентична до послідовності нуклеотидів відповідно до пунктів а) або б) принаймні на 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 92 %, 95 %, 96 %, 97 %, 98 %, 99 %.

94 %, 96 %, 97 %, 98 %, 99 %. Такі послідовності нуклеотидів, що кодують похідні або отримані послідовності амінокислот, переважно можуть бути виготовлені прямим або непрямим способом (наприклад, шляхом ампліфікації або послідовних реплікацій) із початкової послідовності нуклеотидів, що відповідає всій довжині або принаймні частині послідовності SEQ ID NO: 1 чи іншій послідовності, описаній у цьому винаході.

Також варіантом реалізації цього винаходу є молекула нуклеїнової кислоти, що містить послідовність нуклеотидів відповідно до пункту д). Ця послідовність нуклеотидів кодує поліпептид, що містить послідовність амінокислот, ідентичну до послідовності амінокислот, яка кодується послідовністю нуклеотидів відповідно до а) або б) принаймні на 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 92 %, 94 %, 96 %, 97 %, 98 %, 99 %.

Іншим варіантом реалізації цього винаходу є отримання молекули нуклеїнової кислоти, що містить послідовність нуклеотидів відповідно до пункту е). Така послідовність нуклеотидів кодує принаймні один нуклеотид-зв'язуючий домен (NBS), що відповідає позиціям амінокислот 168-227 у послідовності SEQ ID NO: 2 або 182-241 у послідовності SEQ ID NO: 3, принаймні один лейцин-збагачений домен (LRR) відповідно до позицій амінокислот 591-613 у послідовності SEQ ID NO: 2 або 605-627 у послідовності SEQ ID NO: 3 та/або принаймні один внутрішній повторюваний домен (IR) відповідно до позицій амінокислот 1013-1072 у послідовності SEQ ID NO: 2 або 1027-1086 у послідовності SEQ ID NO: 3. Послідовність нуклеотидів переважно відповідає за кодування поліпептиду, що містить принаймні один нуклеотид-зв'язуючий домен (NBS) відповідно до позицій амінокислот 168-227 у SEQ ID NO: 2 або 182-241 у SEQ ID NO: 3, принаймні один лейцин-збагачений домен (LRR) відповідно до позицій амінокислот 591-613 у SEQ ID NO: 2 або 605-627 у SEQ ID NO: 3 та принаймні один внутрішній повторюваний домен (IR) відповідно до позицій амінокислот 1013-1072 у SEQ ID NO: 2 або 1027-1086 у SEQ ID NO: 3. Ці домени, як правило, організовані у поліпептидах послідовно від N- до C- кінця у порядку NBS-LRR-IR, де між доменами в окремих випадках можуть бути присутні одна або більше амінокислот.

Цей винахід також належить до поліпептиду, який здатний зумовлювати стійкість до патогену у рослині, де експресується цей поліпептид, та який кодується молекулою нуклеїнової кислоти, описаною у цьому винаході. При цьому патоген переважно представлений BNYVV та/або рослини переважно представлені родом буряків, а саме рослинами виду *Beta vulgaris* (буряк столовий). Поліпептид переважно має амінокислотну послідовність відповідно до SEQ ID NO: 2 або до SEQ ID NO: 3. Поліпептид може бути ізольованим.

Ще один аспект цього винаходу належить до вектору, що містить молекулу нуклеїнової кислоти, як описано у винаході. Вектором може бути плазміда, косміда, фаг або експресійний вектор, трансформаційний вектор, човниковий вектор, клонуючий вектор, це також може бути двонитковий або одонитковий, лінійний чи колоподібний вектор, він також може трансформувати прокаріотичних або еукаріотичних хазяїв шляхом інтеграції у їхній геном або позахромосомним шляхом. Як показано у цьому винаході, молекула нуклеїнової кислоти переважно функціонально зв'язана з однією чи більше регуляторними послідовностями в експресійному векторі, що робить можливою транскрипцію та, в деяких випадках, експресію у прокаріотичній або еукаріотичній клітині-хазяїні. Як приклад, можна навести молекулу нуклеїнової кислоти, яка знаходиться під контролем відповідного промотора або термінатора. Придатними можуть бути промотори, що постійно індукуються (наприклад, промотор 35S вірусу мозаїки цвітної капусти (Odell та співавт.1985), а особливо придатними є промотори, що індукуються патогеном (наприклад, промотор PR1 петрушки (Rushton та співавт., 1996). Особливо підходять патоген-індуковані синтетичні або гібридні промотори, яких немає в природі, сформовані із певних елементів та які містять мінімальний промотор, а також у положенні вище мінімального промотора мають принаймні один цис-регуляторний елемент, що служить зв'язуючою точкою для спеціальних факторів транскрипції. Гібридні промотори розробляються у відповідності до необхідних вимог та індукуються або репресуються різними факторами. Приклади таких промоторів можна знайти у WO 00/29592, WO 2007/147395 та WO 2013/091612. Придатним термінатором є, наприклад, термінатор нопалін синтази (nos) (Depicker et al., 1982).

Крім вищеописаних векторів, цей винахід також пропонує метод включення описаного вектора до клітини-хазяїна. Вектор може бути включений, наприклад, за допомогою кон'югації, мобілізації, біолістичної трансформації, трансформації за допомогою агробактерій, трансфекції, трансдукції, вакуумної інфільтрації або електропорації. Спеціаліст у цій сфері має інформацію про ці методи, а також про методи підготовки описаних векторів (Sambrook та співавт.2001).

Також цей винахід пропонує включення до клітини-хазяїна молекули нуклеїнової кислоти або вектору, описаного у винаході вище. Клітина-хазяїн у сенсі цього винаходу може бути

прокаріотичною (наприклад, бактерія) або еукаріотичною (наприклад, клітина рослини або дріжджового грибка). Клітина-хазяїн - це переважно агробактерія, наприклад, *Agrobacterium tumefaciens* або *Agrobacterium rhizogenes*, або рослинна клітина, яка містить молекулу нуклеїнової кислоти, як описано у цьому винаході. Численні методи, наприклад, кон'югація або електропорація є відомими для спеціаліста у цій галузі, таким чином, обізнаний спеціаліст може включити молекулу нуклеїнової кислоти або вектор, описані в цьому винаході, у агробактерію; а методи, похідні від трансформації (біолістична трансформація, агробактеріо-опосередкована трансформація) є також відомими для обізнаного фахівця, таким чином такий фахівець може включити молекулу нуклеїнової кислоти або вектор, описані в цьому винаході, до рослинної клітини (Sambrook та співавт.2001).

Ще один аспект цього винаходу стосується трансгенної рослинної клітини, що містить в якості трансгена молекулу нуклеїнової кислоти або вектор, описані в цьому винаході. Трансгенна рослинна клітина цього типу - це, наприклад, рослинна клітина, що була трансформована, бажано стабільно, та має молекулу нуклеїнової кислоти або вектор, як це описано у винаході. Більш бажаним варіантом реалізації є трансгенні рослинні клітини з оперативним приєднання молекули нуклеїнової кислоти до однієї або більшої кількості регуляторних послідовностей, що дозволяє транскрипцію та, в деяких випадках, експресію у рослинній клітині. Таким чином, згідно цього винаходу, уся подальша побудова молекули нуклеїнової кислоти та регуляторних послідовностей містить трансген. Такі регуляторні послідовності - це, наприклад, промотори або термінатори. Численні функціональні промотори або термінатори, що підходять для рослин, є добре відомими обізнаному у цій галузі фахівцю. Трансгенна рослинна клітина, як описано у цьому патенті, це, зокрема, клітина рослини роду буряків, яка переважно демонструє більшу стійкість до патогену, а саме до ВНПЖБ, ніж відповідна нетрансформована клітина рослини (рослинна клітина, що не містить трансгену). Рівень стійкості, наприклад, до ВНПЖБ у рослин роду буряків можна оцінити якісно шляхом визначення показників за відповідними шкалами (шкали для рослин роду буряків є доступними з попередніх робіт, наприклад, у Mechelke (1997) для цукрового буряка). Вища стійкість проявляється у покращенні резистентності принаймні за однією шкалою, принаймні за двома шкалами, або принаймні за трьома або більше шкалами. У цьому винаході, крім того, описано метод отримання трансгенних рослинних клітин, включаючи стадію включення молекули нуклеїнової кислоти або вектора в рослинну клітину. Наприклад, включення може відбуватися шляхом трансформації, переважно стабільної. Також, відповідні техніки включення, такі як біолістична трансформація, агробактеріо-опосередкована трансформація або електропорація, є відомими для обізнаного фахівця (Sambrook та співавт.2001).

Також, цей винахід відноситься до трансгенних рослин, або їх частин, що містять трансгенні рослинні клітини, як описано вище. У цьому випадку частиною рослини може бути клітина, тканина, орган або комбінація певної кількості клітин, тканин або органів. Комбінація певної кількості органів - це, наприклад, квітка або насінина. Зокрема, у цьому варіанті реалізації винаходу описано отримання насіння з трансгенної рослини, яке містить молекулу нуклеїнової кислоти у якості трансгену, згідно цього винаходу. Трансгенні рослини, як описано у цьому винаході, а саме рослини роду буряків, переважно мають вищу стійкість до патогену, а саме до ВНПЖБ, ніж відповідні нетрансформовані рослини (рослини без трансгену). Рівень стійкості, наприклад, до ВНПЖБ у рослин роду буряків можна оцінити якісно за допомогою визначення показників відповідних шкал (шкали для рослин роду буряків можна знайти в попередніх роботах, наприклад, у Mechelke (1997) для цукрового буряка). Вища стійкість проявляється у кращій резистентності за щонайменше однією шкалою, або принаймні двома шкалами, або принаймні трьома чи більше шкалами. Цей винахід також належить до методу отримання трансгенних рослинних клітин, включаючи стадію введення молекули нуклеїнової кислоти або вектора в рослинну клітину, відповідно до цього винаходу, а також в деяких випадках - до стадії селекції трансгенної рослинної клітини. Крім того, такий метод для отримання трансгенної рослини передбачає подальший крок, який включає регенерацію трансгенної рослини із трансгенної клітини, отриманої у попередній стадії. Методи регенерації відомі фахівцям, обізнаному у цій галузі, із попередніх робіт.

Також цей винахід пропонує метод для отримання або збільшення стійкості до патогену, а саме до ВNYVV, у рослин, а саме у рослин роду буряків; цей метод включає стадію трансформування рослинної клітини за допомогою молекули нуклеїнової кислоти або вектора, як описано у цьому винаході. Використання цього методу переважно приводить до покращення резистентності принаймні за однією шкалою, а в кращому випадку - до підвищення резистентності за двома, трьома або більше шкалами. Відповідні показники шкал для рослин

роду буряків відомі з попередніх робіт, наприклад, з роботи Mechelke (1997) для цукрового буряка.

Також цей винахід стосується регуляторної послідовності промотора, який контролює експресію гену, що містить молекулу нуклеїнової кислоти, описану у цьому винаході, ця регуляторна послідовність може зумовлювати або модулювати експресію гетерологічної послідовності ДНК у відповідь на інфікування патогеном, також ця регуляторна послідовність містить молекулу нуклеїнової кислоти, яка має послідовність нуклеотидів відповідно до SEQ ID NO: 1 у нуклеотидах 1-1403. Гетерологічна послідовність ДНК - це, переважно, послідовність нуклеотидів, яка кодує компонент захисту рослини від патогену (наприклад, ген резистентності (R-gene) або ген, що кодує ензими, задіяні у передачі сигналу, наприклад, кінази або фосфатази, а також G- протеїн), або та, що кодує ефектор патогену (відомий як ген авірулентності (avr)). Крім того, цей винахід включає в себе рекомбінантну молекулу ДНК, яка містить регуляторну послідовність, описану вище. Рекомбінантна молекула ДНК переважно функціонально зв'язана з гетерологічною послідовністю ДНК.

Також цей винахід належить до клітини-хазяїна, яка трансформована за допомогою вищеописаної регуляторної послідовності або яка містить описану рекомбінантну молекулу ДНК; а також трансгенної рослини, рослинної тканини або клітини, яка містить регуляторну послідовність або рекомбінантну молекулу ДНК як трансгену. Крім того, цей винахід пропонує спосіб одержання трансгенних рослин, включаючи стадію введення регуляторної послідовності або рекомбінантної молекули ДНК у рослинну клітину, як описано у цьому винаході, а також, в деяких випадках, стадію відбору трансгенних рослинних клітин. Крім того, такий спосіб для отримання трансгенної рослини передбачає подальший крок, який включає регенерацію трансгенної рослини із трансгенної клітини, отриманої на попередній стадії.

Як зазначено вище, ген резистентності RZ-3 був визначений шляхом послідовного процесу клонування на основі картування. Цей процес, зокрема, включав у себе наступні кроки: точне генетичне картування, фізичне картування, формування дуже великої популяції для сплайсингу, що містила більш ніж 8000 F2 сплайсингових нащадків, рекомбінантний скринінг, розробку маркера у таргетному регіоні, порівняльний аналіз послідовностей за допомогою штучних бактеріальних хромосом (BAC-секвенування) у резистентних та чутливих генотипах, біоінформаційний аналіз, попередню оцінку білків та порівняння білків. Така робота є надзвичайно трудомісткою та витратною, крім того, невідомо, чи буде її результатом вдале виявлення гену. Після інтеграції локуса RZ-3 із *Beta vulgaris* підвиду *maritima* до рослини роду буряків, а саме до цукрового буряка (*Beta vulgaris* subsp. *vulgaris* var. *altissima*), маркери, що давали хороші значення, використовувалися для відстеження сегмента генома RZ-3 у точному картуванні, що виявилось особливо важким, оскільки таргетна область має повтори на великих проміжках. Однак, неочікувано, виявилось можливим вдало розробити кілька ідентифікаційних маркерів, які до того ж функціонували лише при використанні певної техніки маркування, а саме піросеквенування, тобто були PSQ-маркерами, або нуль-алельними маркерами.

Незважаючи на описані технічні труднощі, шляхом комплексного аналізу за допомогою цих маркерів ми змогли відмежувати локус RZ-3 у вигляді геномного регіону довжиною 0,67 сМ. Це відповідає фізичній довжині приблизно 340,000 пар азотистих основ. Незважаючи на інтенсивні зусилля, в подальшому було можливо лише частково зменшити інтрогресію у гені *Beta vulgaris* subsp. *maritima* шляхом маркування та визначити можливі гени, які могли би бути геном RZ-3. Подальше скорочення інтрогресії, однак, є бажаним у будь-якому випадку, з точки зору культивування, щоб уникнути будь-якого потенційного "перетягування зв'язку", що дуже характерно для гену RZ-3. В результаті, ми змогли обмежити таргетний регіон приблизно до 0,07 сМ шляхом кількостадійного точного картування та за допомогою інформації про послідовності із фізичних генетичних карт. Однак, це стало можливим лише завдяки тому, що було перевірено загалом 8004 екземплярів, включаючи інформативні рекомбінантні рослини BC2S1 або BC2S2, для яких були інтенсивно проаналізовані 90-180 нащадків, у кожному випадку. Це було необхідно, оскільки з невідомих причин ми не завжди могли зрозуміти експресію резистентності. Описані нащадки були індивідуально генотиповані та паралельно фенотиповані. За допомогою статистичних методів (Т-тест, аналіз сили) були визначені фенотипи інформативних рекомбінантів (гомозиготні резистентні - RR; гетерозиготні резистентні - Rs; гомозиготні чутливі - ss), таким чином, ми змогли зробити висновки щодо генотипу інформативних рекомбінантів.

У відносно малому таргетному регіоні довжиною приблизно 38,000 пар азотистих основ, до генотипу чутливості можна було віднести десять генів. У резистентному BAC-пулі було визначено перекриваючі клони для цього таргетного регіону за допомогою нових специфічних до нього маркерів, а згодом їх було секвеновано. Через те, що таргетний регіон повторюється,

послідовність у чутливому генотипі була визначена у вигляді невеликих частин з невідомим змістом. Тому поєднання послідовностей RR та ss було надзвичайно важливим. Однак, виявилось можливим визначити гіпотетичний ген резистентності. Він містився у всіх без винятку ss-генотипах у вигляді ретротранспозону, що мав довжину приблизно 8000 пар азотистих основ між LRR доменом та доменом IR, що неможливо було визначити у генотипах RR. Послідовність амінокислот, яка передбачалася у гіпотетичному гені резистентності, показала, що цей ген ймовірно кодує білок NB-ARC-LRR. Такий висновок можна зробити, виходячи з того, що введення ретротранспозона вимикає функцію гена у чутливих генотипах ss, оскільки він відділяє внутрішній повторюваний домен (IR) від двох інших доменів (NB-ARC та LRR).

Порівняння гену NBS-LRR у генотипах ss з цим геном у генотипах RR також показало значний поліморфізм, що продемонстровано на фіг. 1, 2 та 3. На основі цього поліморфізму у гені NBS-LRR було розроблено та випробувано маркери у широкому наборі генотипів, що складав приблизно 100 ss та RR генотипів. Маркерні панелі, а також порівняльні визначення послідовностей у таргетному гені підтвердили, що включення практично завжди пов'язано з чутливістю. Однак, кілька генотипів ss, як виявилось, не мали інсерції ретротранспозона і все одно залишалися чутливими. Ці генотипи ss, однак, можна було відрізнити від генотипів RR за допомогою маркерів, що показують значущий поліморфізм, як показано на фіг. 1, 2 та/або 3.

У проаналізованій популяції було визначено рекомбіанти за таргетним регіоном, що демонструє рекомбінацію між геном NBS-LRR та гіпотетичним геном, що знаходиться нижче у суміжній ділянці та який міг би кодувати білок з анкіриновим повтором. У випадку двох рослин, рекомбінації можна виявити між геном NBS-LRR та гіпотетичним геном, що знаходиться вище у суміжній ділянці, що міг би кодувати білок DUF565 (білок з невідомою функцією). За допомогою аналізу резистентності у нащадків цих рекомбінантних рослин (видалення одного гену у ділянці вище та нижче гену NBS-LRR), можна цілком чітко продемонструвати, що ген між геном білка з анкіриновим повтором та геном білка DUF565, а саме ген NBS-LLR, описаний в цьому винаході, відповідає за стійкість у генотипі RR. На фіг. 4 наведено фізичну карту таргетного регіону RZ-3 з розробленими маркерами. Дані генотипу вісьмох цільних рекомбінантних ліній, а також статистичний аналіз їхніх нащадків показано на фіг. 5.

Також винаходом пропонується метод визначення молекули нуклеїнової кислоти, що кодує білок, здатний зумовлювати стійкість до патогену BNYYV у рослинах роду буряків, де експресується цей білок. Метод включає в себе виявлення відсутності інсерцій до послідовності кодуючих нуклеотидів у молекулі нуклеїнової кислоти. Метод переважно включає виявлення відсутності інсерції, а саме інсерції ретротранспозона, до послідовності кодуючих нуклеотидів у молекулі нуклеїнової кислоти. Ретротранспозон може бути довжиною, наприклад, приблизно 500 пар азотистих основ, приблизно 1000, 2000, 4000, 8000 або більше пар азотистих основ. У цьому випадку, при описі методу, зазначена молекула нуклеїнової кислоти - це молекула нуклеїнової кислоти, описана вище у цьому винаході, а також та, що кодує ген резистентності RZ-3 або функціональний гомолог гену RZ-3. Рослини роду буряків - це переважно *Beta vulgaris* subsp. *maritima* або *Beta vulgaris* subsp. *vulgaris* var. *altissima* (цукровий буряк). Обізнаний у цій галузі фахівець володіє знаннями про те, які методи підходять для виявлення відсутності інсерції. Наприклад, такий фахівець, знаючи молекулу нуклеїнової кислоти, описану у цьому винаході, може розробити молекулярні маркери, що виявляють присутність або відсутність інсерції у вищеописаному регіоні гену NBS-LLR (див. приклади для приблизного розуміння). Цей винахід включає вищезгадані маркери та їх використання для виявлення присутності або відсутності інсерції і для селекції резистентних, а саме резистентних до BNYYV, рослин, а саме *Beta vulgaris* subsp. *maritima* або *Beta vulgaris* subsp. *vulgaris* var. *altissima* (цукровий буряк). Такі маркери переважно описують локуси у точках інсерції ретротранспозону. Точки інсерції - це точки переходу між геномною ДНК та ретротранспозоном на 5' та/або 3' інсерції. Точки переходу потрібно визначати в широкому діапазоні, а локуси маркерів можуть бути організовані на ДНК на відстані менше 1000 нуклеотидів, бажано менше 800 або 600 нуклеотидів, найбільш бажано - менше ніж 400, 200, 150, 100, 50, 40, 30, 20 або 10 нуклеотидів вище та нижче від точки інсерції. Альтернативним або додатковим кроком виявлення присутності або відсутності інсерції до кодуючої послідовності нуклеотидів молекули нуклеїнової кислоти є метод, що зокрема включає виявлення поліморфізму у кодуючій послідовності нуклеотидів молекули нуклеїнової кислоти за допомогою молекулярних маркерів, що визначають поліморфізми, а саме значущі поліморфізми; метод включає виявлення принаймні одного поліморфізму, згідно фіг. 1, 2 та/або 3, бажано виявлення принаймні 2 чи 3 поліморфізмів, як показано на фіг. 1, 2 та/або 3, найбільш бажано виявлення чотирьох, п'яти або більше поліморфізмів, як показано на фіг. 1, 2 та/або 3, згідно цього винаходу. Це виявлення переважно виконується за допомогою

принаймні одного маркера для одного поліморфізму, а саме для значущого поліморфізму. Фахівець, обізнаний у цій галузі, має інформацію про те, які техніки маркування слід застосувати для визначення відповідного поліморфізму та як отримати молекулярні маркери для цього (література). Крім того, цей винахід включає молекулярні маркери, що описують або визначають поліморфізм згідно до фіг. 1, 2 та/або 3, а також застосування молекулярних маркерів для визначення поліморфізму як показано на фіг. 1, 2 та/або 3. Також, методи ідентифікації, описані вище, включають у себе методи для селекції рослини, яка має стійкість до BNYVV. Методи селекції включають остаточний етап селекції резистентних рослин.

Крім того, ми змогли показати, що у вивчених генотипах RR знаходилася частина геномної послідовності ДНК, яка відповідала області SEQ ID NO: 4, розташованій вище гена RZ-3 (SEQ ID NO: 1), а також частина геномної послідовності ДНК, що відповідала SEQ ID NO: 5, розташованій нижче гена RZ-3 (SEQ ID NO: 1), та які були щільно пов'язані з геном RZ-3 і, як наслідок, гарно підходили як регіони ДНК для розробки діагностичних маркерів для гену RZ-3. Таким чином, цей винахід належить до методу селекції рослини, яка мала би стійкість до BNYVV. Цей метод включає в себе використання молекулярного маркера на послідовності ДНК, що відповідає SEQ ID NO: 4 та/або на послідовності ДНК що відповідає SEQ ID NO: 5, та заключний етап селекції резистентної рослини. Обізнаний у цій галузі фахівець має інформацію про те, як розробити та використовувати маркери на основі наведеної вище інформації про послідовність.

За допомогою цього винаходу можна отримати наступні переваги при культивуванні та розробці нових резистентних ліній рослин роду буряків:

Інформація про послідовність, а також визначені поліморфізми, що дозволяють відрізнити резистентні RR та чутливі ss алелі в описаному гені, роблять можливою розробку маркера напряму у гені, що дозволяє розробити оптимальні елітні лінії без "перетягування зв'язку" та зумовлює значну користь для агронома. Також, інформація про структуру послідовності може бути використана для ідентифікації наступних генів резистентності, а саме резистентності проти ризоманії, які, як варіант, були би частково гомологічними.

Використання алельного гену стійкості, описаного в цьому винаході, у цис- або трансгенних дослідженнях надає можливість виведення нових резистентних видів рослин роду буряків, які, зважаючи на дозозалежний ефект, матимуть більш високу стійкість або в яких, в результаті стекингу/укладання гену, описаного в цьому винаході, з іншими генами резистентності, можна уникнути переривання та спостерігатиметься оптимізована експресія резистентності. Модифікація гена шляхом культивування або селекційної інженерії є також можливою для розробки нових алелей резистентності.

Цей винахід також відноситься до використання ідентифікованого гена резистентності RZ3 при генетичному або молекулярному стекингу з іншими генетичними елементами, які можуть зумовлювати агрономічно вигідні властивості у рослині. В результаті може бути значно збільшена економічна цінність врожайних рослин, наприклад, за рахунок підвищення плодоносності, також можлива розробка нових посівних площ, що не були доступні раніше для культивування певних рослин, в тому числі, за рахунок біотичних факторів, таких як важке навантаження патогенів, або абіотичних факторів, таких як сухість ґрунту/клімату. Агрономічно вигідною властивістю є, наприклад, стійкість до гербіцидів, таких як гліфосат, глюфосинат або інгібітори ацетолататсинтази. Інші численні гербіциди та можливості їх застосування відомі з попередніх досліджень за цією темою для відповідно фахівця. Обізнаний в цій галузі фахівець може звернутися до попередніх робіт за цією темою, щоб отримати інформацію щодо генетичних елементів, які потрібно застосовувати, та способів їх застосування для отримання бажаної стійкості у рослин. Як приклад можна навести агрономічно вигідну властивість у вигляді додаткової стійкості до патогену, де патогенами можуть бути комахи, віруси, круглі черви, бактерії або гриби. Прикладом також може бути комбінування стійкості/толерантності до різних видів патогенів, де широкий спектр захисту рослини від патогенів може бути досягнутий завдяки тому, що генетичні елементи можуть мати доповнюючі ефекти. Відповідно освіченому фахівцю відомі численні гени стійкості, які можна використати як генетичні елементи для досягнення цієї мети. Також, прикладом агрономічно вигідної властивості може бути стійкість до низької температури та морозу. Рослини, що мають таку властивість, можна сіяти раніше або, наприклад, вони можуть залишатися довше на полі, навіть під час заморозків, що, зокрема, може привести до більшого врожаю. Також обізнаний у цій галузі фахівець може звернутися до попередніх робіт з цієї теми, щоб знайти відповідні генетичні елементи. Також, прикладом агрономічно вигідної властивості може бути ефективне засвоєння води, ефективне засвоєння азоту та врожайність. Генетичні елементи, що можуть зумовлювати такі властивості, можна знайти у попередніх роботах з цієї теми.

Відповідно освіченому фахівцю також відомі численні модифікації захисту від патогену. Крім часто описуваних родин R-генів, також з користю можна використовувати досягнення по дослідженню Avr/R, комплементції Avr гену (WO 2013/127379), автоактивації гену R (WO 2006/128444), дослідження HIGS (пригнічення гену індукване клітиною-хазяїном) (наприклад, WO2013/050024) або VIGS (пригнічення гену індукване вірусом). Зокрема, автоактивація гену R може мати велике значення для цього винаходу. Для досягнення цієї мети потрібно розробити нуклеїнову кислоту, яка кодує білок автоактивованої резистентності, який, в свою чергу, зумовлює стійкість до патогенів у рослин. Таким чином, ця нуклеїнова кислота матиме лише частину гену резистентності NBS-LRR, наприклад, гену RZ3, поширену від нижчележачої ділянки із 5'-кінця кодуєчого регіону гену стійкості NBS-LRR до початку домену NBS в гені стійкості NBS-LRR, за умови, що ген NBS-LRR не є геном стійкості TIR-NBS-LRR.

Крім того, цей винахід включає в себе використання алельного гену стійкості RZ3, визначеного за допомогою вищеописаного методу, для комбінування з однією із вищеописаних модифікацій або з одним з вищеописаних генетичних елементів, які можуть зумовлювати одну або більше агрономічно вигідних властивостей в рослин.

Варіанти та способи реалізації цього винаходу будуть описані на прикладах із посиланням на прикладені фігури та послідовності:

Послідовності:

SEQ ID NO: 1 послідовність геномної ДНК у гені стійкості RZ-3. Послідовність включає нуклеотиди від 1 до 1403 регуляторної області промотора.

SEQ ID NO: 2 передбачувана амінокислотна протеїну резистентності RZ-3\_1

SEQ ID NO: 3 передбачувана амінокислотна протеїну резистентності RZ-3\_2

SEQ ID NO: 4 суміжні хромосомальні області, що розташовані вище гену RZ-3 (SEQ ID NO:

1)

SEQ ID NO: 5 суміжні хромосомальні області, що розташовані нижче гену RZ-3 (SEQ ID NO:

1)

SEQ ID NO: 6 консенсусна послідовність у геномній послідовності гену RZ-3 в генотипі ss

SEQ ID NO: 7 таргетна послідовність у гені RZ3 конструкта PHKi у векторі pZFN-C48-RNAi.

Фігури:

Фіг.1 A-I: порівняння послідовності нуклеотидів у геномній консенсусній послідовності гену RZ-3 у генотипі ss (SEQ ID NO: 6) та гену RZ-3 у генотипі (SEQ ID NO: 1). Значущі поліморфізми показано сірим та чорним шрифтом. Незначущі поліморфізми підкреслено. Потенційні початкові точки транскрипції у гені позначені стрілками. Вони ведуть до двох варіантів поліпептида RZ-3\_1 та RZ-3\_2. Позиція ретротранспозона позначена чорним трикутником згори.

Фіг. 2 A-L: порівняння послідовності амінокислот передбачуваного поліпептида по генотипу RR (RZ-3\_1; SEQ ID NO: 2) та поліпептидів по 22 різним генотипам ss. Значущі поліморфізми показано сірим та чорним шрифтом. Незначущі поліморфізми підкреслено.

Фіг. 3 A-L: порівняння послідовності амінокислот передбачуваного поліпептида по генотипу RR (RZ-3\_2; SEQ ID NO: 3) та поліпептидів по 22 різним генотипам ss. Значущі поліморфізми показано сірим та чорним шрифтом. Незначущі поліморфізми підкреслено.

Фіг.4: фізична карта таргетної області гену RZ-3. П'ять генів були позначені на представленій таргетній області у референтному чутливому генотипі: ("2" (DUF565), "3" (гіпотетичний білок), "4" (NBS-LRR ген-кандидат), "5" (ретротранспозон) та "6" (анкіриновий повтор)). Ген-кандидат NBS-LRR ("4") містить ретротранспозон ("5") у чутливій референтній послідовності. Цей ретротранспозон повністю відсутній у резистентній послідовності, і таким чином, лише чотири гени можуть бути позначені як гени генотипу резистентності "2", "3", "4" та "6") Позиції найбільш щільних рекомбінацій (рекомбіанти: 111T\_3515/ ZR11007\_03075 під номером "7" та 111PB3645/ZR08093\_05621 під номером "8") зображені згори. Виходячи з вищезазначеного, можна відмежувати коротшу таргетну область "1". Маркери, розроблені для цієї мети шляхом рекомбінантного аналізу, зображені у вигляді чорного пунктиру у нижній частині фігури. Сегмент гену ("9"), що обраний та виділений як таргетна послідовність із домену "10", був використаний для оцінки гену у роботі з PHKi для сплайсингу генів у алельному гені резистентності RZ-3.

Фіг. 5: аналіз маркерів рекомбінантів по найбільш близько розміщених локусах у таргетній області RZ-3 (маленькі літери у жирній рамці є маркерними даними, отриманими in silico). Вісім рекомбінантних ліній були фенотиповані та генотиповані у 1051 нащадків. Нащадки були розділені на 3 групи (RR резистентні гомозиготи, Rs гетерозиготи, ss чутливі гомозиготи) на основі даних маркування за геном-кандидатом NBS-LRR або сплайсингу фланкуючих областей у випадках, коли ген-кандидат NBS-LRR був гомозиготним RR або ss. Крім того, представлено відповідні дані ІФА. Статистичну обробку даних сплайсингу, а також випадків, коли сплайсинг не



відбувся, було виконано за допомогою т-тесту та критеріїв Вілкоксона. На основі результатів можна було чітко відокремити ген-кандидат між маркерами s3e5800s01 та s3e5873s01.

Фіг. 6: Вектор трансформації pZFN-C48-RNAi: d35S-промотор; C48 s: положення смислової послідовності C48; AtAAP6 інтрон2: інтрон амінокислоти пермеази 6 Арабідопсису (*Arabidopsis thaliana*); C48: положення антисмислової послідовності C48; Nos-T: термінатор Нопалін синтази; LB фланкуюча ділянка: фланкуюча ділянка лівого краю; ZFN ділянка: ділянка розпізнавання білкового домена "цинкові пальці" (комплементарна); Pnos: промотор Нопалін синтази; NPT: кодує послідовність; ген неоміцин фосфотрансферази (npt); pAG7: термінатор pAG7; Bvpa13'UTR: 3' -нетрансльована область гену Pal буряка цукрового; LB: лівий край; aadA: кодує послідовність; аміноглікозид-3"-аденилтрансферази (AAD); pVS1-REP: початок реплікації pVS1;

ColE1 ori: початок реплікації ColE1; RB: правий край.

Приклади:

Картування та точне картування гену RZ-3/ фізична генетична карта.

Ген резистентності RZ-3 (також позначений як стійкість-C48 або C48) було картовано шляхом кількох кроків за допомогою картування та точного картування хромосоми 3 між 57.1 та 57.8 cM (внутрішня референтна карта), тобто на генетичній відстані між двома фланкуючими маркерами 0.0714 cM на генетичній карті. Для картування всього досліджено 8004 рослин, схрещених за S504 (чутливий генотип) x T74 (стійкий генотип). Паралельно до QTL картування C48, після кожного кроку картування розроблялися нові інформативні маркери, що орієнтовані на таргетну область та застосовані для відмежування таргетної області C48.

Координати було додатково підтверджено при точному картуванні за допомогою аналізу нащадків інформативних рекомбінантів. Для цієї мети були детально проаналізовані інформативні рекомбінантні рослини BC2S1 або BC2S2, у кожному випадку аналізували 90-180 нащадків. Ці нащадки були генотиповані та фенотиповані паралельно та індивідуально для кожної рослини. За допомогою статистичних методів (т-тест, аналіз потужності), було визначено фенотипи інформативних рекомбінантів (резистентні гомозиготи RR/гетерозиготи Rs/чутливі гомозиготи ss), і, таким чином, можна було зробити висновки стосовно генотипу інформативних рекомбінантів. Отримано гомозиготні класи нащадків (RR проти ss), що відрізняються за параметрами резистентності, ген був присутній у гетерозиготній ділянці (Rs) батьківської рослини, в іншому випадку він був присутній у гомозиготній ділянці (RR або ss) батьківської рослини.

Фізична карта була розроблена для чутливих до ризоманії генотипів за допомогою визначення маркерів та їхніх генетичних позицій у хромосомних послідовностях. Після QTL визначення області C48 було розроблено нові інформативні маркери на основі референтної послідовності та було проведено додаткове порівняльне секвенування у резистентних генотипах (секвенування наступного покоління та секвенування за Сангером).

Область, що була визначена за допомогою точного картування, містить послідовність довжиною 37996 пар основ (положення фланкуючих маркерів SNP) у чутливій референтній послідовності. Колінеарність між генетичною та фізичною картами у таргетній області є суттєвою (послідовність 12 маркерів у таргетній області).

Ідентифікація та секвенування резистентних BAC- клонів

Бібліотека BAC була розроблена для вибраного резистентного генотипу RZ-3 (C48). Із цього BAC-банку було взято зразки та промарковано C48 QTL області. Для дослідження описаної вище таргетної області було вибрано декілька BAC-клонів. Із них для секвенування були обрані три BAC- клони різної довжини, де було повністю визначено таргетну область. BAC-клони були секвеновані, та на основі отриманих даних секвенування було виконано збірку "de novo". Серед отриманих сегментів у послідовності, що кодує стійкість, найбільша послідовність мала довжину 110909 пар азотистих основ (34537 зчитуваних фрагментів) та повністю містила таргетну область.

Порівняння чутливих та резистентних послідовностей - оцінка послідовностей

Колінеарність двох ss та RR послідовностей була порівняна за допомогою різноманітних програмних інструментів. Для обох - резистентної та чутливої послідовностей - анотація гену виконувалася за допомогою програм Maker та Pedant. Анотація гену в обох послідовностях показала одну й ту саму послідовність передбачуваних генів. Однак, неочікувано ми виявили значну відмінність в одному із цих генів, а саме у гені, що є цим винаходом (RZ-3). Ретротранспозон можна було анотувати до чутливого генотипу у цьому ідентифікованому гені NBS-LRR. Введення ретротранспозону виконувалося у гені між двома доменами - LRR та IR. Резистентний генотип не має цієї інсерції та відтворюється згідно послідовності SEQ ID NO: 1.

Крім того, передбачувані послідовності поліпептидів були зрештою порівняні та оцінені (частково показано на Фіг. 2 та 3).

#### Порівняльне секвенування гена-кандидата NB-ARC-LRR

Ген-кандидат NB-ARC-LRR був секвенований за допомогою порівняльного методу шляхом двох кроків. У наборі генотипів, що містив всього 92 резистентних та чутливих генотипи була верифікована точка інсерції ретротранспозона. Цей аналіз показав, що жоден з резистентних генотипів не мав інсерції ретротранспозона. Тим часом, в чутливих генотипів інсерції можна було визначити у більш ніж 90 % випадків. Визначення інсерції таким чином виявилось пов'язаним з генотипом чутливості. Однак, завдяки знайденим невідповідностям (приблизно 10 % решти чутливих генотипів без інсерції), на другому етапі секвенування всього гену перед точкою інсерції було розширене та охопило область промотора (SEQ ID NO: 1). Загалом було просеквеновано та проаналізовано 31 вибраний резистентний та чутливий генотип, включаючи суперечливі генотипи. В результаті всі резистентні генотипи, де стійкість мала сім різних джерел походження, були на 100 % ідентичними при порівнянні приблизно 4100 пар основ. Також, у послідовності нуклеотидів, із яких певна кількість призводила до заміни амінокислот у білковій послідовності було виявлено безперечно значущі поліморфізми (див. Фіг. 1, 2 та 3). Деякі із цих заміни, а саме в областях доменів, могли призводити до функціональної втрати визначеного білка резистентності у генотипах ss. Крім того, три інсерційно-делеційні поліморфізми (INDELs), повністю зв'язані з резистентністю (зміщення рівноваги зв'язку = 1), також були виявлені в області промотора (Фіг. 1). Ці інсерційно-делеційні поліморфізми (INDELs) також можна вважати потенційними кандидатами на ті, що зумовлюють функціональну втрату відповідного білка.

#### Верифікація гену за допомогою рекомбінантів за близько розміщеними локусами

В проаналізованій популяції, що містила 8004 рослин, було виявлено 16 рекомбінантів у таргетній області (область точно картована і визначено 37996 пар основ). Із цих 16 генотипів 9 рослин містили рекомбінацію між білком NB-ARC-LRR та суміжним білком з анкіриновим повтором на правій стороні. У випадку двох рослин, рекомбінації були знайдені між білком NB-ARC-LRR та суміжними білками DUF565 зліва (білок з невідомою функцією). За допомогою аналізу нащадків всіх цих рекомбінантних рослин (довжина гену вліво та вправо), ми змогли цілком чітко продемонструвати, що ген лежить між DUF565 та білком із анкіриновим повтором, а особливо те, що лише білок NB-ARC відповідає за стійкість.

#### Приблизне визначення відсутності інсерції транспозона

Для визначення інсерції ретротранспозона, було розроблено 3 спеціальних комбінації праймерів. Перша та друга комбінації праймерів здатні визначити інсерції, оскільки в кожному випадку один праймер із кожної з двох пар знаходиться в ретротранспозоні (ліве або праве фланкування ретротранспозона), а другий праймер приєднаний безпосередньо перед ретротранспозоном або після нього. Третя пара праймерів визначає відсутність ретротранспозона, і таким чином праймери мають точки приєднання перед ретротранспозоном та після нього. Продукт ПЛР зрештою може бути отриманий при стандартних умовах лише якщо немає ретротранспозона, в іншому випадку, коли

коли ретротранспозон присутній, продукт ПЛР буде занадто великим і в такому випадку неможливо буде створити амплікон.

#### Верифікація гену за допомогою РНКі

Крім описаної вище верифікації генів за допомогою рекомбінантів, подальше визначення ефекту резистентності на ген було виконане за допомогою РНК-інтерференції. Для цієї мети, стандартний генотип резистентного цукрового буряка був трансформований за допомогою ДНК-конструкта, який кодує двохспіральну шпилькову РНК. Ця дсРНК може впливати на пост-транскрипційне пригнічення гена, що може зменшити або вимкнути ефект алельного гену резистентності RZ-3, і таким чином попередньо резистентний генотип цукрового буряка повинен стати чутливим до ризоманії.

Для того, щоб отримати придатний ДНК-конструкт, було вибрано певну таргетну область алельного гену резистентності RZ3, що мала довжину 434 пар азотистих основ (SEQ ID NO: 7; Фіг. 4), вона була ампліфікована за допомогою ПЛР та клонована як в смислового та і в антисмислового напрямках у векторі pZFN, що є придатним для синтезу шпилькових структур (Фіг. 6). Цей вектор має подвоєний промотор CaMV 35S, декілька точок клонування, інтрон із гену AtAAP6, що кодує амінокислоту пермеази в Арабідопсиса (*Arabidopsis thaliana*), також кілька точок клонування та термінатор нопалін синтази. Трансформація цукрового буряка за допомогою отриманого вектора була виконана у відповідності з протоколом Ліндсей та Галуа (1990) з використанням антибіотика канаміцина як маркера селекції. Після декількох стадій селекції, було проаналізовано вдалу трансформацію на трансгенних паростках шляхом ПЛР за

допомогою визначення присутності гену *nptII*, інтрону *AAP6* та двох т-ДНК крайніх послідовностей (LB/RB), а також відсутності *vir*. Позитивні паростки були розмножені за допомогою клонування *in vitro* до 30 паростків з кожного, висаджені та перенесені у землю в теплиці. Приблизно 2 тижні потому рослини трансгенного цукрового буряка були висаджені у 5 ґрунт, заражений ризоманією, де вони зростали протягом 8-10 тижнів. В якості контрольної групи було використано нетрансформовані рослини, взяті за основу для трансформації того ж генетичного стандарту резистентності. Для визначення поширення ризоманії та сили атаки BNYVV корені цукрового буряка було зібрано та оцінено шляхом ІФА тесту, в якому низький показник за ІФА був показником резистентності, а високий показник вказував на чутливість 10 (Mechelke 1997, Clark & Adams 1977). Показник ІФА у трансформованого цукрового буряка із середнім значенням 3.55, був достовірно вищим ніж ІФА показник у контрольній групі, що також була резистентною та мала середнє значення 1.27, порівнянне з чутливим стандартом D108\_ss (таблиця 1). Результати тесту ІФА відповідно показали, що попередньо резистентна рослина була чутлива до ВНПЖБ в результаті пригнічення специфічного гену, а саме алельного гену 15 резистентності RZ-3 на фоні трансформації. Таким чином ген, описаний в цьому винаході може бути чітко названий геном резистентності RZ3.

Таблиця 1

Результати тесту ІФА та статистичного аналізу даних  
(D108\_ss = чутливий стандарт; 6921\_RR = резистентна трансформована основа;  
6921\_RNAi = резистентна трансформована основа з дсРНК, направлена проти гену RZ3)

	D108_ss	6921_RR	6921_RNAi
n	6	25	64
Середнє значення	3.98	1.27	3.55
Стандартна похибка	0.02	0.25	0.11
Стандартне відхилення	0.06	1.24	0.87

T-тест (рівень достовірності):  $p < 0.0001$

Джерела інформації:

- 20 Clark, M.F.; Adams, A.N. (1977): Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. J. Gen. Virol. 34, 475-483
- Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM (1982) Nopaline synthase: transcript mapping and DNA sequence. J Mol Appl Genet. 1(6): 561-73.
- 25 Esser K (2000) Kryptogamen 1: Cyanobakterien Algen Pilze Flechten Praktikum und Lehrbuch. Springer Publishing House, Berlin, Heidelberg, 3rd edition. 2000.
- Gidner S, Lennefors BL, Nilsson NO, Bensefelt J, Johansson E, Gyllenspetz U, Kraft T (2005) QTL mapping of BNYVV resistance from the WB41 source in sugar beet. Genome 48: 279-285.
- Larson RL, Wintermantel WM, Hill A, Fortis L, Nunez A (2008) Proteome changes in sugar beet in response to Beet necrotic yellow vein virus. Physiological and Mol. Pl. Pathol. 72: 62-72.
- 30 Lindsey, K., and P. Gallois (1990) "Transformation of sugar beet (*Beta vulgaris*) by *Agrobacterium tumefaciens*." Journal of experimental botany 41.5: 529-536.
- Martin GB, Bogdanove AJ; Sessa G (2003) Understanding the functions of plant disease resistance proteins. Annual Review of Plant Biology 54: 23-61.
- 35 Mechelke W (1997) Probleme in der Rizomaniaresistenzzüchtung, Vorträge für Pflanzenzüchtung, Resistenzzüchtung bei Zuckerrüben, Gesellschaft für Pflanzenzüchtung e.V., 113-123.
- Odell JT, Nagy F, Chua N-H (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313, 810 - 812
- Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, and Somssich IE (1996) Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. EMBO J. 15(20): 5690-5700.
- 40 Sambrook J, Russell DW (2001) Molecular Cloning. A laboratory manual, Cold Spring Harbor Laboratory Press, 3. Aufl. 2001.
- Schmidlin LEDEB, Weyens G, Lefebvre M, Gilmer D (2008) Identification of differentially expressed root genes upon rhizomania disease. Mol. Plant Pathol. 9(6):741-51.
- 45 Scholten OE, Bock TSMD, Klein-Lankhorst RM, Lange W (1999) Inheritance of resistance to Beet necrotic yellow vein virus in *Beta vulgaris* conferred by a second gene for resistance. Theor. Appl. Genet. 99:740-746.

Sohi HH, Maleki M(2004) Evidence for presence of types A and B of beet necrotic yellow vein virus (BNYVV) in Iran. *Virus Genes* 29(3): 353-8.

Van Ooijen G, Mayr G, Kasiem MMA, Albrecht M, Cornelissen BJC, Takken FLW (2008) Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. *Journal of Experimental Botany*, 59(6): 1383-1397

WO/2000/29592 (Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.). Chimeric promoters capable of mediating gene expression in plants upon pathogen infection and uses thereof.

WO/2006/128444 (KWS SAAT AG). AUTOACTIVATED RESISTANCE PROTEIN.

WO/2007/147395 (KWS SAAT AG). Pathogen induzierbarer synthetischer Promotor.

WO/2013/127379 (KWS SAAT AG). PATHOGEN-RESISTANT TRANSGENIC PLANT.

WO/2013/050024 (KWS SAAT AG). TRANSGENIC PLANT OF THE SPECIES BETA VULGARIS HAVING ENHANCED RESISTANCE TO CERCOSPORA.

WO/2013/091612 (KWS SAAT AG). NOVEL PLANT-DERIVED CIS-REGULATORY ELEMENTS FOR THE DEVELOPMENT OF PATHOGEN-RESPONSIVE CHIMERIC PROMOTORS.

## Список последовательностей

<110> KWS SAAT AG  
 <120> Resistenzgen gegen Rizomania  
 <130> KWS 0209 PCT  
 <150> DE102013010026.7  
 <151> 2013-06-17  
 <160> 7  
 <170> PatentIn version 3.5  
 <210> 1  
 <211> 5009  
 <212> DNA  
 <213> Beta vulgaris  
 <400> 1  
 caaatcttct ggcacatg gcggtgttgc cgttcatcaa tttaacatca atggaggtaa 60  
 gagtcatgtt ttttaacaa tataaaactt atacttcctc tggtctgttt taaatgaaac 120  
 gtttgttttc tcacgcaacc caaccactt ttttaataat aaatatTTTT agttgtgtgc 180  
 acgtaaaaaa tataaaaaag ttataatttg atagtatctt gtttgagatt gtgattatta 240  
 agagagtcaa gtctcacaat attcgaaagt ctacgtaatc cacctcaaat tgacgaagaa 300  
 aacaagcagg aaaggattaa gtaagttcgt ggaaccacta gaattgattt tcaaatatag 360  
 ctctacctaa tatatggcct acttttaatt ttaataagg agaaggtaat gtgattagaa 420  
 acaaattggt cttaaattat tcattaagct taataatgta taaacataat caagtgttat 480  
 cttcttttca gggccgtctt gaagattttg gggcccgggt ctattatgaa aattgggccc 540  
 ctaaatttat agaaaataaa gatggaagggt tagagttcta aagatagaaa gttgaaaatc 600  
 taaatataaa tcattgacaa atttattaag ggtgagaaac aagggtgttt tcttcaaata 660  
 tgaagcaaaa ttttcaaaat aatacttcct ccgtttctaa ataagtgcaa catttgcata 720  
 atgtttacta ttcacagttt aaactttaat tagctttggt gattttacatt ttaggaaaaa 780  
 acatagtcac gtgggatctt gttagattcg tctgaatgtg aattttttta atatcaactt 840  
 tttataattt ttacttattg ataattgaag atattaatgg ttaaaataat gcattggcaa 900  
 acgtgaaaac aagaagtgtt gcacttattt agaaacggag gaagtatatt ttgggccttt 960  
 ataattttgg agaccctgcg ctgttgggct ccttgcacac cttcatctac ccctctgctt 1020  
 ctttgataac aatttttcag cgacatgatt gtcgattgat gcatatatta ttgtatactc 1080  
 gatccatatt gtttaagatg aattgtttgt ctttgatggt ctccaatgca tattttgtat 1140  
 acttaggaat tctaattatg tactattagt agacattgag atgaatacat aattgccata 1200  
 atgaagtatg attatttttag ttatatactt tctccattcc aaatatataa atgtaacact 1260  
 tgtgtacttt atgcgtacta atgcataaca acgtgcactc tcatgtgttt aattatatac 1320  
 tttttgagag aagtgttaca ttggggacca tgggactgtg tataatttga ccgcaaaatc 1380



gaagtgtcgc	at ttgattga	aaatggagag	agtagtatat	agatggaaca	cagcagagac	1440
tgctgggcat	ctttggccaa	caaaccctaa	attgatatta	atcccttatt	caggtcattt	1500
catctttttg	acacaaaatg	gatgtttag	gcaactgcgt	atctgctgcc	caatctctgt	1560
ttgcagccct	gcaaagttct	gagctcaaag	agatcctctc	gatctttggc	tacaaatccc	1620
gacttgatga	cctccaacgc	actgtctcta	ccatcaacgc	tgtattccgt	gatgctgaga	1680
ccaaacagga	gctcactcat	gaagcacagc	attggctcga	ggaactcaag	gatgctgtct	1740
ttgaagcaga	tgatctgttc	gacgagtttg	tcactcttgc	cgagcagaag	caacttgtag	1800
aggctgggtg	cagtctttcc	aaaaagatgc	gccaatctct	ttctgattcc	aacccccctg	1860
gcattgctta	taggatgtca	cgaggggtta	agaagatcaa	gaagaagttg	gatgctatcg	1920
cttacaatca	tcaatttagc	tttaagattg	atcttgagcc	tatgaaagag	agaaggctag	1980
agactgggtc	tgctgtgaac	gcaggtgata	tcattggaag	agaggacgac	ttggagaaga	2040
ttgtaggttt	gttgcttgat	tctaacatcc	aacgtgatgt	gtctttcctt	actattgtgg	2100
gaatgggagg	gttgggtaaa	actgctcttg	cccaactcgt	gtacaatgat	ccaagggtca	2160
gaactgcttt	tccattgaga	tggtggaatt	gtgtgtctga	tcaagatcaa	aagcaactag	2220
atgtgaaaga	aattttgggt	aagattcttg	ctacagctac	tggtagaagt	catgaggggt	2280
caaccatgga	tcaggtgcaa	acccaactac	gagaacaact	atgtggcaag	agatacttgc	2340
ttgttttgga	tgatgtatgg	aatgagaatc	ctaataatt	gcgtgatctg	gtagaattct	2400
tcattgggag	tcgaagcaga	aattggattg	tggttaactac	gcgttcgcac	gagacagcga	2460
gaattataag	agatggtcca	ttgcacaagc	tccaagggtt	gtctgaggaa	aactcttggc	2520
gtttatttgt	cagggtggacc	ttcggatcag	tgcaagcaaa	attccctaatt	gactttatca	2580
tgattgcacg	agatatagtt	gacaaaatgt	ctcgaaaccc	tctggctata	agagtggtag	2640
gaagtctttt	gtgtggtcaa	gacaagagta	agtggctttc	atttcatgag	atcgatttag	2700
gcaacattag	aaagagccat	aatgatatac	tgccaatact	gaacctaagt	taccatcatc	2760
ttgaacctcc	aattaagaga	tgcttttagt	attgtgcagt	gtttccaaag	gatttcctta	2820
taggggaagc	gacgctgata	aacctctgga	tggcacaagg	ttatattgtt	ccgttagaca	2880
aagatcaaag	catagatgat	gctagtgagg	aatacatatc	aattttgttg	cggagatggt	2940
ttttcgaaaa	tgctggagca	gaaaaagatg	gtgttattaa	gatccatgat	ctcatgcatg	3000
atattgctca	aaatgtcatg	gggaaggagc	tttgtagcac	taaaaacatt	agtggcagct	3060
tggtataaaag	tgctcgccat	ctatctcttg	ccagaactag	ttttgcaaga	tactctttca	3120
atgcaactca	tattcgctcc	tatttctgtg	ctggctactg	gtgtcaggat	gctgagataa	3180
accagttttc	agttgaggca	ttagtaccaa	actgtttgta	cctaagggca	atggacctcg	3240
cttggctgaa	gataaaaagt	ttaccagact	cgattgggtg	attgttgcat	ttgaggtagt	3300
tagatctttc	gtataacgaa	gatctggaag	tacttccaaa	ctcaattgct	aaactatata	3360
atctacaaac	cttacaattg	aagggttgca	agagattgga	agggttacca	aaacatttga	3420



```
gcaggctggt taagcttcaa actttggata tacatggttg caacaatgta acttatatgc 3480
ccaaagggcat gggtaagttg acttgccttc acactctcag taagtttata gtgggtggag 3540
aaggggagttg ttcaagttgg aagcaatggt ttgatgggtt ggaagatcta aaggctctca 3600
ataacctaaa gggcatctg gaaatccaaa tcaggtggcc caaaaatact acagatgctg 3660
tcaaagaaga tgttacgagg gaaggattat acctgaatca taaggaacat ctcaatcaca 3720
ttgtggttga ttttagatgt gaggaggggtg gtggaagaat ggatgatgag gaagcaagaa 3780
gattgatgga agagttgcgg ccacatcctt atcttgaaaa tttggctgtg aaagcatatt 3840
atggtgtgaa aatgcctggt tgggcaaccc ttctcccaa tcttacagag ctttttcttt 3900
ctgattgtgg ggaactggag aaccttccat gcctgggaaa cttggatcat ctaaaagtcc 3960
tccgactttc gcatttggca aaattggagt acattgaaga agatagctca tcagctaatt 4020
tcaggtgtag gcctggacca gaaagtgcag gactatcatt atacttcccc tcccttgaac 4080
gccttgagtt gaagcgtttg tgtaagttaa aaggatggag gagaggggaa gggtaggag 4140
atgatcacca gccttttaat gaaagcagca gcaatacaca agtccaatta caattatgtc 4200
ttcctcaatt gaagtcattg agaatagaaa gatgcccatt gctgacattt atgccgctgt 4260
gtcccaagac agaaaaactg catttagttg tatttaatga acgactccgg atagtgcag 4320
ctaagagaga tgagaatttc tatgctccat tacattcatc atcatctgat cctgaaaacc 4380
cgaggaacac tattcccat cccatgttta gagaggtata cataaacaat gtggcgtggc 4440
taaattcgct gcctatggag gcttttaggt gtctcactca tatgacaata aaaaacgacg 4500
aggtagagag tttgggagaa gttggagagg tgtttcggag ctgctcatct tctttgcat 4560
ccttgaatat cacaggttgc tccaacttaa gaagtgttgc tggagggctg gagcatctca 4620
ctgctttgga gatgttagaa atatacgaca ccataagct gagtctatca gaagaccag 4680
aagggtgtgt gccatggaaa tcccttcac actccctcag ctacttgcaa ctgatgaatc 4740
tcccacagct ggtcaacctg cctgattcga tgcagttctt ggctgccctc cgaactcttt 4800
caatagtgca ttgactaaa ctgcaatcag tgccagattg gatgcccaga ctacttctc 4860
tcaggaagct tatggtttca ttctgttccg cacatctgga gagaagatgc caaaatccaa 4920
ctgggggtgga ctggcctaac attcaacaca tcccctccat tgatgtcacc tctagccttc 4980
ctaagttttt agtgttgccg tatgaatag 5009
```

```
<210> 2
<211> 1163
<212> PRT
<213> Beta vulgaris
```

```
<400> 2
```

```
Met Asp Val Val Gly Thr Ala Leu Ser Ala Ala Gln Ser Leu Phe Ala
1          5          10          15
```

Ala Leu Gln Ser Ser Glu Leu Lys Glu Ile Leu Ser Ile Phe Gly Tyr  
20 25 30

Lys Ser Arg Leu Asp Asp Leu Gln Arg Thr Val Ser Thr Ile Asn Ala  
35 40 45

Val Phe Arg Asp Ala Glu Thr Lys Gln Glu Leu Thr His Glu Ala Gln  
50 55 60

His Trp Leu Glu Glu Leu Lys Asp Ala Val Phe Glu Ala Asp Asp Leu  
65 70 75 80

Phe Asp Glu Phe Val Thr Leu Ala Glu Gln Lys Gln Leu Val Glu Ala  
85 90 95

Gly Gly Ser Leu Ser Lys Lys Met Arg Gln Phe Phe Ser Asp Ser Asn  
100 105 110

Pro Leu Gly Ile Ala Tyr Arg Met Ser Arg Gly Val Lys Lys Ile Lys  
115 120 125

Lys Lys Leu Asp Ala Ile Ala Tyr Asn His Gln Phe Ser Phe Lys Ile  
130 135 140

Asp Leu Glu Pro Met Lys Glu Arg Arg Leu Glu Thr Gly Ser Val Val  
145 150 155 160

Asn Ala Gly Asp Ile Ile Gly Arg Glu Asp Asp Leu Glu Lys Ile Val  
165 170 175

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

Ile Val Gly Met Gly Gly Leu Gly Lys Thr Ala Leu Ala Gln Leu Val  
195 200 205

Tyr Asn Asp Pro Arg Val Arg Thr Ala Phe Pro Leu Arg Cys Trp Asn  
210 215 220

Cys Val Ser Asp Gln Asp Gln Lys Gln Leu Asp Val Lys Glu Ile Leu  
225 230 235 240

Gly Lys Ile Leu Ala Thr Ala Thr Gly Lys Asn His Glu Gly Ser Thr  
245 250 255

Met Asp Gln Val Gln Thr Gln Leu Arg Glu Gln Leu Cys Gly Lys Arg  
260 265 270

Tyr Leu Leu Val Leu Asp Asp Val Trp Asn Glu Asn Pro Asn Gln Leu  
275 280 285



Arg Asp Leu Val Glu Phe Phe Met Gly Gly Arg Ser Arg Asn Trp Ile  
 290 295 300  
 Val Val Thr Thr Arg Ser His Glu Thr Ala Arg Ile Ile Arg Asp Gly  
 305 310 315 320  
 Pro Leu His Lys Leu Gln Gly Leu Ser Glu Glu Asn Ser Trp Arg Leu  
 325 330 335  
 Phe Val Arg Trp Thr Phe Gly Ser Val Gln Ala Lys Phe Pro Asn Asp  
 340 345 350  
 Phe Ile Met Ile Ala Arg Asp Ile Val Asp Lys Cys Ala Arg Asn Pro  
 355 360 365  
 Leu Ala Ile Arg Val Val Gly Ser Leu Leu Cys Gly Gln Asp Lys Ser  
 370 375 380  
 Lys Trp Leu Ser Phe His Glu Ile Asp Leu Gly Asn Ile Arg Lys Ser  
 385 390 395 400  
 His Asn Asp Ile Met Pro Ile Leu Asn Leu Ser Tyr His His Leu Glu  
 405 410 415  
 Pro Pro Ile Lys Arg Cys Phe Ser Tyr Cys Ala Val Phe Pro Lys Asp  
 420 425 430  
 Phe Leu Ile Gly Lys Gln Thr Leu Ile Asn Leu Trp Met Ala Gln Gly  
 435 440 445  
 Tyr Ile Val Pro Leu Asp Lys Asp Gln Ser Ile Asp Asp Ala Ser Glu  
 450 455 460  
 Glu Tyr Ile Ser Ile Leu Leu Arg Arg Cys Phe Phe Glu Asn Val Gly  
 465 470 475 480  
 Ala Glu Lys Asp Gly Val Ile Lys Ile His Asp Leu Met His Asp Ile  
 485 490 495  
 Ala Gln Asn Val Met Gly Lys Glu Leu Cys Thr Thr Lys Asn Ile Ser  
 500 505 510  
 Gly Ser Leu Asp Lys Ser Val Arg His Leu Ser Leu Ala Arg Thr Ser  
 515 520 525  
 Phe Ala Arg Tyr Ser Phe Asn Ala Thr His Ile Arg Ser Tyr Phe Cys  
 530 535 540  
 Ala Gly Tyr Trp Cys Gln Asp Ala Glu Ile Asn Gln Phe Ser Val Glu  
 545 550 555 560

Ala Leu Val Pro Asn Cys Leu Tyr Leu Arg Ala Met Asp Leu Ala Trp  
565 570 575

Ser Lys Ile Lys Ser Leu Pro Asp Ser Ile Gly Gly Leu Leu His Leu  
580 585 590

Arg Tyr Leu Asp Leu Ser Tyr Asn Glu Asp Leu Glu Val Leu Pro Asn  
595 600 605

Ser Ile Ala Lys Leu Tyr Asn Leu Gln Thr Leu Gln Leu Lys Gly Cys  
610 615 620

Lys Arg Leu Glu Gly Leu Pro Lys His Leu Ser Arg Leu Val Lys Leu  
625 630 635 640

Gln Thr Leu Asp Ile His Gly Cys Asn Asn Val Thr Tyr Met Pro Lys  
645 650 655

Gly Met Gly Lys Leu Thr Cys Leu His Thr Leu Ser Lys Phe Ile Val  
660 665 670

Gly Gly Glu Gly Ser Cys Ser Ser Trp Lys Gln Cys Phe Asp Gly Leu  
675 680 685

Glu Asp Leu Lys Ala Leu Asn Asn Leu Lys Gly His Leu Glu Ile Gln  
690 695 700

Ile Arg Trp Pro Lys Asn Thr Thr Asp Ala Val Lys Glu Asp Val Thr  
705 710 715 720

Arg Glu Gly Leu Tyr Leu Asn His Lys Glu His Leu Asn His Ile Val  
725 730 735

Val Asp Phe Arg Cys Glu Glu Gly Gly Arg Met Asp Asp Glu Glu  
740 745 750

Ala Arg Arg Leu Met Glu Glu Leu Arg Pro His Pro Tyr Leu Glu Asn  
755 760 765

Leu Ala Val Lys Ala Tyr Tyr Gly Val Lys Met Pro Gly Trp Ala Thr  
770 775 780

Leu Leu Pro Asn Leu Thr Glu Leu Phe Leu Ser Asp Cys Gly Glu Leu  
785 790 795 800

Glu Asn Leu Pro Cys Leu Gly Asn Leu Asp His Leu Lys Val Leu Arg  
805 810 815

Leu Ser His Leu Ala Lys Leu Glu Tyr Ile Glu Glu Asp Ser Ser Ser

820										825					830				
Ala	Asn	Phe	Arg	Cys	Arg	Pro	Gly	Pro	Glu	Ser	Ala	Gly	Leu	Ser	Leu				
		835					840					845							
Tyr	Phe	Pro	Ser	Leu	Glu	Arg	Leu	Glu	Leu	Lys	Arg	Leu	Cys	Lys	Leu				
	850					855					860								
Lys	Gly	Trp	Arg	Arg	Gly	Glu	Gly	Leu	Gly	Asp	Asp	His	Gln	Pro	Phe				
865					870					875					880				
Asn	Glu	Ser	Ser	Ser	Asn	Thr	Gln	Val	Gln	Leu	Gln	Leu	Cys	Leu	Pro				
				885					890					895					
Gln	Leu	Lys	Ser	Leu	Arg	Ile	Glu	Arg	Cys	Pro	Leu	Leu	Thr	Phe	Met				
			900					905					910						
Pro	Leu	Cys	Pro	Lys	Thr	Glu	Lys	Leu	His	Leu	Val	Val	Phe	Asn	Glu				
		915					920					925							
Arg	Leu	Arg	Ile	Val	His	Ala	Lys	Arg	Asp	Glu	Asn	Phe	Tyr	Ala	Pro				
	930					935					940								
Leu	His	Ser	Ser	Ser	Ser	Asp	Pro	Glu	Asn	Pro	Arg	Asn	Thr	Ile	Pro				
945					950					955					960				
Ile	Pro	Met	Phe	Arg	Glu	Val	Tyr	Ile	Asn	Asn	Val	Ala	Trp	Leu	Asn				
				965					970					975					
Ser	Leu	Pro	Met	Glu	Ala	Phe	Arg	Cys	Leu	Thr	His	Met	Thr	Ile	Lys				
			980					985					990						
Asn	Asp	Glu	Val	Glu	Ser	Leu	Gly	Glu	Val	Gly	Glu	Val	Phe	Arg	Ser				
		995					1000					1005							
Cys	Ser	Ser	Ser	Leu	Arg	Ser	Leu	Asn	Ile	Thr	Gly	Cys	Ser	Asn					
	1010					1015					1020								
Leu	Arg	Ser	Val	Ser	Gly	Gly	Leu	Glu	His	Leu	Thr	Ala	Leu	Glu					
	1025					1030					1035								
Met	Leu	Glu	Ile	Tyr	Asp	Thr	His	Lys	Leu	Ser	Leu	Ser	Glu	Asp					
	1040					1045					1050								
Pro	Glu	Gly	Val	Val	Pro	Trp	Lys	Ser	Leu	His	His	Ser	Leu	Ser					
	1055					1060					1065								
Tyr	Leu	Gln	Leu	Met	Asn	Leu	Pro	Gln	Leu	Val	Asn	Leu	Pro	Asp					
	1070					1075					1080								



Ser Met Gln Phe Leu Ala Ala Leu Arg Thr Leu Ser Ile Val His  
1085 1090 1095

Cys Thr Lys Leu Gln Ser Val Pro Asp Trp Met Pro Arg Leu Thr  
1100 1105 1110

Ser Leu Arg Lys Leu Met Val Ser Phe Cys Ser Ala His Leu Glu  
1115 1120 1125

Arg Arg Cys Gln Asn Pro Thr Gly Val Asp Trp Pro Asn Ile Gln  
1130 1135 1140

His Ile Pro Ser Ile Asp Val Thr Ser Ser Leu Pro Lys Phe Leu  
1145 1150 1155

Val Leu Pro Tyr Glu  
1160

<210> 3  
<211> 1177  
<212> PRT  
<213> Beta vulgaris

<400> 3

Met Glu Arg Val Val Tyr Arg Trp Asn Thr Ala Glu Thr Ala Gly His  
1 5 10 15

Leu Trp Pro Thr Asn Pro Lys Leu Ile Leu Ile Pro Tyr Ser Ala Leu  
20 25 30

Gln Ser Ser Glu Leu Lys Glu Ile Leu Ser Ile Phe Gly Tyr Lys Ser  
35 40 45

Arg Leu Asp Asp Leu Gln Arg Thr Val Ser Thr Ile Asn Ala Val Phe  
50 55 60

Arg Asp Ala Glu Thr Lys Gln Glu Leu Thr His Glu Ala Gln His Trp  
65 70 75 80

Leu Glu Glu Leu Lys Asp Ala Val Phe Glu Ala Asp Asp Leu Phe Asp  
85 90 95

Glu Phe Val Thr Leu Ala Glu Gln Lys Gln Leu Val Glu Ala Gly Gly  
100 105 110

Ser Leu Ser Lys Lys Met Arg Gln Phe Phe Ser Asp Ser Asn Pro Leu  
115 120 125

Gly Ile Ala Tyr Arg Met Ser Arg Gly Val Lys Lys Ile Lys Lys Lys  
130 135 140

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

Asp Ile Met Pro Ile Leu Asn Leu Ser Tyr His His Leu Glu Pro Pro  
 420 425 430  
 Ile Lys Arg Cys Phe Ser Tyr Cys Ala Val Phe Pro Lys Asp Phe Leu  
 435 440 445  
 Ile Gly Lys Gln Thr Leu Ile Asn Leu Trp Met Ala Gln Gly Tyr Ile  
 450 455 460  
 Val Pro Leu Asp Lys Asp Gln Ser Ile Asp Asp Ala Ser Glu Glu Tyr  
 465 470 475 480  
 Ile Ser Ile Leu Leu Arg Arg Cys Phe Phe Glu Asn Val Gly Ala Glu  
 485 490 495  
 Lys Asp Gly Val Ile Lys Ile His Asp Leu Met His Asp Ile Ala Gln  
 500 505 510  
 Asn Val Met Gly Lys Glu Leu Cys Thr Thr Lys Asn Ile Ser Gly Ser  
 515 520 525  
 Leu Asp Lys Ser Val Arg His Leu Ser Leu Ala Arg Thr Ser Phe Ala  
 530 535 540  
 Arg Tyr Ser Phe Asn Ala Thr His Ile Arg Ser Tyr Phe Cys Ala Gly  
 545 550 555 560  
 Tyr Trp Cys Gln Asp Ala Glu Ile Asn Gln Phe Ser Val Glu Ala Leu  
 565 570 575  
 Val Pro Asn Cys Leu Tyr Leu Arg Ala Met Asp Leu Ala Trp Ser Lys  
 580 585 590  
 Ile Lys Ser Leu Pro Asp Ser Ile Gly Gly Leu Leu His Leu Arg Tyr  
 595 600 605  
 Leu Asp Leu Ser Tyr Asn Glu Asp Leu Glu Val Leu Pro Asn Ser Ile  
 610 615 620  
 Ala Lys Leu Tyr Asn Leu Gln Thr Leu Gln Leu Lys Gly Cys Lys Arg  
 625 630 635 640  
 Leu Glu Gly Leu Pro Lys His Leu Ser Arg Leu Val Lys Leu Gln Thr  
 645 650 655  
 Leu Asp Ile His Gly Cys Asn Asn Val Thr Tyr Met Pro Lys Gly Met  
 660 665 670  
 Gly Lys Leu Thr Cys Leu His Thr Leu Ser Lys Phe Ile Val Gly Gly  
 675 680 685



Glu Gly Ser Cys Ser Ser Trp Lys Gln Cys Phe Asp Gly Leu Glu Asp  
 690 695 700  
 Leu Lys Ala Leu Asn Asn Leu Lys Gly His Leu Glu Ile Gln Ile Arg  
 705 710 715 720  
 Trp Pro Lys Asn Thr Thr Asp Ala Val Lys Glu Asp Val Thr Arg Glu  
 725 730 735  
 Gly Leu Tyr Leu Asn His Lys Glu His Leu Asn His Ile Val Val Asp  
 740 745 750  
 Phe Arg Cys Glu Glu Gly Gly Gly Arg Met Asp Asp Glu Glu Ala Arg  
 755 760 765  
 Arg Leu Met Glu Glu Leu Arg Pro His Pro Tyr Leu Glu Asn Leu Ala  
 770 775 780  
 Val Lys Ala Tyr Tyr Gly Val Lys Met Pro Gly Trp Ala Thr Leu Leu  
 785 790 795 800  
 Pro Asn Leu Thr Glu Leu Phe Leu Ser Asp Cys Gly Glu Leu Glu Asn  
 805 810 815  
 Leu Pro Cys Leu Gly Asn Leu Asp His Leu Lys Val Leu Arg Leu Ser  
 820 825 830  
 His Leu Ala Lys Leu Glu Tyr Ile Glu Glu Asp Ser Ser Ser Ala Asn  
 835 840 845  
 Phe Arg Cys Arg Pro Gly Pro Glu Ser Ala Gly Leu Ser Leu Tyr Phe  
 850 855 860  
 Pro Ser Leu Glu Arg Leu Glu Leu Lys Arg Leu Cys Lys Leu Lys Gly  
 865 870 875 880  
 Trp Arg Arg Gly Glu Gly Leu Gly Asp Asp His Gln Pro Phe Asn Glu  
 885 890 895  
 Ser Ser Ser Asn Thr Gln Val Gln Leu Gln Leu Cys Leu Pro Gln Leu  
 900 905 910  
 Lys Ser Leu Arg Ile Glu Arg Cys Pro Leu Leu Thr Phe Met Pro Leu  
 915 920 925  
 Cys Pro Lys Thr Glu Lys Leu His Leu Val Val Phe Asn Glu Arg Leu  
 930 935 940  
 Arg Ile Val His Ala Lys Arg Asp Glu Asn Phe Tyr Ala Pro Leu His

945		950		955		960
Ser Ser Ser Ser	Asp 965	Pro Glu Asn Pro	Arg 970	Asn Thr Ile Pro	Ile 975	Pro
Met Phe Arg	Glu 980	Val Tyr Ile Asn	Asn 985	Val Ala Trp Leu	Asn 990	Ser Leu
Pro Met	Glu 995	Ala Phe Arg Cys	Leu 1000	Thr His Met Thr	Ile 1005	Lys Asn Asp
Glu Val	Glu 1010	Ser Leu Gly	Glu 1015	Val Gly Glu Val	Phe 1020	Arg Ser Cys
Ser Ser	Ser 1025	Leu Arg Ser	Leu 1030	Asn Ile Thr Gly	Cys 1035	Ser Asn Leu
Arg Ser	Val 1040	Ser Gly Gly	Leu 1045	Glu His Leu Thr	Ala 1050	Leu Glu Met
Leu Glu	Ile 1055	Tyr Asp Thr	His 1060	Lys Leu Ser Leu	Ser 1065	Glu Asp Pro
Glu Gly	Val 1070	Val Pro Trp	Lys 1075	Ser Leu His His	Ser 1080	Leu Ser Tyr
Leu Gln	Leu 1085	Met Asn Leu	Pro 1090	Gln Leu Val Asn	Leu 1095	Pro Asp Ser
Met Gln	Phe 1100	Leu Ala Ala	Leu 1105	Arg Thr Leu Ser	Ile 1110	Val His Cys
Thr Lys	Leu 1115	Gln Ser Val	Pro 1120	Asp Trp Met Pro	Arg 1125	Leu Thr Ser
Leu Arg	Lys 1130	Leu Met Val	Ser 1135	Phe Cys Ser Ala	His 1140	Leu Glu Arg
Arg Cys	Gln 1145	Asn Pro Thr	Gly 1150	Val Asp Trp Pro	Asn 1155	Ile Gln His
Ile Pro	Ser 1160	Ile Asp Val	Thr 1165	Ser Ser Leu Pro	Lys 1170	Phe Leu Val
Leu Pro	Tyr 1175	Glu				

<210> 4  
 <211> 987  
 <212> DNA  
 <213> Beta vulgaris



<400> 4  
 ttctctctta gatttccata gatttgaaca aattggggtg attttcatag attattcata 60  
 attctctctc cataattctt ctctctcttt ctccatacat ttttcattcg caatactcag 120  
 gaactgagta ttgggaattg tccccagttg tcagggtgat gtgcagtaac ttttaagagaa 180  
 gactttcctc tcatgcaaca tgtccctgat ctgttgcttg atggtcgtca tctcactatt 240  
 ctctaataag ctcgatttgt atgaaacaga tgatactata ttccgtttcg tgcaatgtgc 300  
 acgaaacgga atatagtatt ttgtgcaagg tgcacgaaac ggattcgatt gtttcgtgca 360  
 cattgcacga aacggaatca actgtttcat gtagtctaca cgaaacggaa tcaattgttt 420  
 cgtgtagtct acacgaaaca gactaatcat gcattacgaa tcataattac gaaaaaaaaat 480  
 taacaacttg aatcacaatg acgaaaaaaaa attcagaaat tatatcagat tgaaattcga 540  
 ttgggtcaaa attatggtcc attaaatatc aaattaaaaat ttgtagatct tcaatgaagt 600  
 tttttatatc taaccgttag agaggaggag agaataatttt tagagagaga aagggttttt 660  
 tagaaagaat gtgataataa ggggtttttt gggttttttt aggctgcgtt agtaaagtga 720  
 ggctgcattt agcaaccttt tttttttggt aaatttcatt tcctcgatga acaaggaaac 780  
 gaaacggcga gatggcggcg ttggtggaat ttcccggcga aacgcagctt cttttcgatt 840  
 catagttgcc ataaatttgc attttaccca gatttcaaatt aatttttact aattcgtcga 900  
 aattgctcat gaaattgttt atttccgcaa attttttgat taaccctcc agaatttgat 960  
 tcgcaaatat ggcgaagcta ttgagta 987

<210> 5  
 <211> 12364  
 <212> DNA  
 <213> Beta vulgaris

<400> 5  
 ttctgttccg cacatctgga gagaagatgc caaaatccaa ctggggtgga ctggcctaac 60  
 attcaacaca tccccctcat tgatgtcacc tctagccttc ctaagttttt agtgttgccg 120  
 tatgaatagg tatatacttc tttgggtttt gttcgtgctt ccatttagct caaattggaa 180  
 atgagcgtat ggcgtcagat ggtgaccaat ctgcagttat tgcgctacgt gtatgttctg 240  
 gtttatattg atggcaatgt tcaatagttc attataatcc caatcaaatt tctttgtcca 300  
 ttgtttataa tcccaaattc aatttctttg gaagattgtg ctgaggagag cttgatgaag 360  
 gaacttgttt aagggttttt ctcttataga ttatctttct accattgttt ataattccaa 420  
 ctgcatcttg gtctgagaag gaattcaata ttttttctag ttttacttga ggttaaggct 480  
 gtttataatc acaggctttg ccaatagtat aatttttata aagtactact gtacagatta 540  
 tgtgaatctt caaagggttg agagaatcgt cctaaattca tgtaaaacttg gagttaagga 600  
 gcaggaaaatg gagttacttc aagtgttaat gcaatcagct caaaaaatta ctaatacaga 660  
 acttattcgt gtcacaactc agaagccttt ataaattata aaagtagtaa agatttcggt 720



tcgaaagtat tattcatatt agagtacaat gaataatttg ctttggcaaa gccatctgaa	780
ggtccttaag aaatgttttag tgcagggtgat ttaacttgct gtgattatgt caccaggaga	840
acttgcttat caaagaattc agtagcaagt tggtcgatac tcgacagaac atccttatca	900
aacatgttag tgcagaggtc agagcatgat accctttcat tagtattata ttttacatta	960
ttaagaaata tatttacatc tgaaattata tatgctgata caataatctt atgcattttg	1020
agttatacat taaacacggg gaaaaacatt atttcgagca tctgcatttg tgcttggtgt	1080
tgtgtaatgg taaattggca attgcttgac ctttgtttaa ggctgggtaa gctttcaaaa	1140
agttgaattg tttagcagta gatcagtaac tatccaaaca aagaccaagc tatacgccaa	1200
ctgtttcttc acgttcactg tacaagtgtt aaagtatctt caatgggaat accgaaatct	1260
aaagtcaatt atccaattaa tagtagtcaa aagtcaaaca tctctgtatg aatatgtaga	1320
ttgaagagtt tctgtatgca ttcaattcaa tgcaacaagt tgtattcgtt cacaccttat	1380
tacttgggtca aaagttgact aattttacac aaggaacttt agcatcaatc atcatcctac	1440
tatctacgag ttgaaagaaa cttgtacaaa aacttgtttt aaccctgata cagttcagtt	1500
aagcctgttt gttgatcagt ctgcaatttg aatcactcgt ccaactcgcac gacttagtgt	1560
gcgatcttgt gtctagtttt ctcgagaccc cgcctcctga ccgtgcccag cccaccccaa	1620
ccattcccag ctctactagc tagtatcagc atacatggga agagccagga ctgtaagcat	1680
gtcgtttacc gagtacagc tgctgtcagc agcaggagat ggagatgagg aattccta	1740
acaagctctt gctactcaac ctattgatta cttcctaacg cggagcaaaa acaaaaatgg	1800
tgaagaacat tgcaatatta tccacattgc agtggttaa gaacaagcaa agttcctcaa	1860
tcgagcattg agtatattac ccatctcaac tctgcatctc cttctctgtc agcaagattt	1920
ttcctacttt agctacaacc ctcttcactg tgcatcttta cgaggtaact ttgctattgt	1980
caagctcctc gtcgagtttt acgaatcatc atcatcttct tcttcgtcat tggttgatcc	2040
aagctgtaag ccatgggttag ccaaggatgt gaacgggaag acgcctttac aggtggcttt	2100
ggataggggt agaggtgaat gtgcattaaa aataatggga ttagatgaag aattgctttg	2160
taatatgggt gataataaag gtaacagccc gctatttcaa gctgtacaaa gaggtagtga	2220
acaaattgct atgaagatct tggcatcagg gcattcttat agtactgggt gcgagtatga	2280
gttgactccc cttcatgttc taccaaattg ctgagggtgt tagtattgat ttgtttttca	2340
atttgttaaa atttcttacg ctttctgtcc cttaattttt ctcacatgtg ggtttgacac	2400
agacatttag tgtaagagc aacttcaatg gtcagctatg cactcttcta acttagcttt	2460
ccacctcaac tacattcaag taacattagt ttcaggctac aacgtccttt tgggtgcatt	2520
tcttcacttg aacttatatc aaaaccgagc ttattactct tgattggacc taatcaacca	2580
tgatacgtgt tttcggaag ttaacagttt cctaatttag ttttcttctg tacaatttca	2640
agataactaa taaagtatta gcaatcttaa ctataaaaaa agaagaaaac actagcctaa	2700
tatcatctga tctgcagagg aagtttgca acttctactt gacaagcatc cagaaatgat	2760



aaaagcagtc gacaaaaatg gacttacaat cctacacaaa tgggcaatga tgggtaaact	2820
atggccattt caatttcttt taaagcaaga aaaaagttct aggttgagga aggacttcat	2880
caacctttta tgtgcaactg agaagtcgac aggcaacaat cttttacaca cagcagctta	2940
ttaccacaat gaagaaactg cgcaggttgt gcagcttctt gtagaagctt atatagatgc	3000
taaggaacaa ggagtggagc ttcagccgag cccttggaaca tgtgagaata tagaaggaga	3060
tacacctttg atggtatcct taatcaacaa acatgaaaaa ttggcactgt atttcatgtc	3120
tgtggatatg gagaattcag ttgtatatgc aaccaagagt gtactatatt gtgctgtact	3180
gcgtggatgt gatgaagttg cagaagaaac agtggcttca gttgatcctg cctgcttcag	3240
cttcatgcag cttaaagacg atgggtggcg aaatgtcttg catgttgcac caaattgcac	3300
aggtgaggtg agagtacctt attgcttgta tatctttctt cttatttgaa aaatcttgga	3360
ggaactaagc acaagcaagg ccaaggcttc gtcgtgctgt ggtgggctca aagaaggcct	3420
atgcacagcc tggcccatat ggggtgcaggc ggcctcttta agaaattgag gaggccgaaa	3480
cacggctggc tttttgggct tgtgctcatt ttcaaaattg atgccatta cagcccacgg	3540
agcactgctg tgatcttggt tcgtgagtca tgacaggcca gtaacgggct tgtgctgggc	3600
aagtaatatg gcataactat ccttatcatt atttataacc ggaacatctc atatatgcaa	3660
acctttaatt ctgactttga tcagtttatt atagaacatg gaaatcgaac atatctcagc	3720
agcatattca agctattgta aacttcttta aataattaag gtggtagacg tatatgcctt	3780
gtattatttc tgggtcaaat gagtatcagg gaaatctaataaaaacatattt aatctttacc	3840
tatatagcaa gaaaacagtt gggattttga aaacgggaaa atcttgagtc ccatccattt	3900
atttgcttca ttaccacgga actggggaaa tttcacatac ttgaagtttc ggtgcttcat	3960
agattctaac attgaagtgt ttatacagag agaacaggca ccttgttgggt ggaaaagcta	4020
gcttggttga tcaacgagcc agatgatgat ggaaagagac cccttgatat agcttcagaa	4080
gttggtaacg catggcttat aaaattactg ctgacaaaag acccttcctc aaacacaagt	4140
gcgccatttg cttggattga agcatgtaaa aaaggctact tatcagcaat acatgctttc	4200
atagaccatt cccctgattt tagaacattt tgtctccaaa gaaaagactc tcctttacat	4260
cacatacaac tgagaagtta caaagaatac caagaatttc ttgctattcc gttgattcaa	4320
gagatgaaaa atatgctcga tttcagtggc tcaacgccct tgcatcgagc attagaacgt	4380
aaggatatcc tccttgctga agcactgctc tctggcgatg gggttcatag aagcatcaaa	4440
gataaaaatg gtaaaactgc taccgacctg ctagtaaagc tgtgcgacca agagtatgaa	4500
tgggtatgtg tctccagtct ccacctaat cttccaatct agaaaactat gattgcatta	4560
ggaaatactt ctatgtcagt tgtcactatc atcccttgct tgctataagt ctacattgtc	4620
ttggcaaaaa cataaaatga gcataaatat aaagggttaat tttttctagg gaaattgtac	4680
tgcgtatacc ctaaactaga atttgccaaa taaccttaga gattcaaaat atatccataa	4740



ctattaactt tggctaattt gtaatgggtc tctcttatac aattttattga ttcaaacctta	4800
tttctgctga ttaaacttga tcattggtaa tattgaagaa aaatttctcat ggccttgcaa	4860
gacagtgaag tttccatgaa tctcggaaaa atgacagcca caggataaaa taagactgct	4920
ttagacagca tagtagctct tatgtttttac tatatacaag aatatgtaaa agccttgatc	4980
tatgaaatga attgttctat atatattatt atgggtgatat gcaagctcct tctttgaatt	5040
caatttcaaa caaaatgcag gatactatgt gcaaacgtac acaaattagt ccgtggctaa	5100
cgacaaacta tatcggaaact tcccttgcta ataaggcctt tagatacaca ggcagtacaa	5160
gacttggtag aacaccatca gcaggagaaa tgcgtagcac actttcagtc gtagcagccc	5220
ttctagcaac ccttacattt gcagctgggt ttacacttcc tggaggcctt aacgaagata	5280
ctggcgaagc catcttagca aagaagggtt catttctagt gtttatacta gcagacacat	5340
acgcgatgtg ttgctccatg ttggtgctgt tctgcctcat atgggtctatg gtttagcgaca	5400
gagataagtc acttctactg attgatcgaa gtgttgtgat actcgtccaa tcactttatg	5460
gaacgttaat agcatttatg gctggagttt acactgctat atcacacaag tctttgtggg	5520
cagctattat agtcattgtt atgtgctctt tcgttgcgat ttcagctaac agagctattc	5580
tggataaagt gcttgataag ttgatccctt cggctgatag taagagaaga aattaaaccc	5640
agacactgga tggatgtctg gatgatgtag gctctcctat aatctttcac tatcttatga	5700
ttttggatat tactgtctgc aaatgttaaa ctcacacatt gctattatag ttctttgtta	5760
tgcaagtatg gattcaactc tggactttgg tcagtctggt aactagttgc agcccatgac	5820
cccaactttt agttattctt atactacctc tgttttgttt ataatacatt ggtacaaatc	5880
ttattcatgt tactagatga acatgctaatt gactagtttt ataatcggct tgtagatgg	5940
agctctgagt acacatttga aaaatatatc taattaaaaa tattaaccg actttaaaact	6000
tgtacgatac atgagattat aggaaataaa gaatataagt tttctaatta ttatacgggt	6060
ttaggaaaat ttagcaatta gtttagttat tatattaatt tgatacttag attttcaatt	6120
agattatact agagtatata aaatttccta attatatgta ttcggttttc aatcagaaaa	6180
taatgattta gttttattat aaaattaaat tacttttttt agacggtgct tttctctcca	6240
aagtttcctc ctttattcac acatgctaaa catggaagaa tatatgtagc attattgttt	6300
tcacacatca ttttctaaag gttgtatgat tcttattcca aagcaaatta ccttcgataa	6360
tgttgttact cgatatctaa taaaaaactt ttccatagat gatatcaaga ccttaatgat	6420
tttaataatt atccaaataa tcgccagagc aactagtact ttttaacaa taatatattt	6480
tttttgacaa tggggtaaac aataatattt cttacataaa cattttcata ttcttagggg	6540
gaaaaaccat taagaaaaat gcatgtatct attggatctc tatacaagtt tattttgata	6600
gttcgagccc taaattttgt cagcaagtca tatgtaagat ttgtgtataa actataaagt	6660
gaactattgt ttatttatta gctatgaatt aggtttcaca aaatattata taaagttgaa	6720
tgatgattaa cggaactat actgatatta tcat ttgaga ttttctcttc atgtaaaaga	6780



ccatttatct	tatcccttat	ctctactagt	ctactttaag	ttcttgtgtt	atgtttaatt	6840
tttgtcatgt	atttacctaa	atgctagttt	tacattcaca	actccttttc	ttcactagag	6900
ctattttaaca	tttcaaaata	cgctataatt	ttatattagc	aaatataaac	gtaatgatcg	6960
ggattcctta	tattttttca	caaattatta	gaataggcgg	tctaattttt	acataaatta	7020
gatgaactta	gaagtgaatt	tttcaaacaa	acccattcca	tttcaactta	acccaacact	7080
atcttagtca	tcccttatct	tttgcttcct	ttgttttctt	gattctcgaa	ctacaacaga	7140
caattttaag	aaataactcg	gtatttttat	cgaacggatt	aaactagtca	ctaaatggat	7200
aaacaagtca	ctgaatgggt	tagtgaatgt	cattcacgaa	atagattaaa	ttggtcacga	7260
aatagagtca	ctatatttaa	aaagggtggc	tgttctctgc	tgaatattag	acttgcaccg	7320
tgcctaattt	taaaagtagg	cgatatctta	caagacaact	gtcatttttc	cacttcctta	7380
ataatgagta	atcatgttca	tgtatcatac	tccttgaaca	tgacatatat	atttttctag	7440
aatgaaaaat	cacctaacac	aaaaagggga	accaattaga	aagagagaaa	gaaaagtaac	7500
acaaacaaca	atcaaaacat	gaaaacaact	agcaaaattt	attaagtact	aattaataca	7560
tctagttacc	taaaatgcac	tctaattact	ttaaaaagtt	caaactccaa	caatagtgca	7620
aattagcata	aacacttggt	acagcaagtt	gtgcaaactc	agacacacag	accacagaag	7680
gcttgagtccc	cacaccagag	gggcaactct	aattttctca	acgtctcctt	tttctttctc	7740
ttctttctct	ttcatcttct	ttgctttctc	cctccagaat	ctttctctct	cttccatttc	7800
caggtttctc	tctcctcttc	tctgcttctt	attttttgaa	agatgcaaac	ttttactgaa	7860
atttatgttt	tgaatagtgt	tacttattgt	tatgctttta	attctgagtt	gggtcacttt	7920
catttttggt	tgaatttaat	gggtttttgc	tgaattatgc	tctttttact	ccagtgaacg	7980
gtttttcagt	ttctgggtgt	taactgtatt	tagttaatta	agattgggtt	gaattcaaaa	8040
aaaacttagg	gtttactatt	ttccatgctt	aatctttatt	ttttaatgtc	tgaatatgta	8100
aaaatgtaaa	aattctatgt	tgaaaaactg	agtaaaatag	tatcaaatca	aagttttgaa	8160
gctttgaatt	actgatatgt	tgtagtttgg	tacttgggtt	gatctgtgag	attattcata	8220
agatgctatc	tttattttct	gtttttcttt	tagtgcaaat	attctgaata	aaatatgcat	8280
tagtttactt	ttatatagaa	tataagtatt	tgggattcta	agttatggga	cactcaattt	8340
tatatgcaga	tccagctggt	ccagactaga	ctaacttggg	agcttgagct	tcacttggtg	8400
tgcttgatc	tgtaagctt	ggaagttttg	ctgatttagc	gccatatggt	ctagatgtat	8460
tgtattagtc	aagtaaagtt	ttgatatcga	aatttggacc	tttagtggca	taagagtggg	8520
ttattcttca	tttagaattt	tgaccttgag	cttagttttg	gaattgagtg	gttgagaaac	8580
ttcaaacact	ttggcttttc	agttttattat	acccgggttt	ttattgagga	aggtagtgag	8640
aaagctccag	gaaaatttga	ctcttgtgtc	tacagaaaag	tcacttagtc	ttctcctata	8700
attttgctgt	aatcctgggt	ctggacctct	aggctctgga	atggcagttg	gcaaaaacag	8760



cagtaacgct	ggatcattaa	ctcggccatg	tcattgtttc	aaggtggcaa	acttgaagga	8820
aactatTTTT	gatgctagcg	agacatccga	gttaaaagat	cgttatgttt	tgggagatca	8880
actaggttgg	gggcagtttg	gtgtgatccg	ggcatgtgct	gataagttta	ctggagaact	8940
actggcgtgc	aagtccattg	ccaaagatag	acttgaaca	caagatgatg	ttcgaagtgt	9000
gaagctcgaa	atcgagatta	tgagcaagtt	gtccgggtcat	cctcatgttg	tcgatctcaa	9060
agctgtttat	gaggaagaag	attatgtcca	cttgggtgatg	gagctttgtg	ccggtgggga	9120
gttgttccac	cgattagaga	aacaagggaag	gtattgctgag	tctcaagcca	aagtcattctt	9180
caggcatcta	atgcaagtag	tcttgatttg	tcatgataat	ggtgttggtc	atagagattt	9240
gaagcctgaa	aatgtttctt	tggcaaccaa	gtcttcttct	tcgccaatta	aattagccga	9300
ttttgggtctt	gctacatata	tcaaaccagg	tagaactctc	ttcttcatct	agtttgtgat	9360
tttagctgtg	ttactcggct	tctttcaatt	cacctcaata	gctgtccatg	ttgaattttg	9420
gataatttga	tcaagtcagt	ctggccccta	aacatgttcc	tgcgcccaac	actaacagtg	9480
tcttaagcct	ttggttactg	gtcaagcagg	ggagagtgtg	catgggacag	tggggagtcc	9540
tttctatata	gctcctgaag	ttctgtcagg	aggttacacc	caggctgctg	atgtatggag	9600
tgctggcgta	attctctaca	ttctcttgag	ttctatgcca	ccattttggg	ggaagacaaa	9660
gtcaaggata	tttgatgcag	ttcgagcagc	tgatctgcgc	ttcccttctg	aactttggga	9720
tcggatatca	gaacctgcc	aggagctgat	caggaaaatg	ctttgtgtag	atcctttgaa	9780
gcgcttgaca	gctgagcaag	ttttagggtat	atttttaatt	tttgccctct	ttgctgaatt	9840
cagatgacag	ttatatgaca	cagtatacat	ttgtagaacc	caagtgtctg	aatagccaat	9900
caccaatgtc	tgacaagttt	tttttggtctt	tcgattacaa	aatcatttat	tacatatttg	9960
cattaactgt	gttatTTTT	acacattaca	tgaaatcttc	attgcttatt	tgtgattctt	10020
tgtgaattgc	tgtacattgg	aagctgtccc	cttttaacat	ttgagatgtc	cgtaagtggg	10080
gagtgttaact	catctgtccc	cacggagaca	aagcttttgg	atgattttag	gacaaatgca	10140
tttaacgttc	tcagcttatc	tgacaatgtt	atccatgggtg	tccgtggaca	catgtaccaa	10200
cactaaacat	ctagtatttc	ttgttggttc	tttagttttg	ctggatatac	ttaagggcta	10260
aggctttatt	tttgttcctg	cagctcactc	atggatggaa	gaggttactg	tagctacgga	10320
agaatcacat	gaacatgatt	tggcctgctc	tgaacattta	aaaacccgag	atagctcatt	10380
ttcagcgtca	tgtatatcca	gggatcacga	tataagcttt	ggcactggat	ctgcggtaaa	10440
ttgtgaacct	caatctccaa	catttacgtg	cagatcttca	ttctcggcat	ttatggcgga	10500
accatccacc	cctaccctta	tatctgctgg	attttctttc	cgtagcagcg	gtgatttcac	10560
tgctcttgag	tttgtttctc	caattccttc	cttgccctagc	ttcacatttt	tcagccctag	10620
gtcagtagat	gagcatggaa	acaaaaataa	ggttttttca	agcaacaggg	agaacactga	10680
cgaaattcat	acaggtaact	cacttttcta	tttgtaataa	tttgttattg	tcactgattc	10740
tctaattgtca	tttttgcata	tctaggtctt	gtaacatttt	aaacgtcata	taaacatcca	10800



```

tgacatttta tgtattttgc tcaactttca agtatttttg tatgactttt aaactttcaa 10860
gttttatctg cgtataactt ttcatagcct aaccgtttct ctgttgata agtatatgcg 10920
aaaaatcatt catcacatga attgtttgaa tctaattgtg gattgtgcag cctaaccaac 10980
tcgagtttct taaaaaaaaat tgttctgtaa ttcaattatg tcctccatat cttatctttc 11040
tttactttta aagcactatt gagaaacaga agaaaacccc agtgtgacga ctgtcagatg 11100
tggtttgatt gagattagac tattttatca gtgttggtt attgagattc cagtgtgtt 11160
aatatcatta gttccatcgt aagcttttaa agttatcaat tcataaaact agatctaattg 11220
cccacgtat ttttagaagg ggccgacttg gagaagcgat ttgaatcacc tcattcatcg 11280
ctgtgtcag gaattgatgc tagggacctg aaagagaagt cagcagactc taagaggagt 11340
ggaggaacag gggtgaggat ttttgggac cataacaaga gaaataggac gattgggctt 11400
ggcgaattca accagcttga tattgtggtc actgaatctg tcatccgatg ggcgtcatgc 11460
accacttac ctaccgcctc atcactcagg tcttctcttg tatgctgata atcatgggcc 11520
gaccacaatg tatcaaatta acatcataaa tcatgatata aagtgggca acacgcaaac 11580
gtgtgaattc tactgtctctg ctacaagatt gaagatataa tgggtttgag tcgcgtgtac 11640
tgttggtgat gatcatccca tattagagag ctaaatgtta gtaactacat tgtaatggat 11700
ctgcagaata agcagattct tcttcaaaag gtgtaaagca tggattttga aagcaatggt 11760
tttctcaacc tttttgatca tgatttagta gatatatata gggggctttt tgttcttatt 11820
agatgtgatt gttaagcctt cttcatgaac aaacataagc actggtgact tgtgggatgg 11880
tacatagaaa gagttccgtt tcatcgattt tgattttgtc cagatcctcc ttttctttcg 11940
aaaaaagggg aggagagtat ccgaaagatt acccattttg tattttcggg tgtctttaaa 12000
agattgatcc attaccttc ggacaatttg attcctataa tactcctatt tctctacttc 12060
atttctgcc atatactact atactaaaga caaaattact taatacgaac ttaaaattaa 12120
gaggagaaac aaggtaaaat ttggtgagga tgatggagga atcatatgaa agcttacgca 12180
cattatatat ttgggagtat gatgtacact agttaatgaa gatcagacac ttttctatat 12240
cttttgattc gtatgttctt faataaatga agatgggata tttaaaattt cggtgtgtatt 12300
tgaaaatatc atgcacatta gacatccaag tctaagatta attcacaact atctattttt 12360
tttt 12364

```

```

<210> 6
<211> 4659
<212> DNA
<213> Beta vulgaris

```

```

<220>
<221> misc_feature
<222> (461)..(485)
<223> n is a, c, g, or t

```

<220>  
 <221> misc\_feature  
 <222> (621)..(696)  
 <223> n is a, c, g, or t

<220>  
 <221> misc\_feature  
 <222> (704)..(707)  
 <223> n is a, c, g, or t

<220>  
 <221> misc\_feature  
 <222> (1115)..(1123)  
 <223> n is a, c, g, or t

<220>  
 <221> misc\_feature  
 <222> (1169)..(1170)  
 <223> n is a, c, g, or t

<220>  
 <221> misc\_feature  
 <222> (1879)..(1880)  
 <223> n is a, c, g, or t

<220>  
 <221> misc\_feature  
 <222> (2194)..(2194)  
 <223> n is a, c, g, or t

<400> 6  
 caaatcttct ggcataatg gcggtgttgc cgttcatacaw tttaacatca atggaggtaa 60  
 gaggcatgtt ttttcaacaa tataaaactt atatgatttt tctgtttttc cccgtatctt 120  
 gtttgagatt gtgattatta agagagtgca gtctcacaat attcgaaagt ctacgtaatc 180  
 cacctcaaat tgacgaagaa aacaagcagg aaaggattaa gtaagttcgt ggaaccayta 240  
 gaattaattt tcaaataatag ctctaccta tatatggcct acttttaatt ttaaataaga 300  
 agaaggtaat gtgattagaa acaaattggt cttaaattat tcattaagct taataatgca 360  
 taaactttat caagtgttat cttcttttca gggccgtctt gaagattttk gggcccrgtt 420  
 ctattatgaa aattgrgccc cttaaattat asaaaataaa nnnnnnnnnn nnnnnnnnnn 480  
 nnnnngatgg aagggttagag yyctaaagat agaaagttga aaatctaaat ataaatcatt 540  
 gacaaattta ttaagggtga gaaacaaggg tgttttcttc aaatatgaag caaaattttc 600  
 aaaataatat awtttkggsc nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660  
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnncttc tttnnntga tacaattttt 720  
 sagcgacatg attgtcgatt gatgcttata ttattgtata ctcgatccat attgtttaag 780  
 atgaattgtt tgtctttgat ggtctccaat gcatattttg tatwmtagg aattytaatt 840  
 atgtactatt agtasayaty gagaygaata caraatygcc ataataagat atgattattt 900  
 tarttatata ctttctccgt tccaaatata taartgtaac acttgtgtac tttatgcgta 960  
 ctaatgcata ayaacgtgca ctctccygtg tttaattata tactttttga gagaagtgwt 1020  
 acattgggga ccatgggact gtgtataatt tgaccgcaa atygaagtgt ygcatttgat 1080



tgaaaaygga garrgtagta tatagrtgga acacnnnnnn nnntgctgrt catctttggc	1140
caacaaaccm aaaattgata ttaatccynn twwtymrgkw yttttcatct ttttgacaca	1200
aatggatgt tgtaggcwct gcgctatctg ctgcccaatc tctgtttgca gccctgcaaa	1260
gttctgagct caaagagatc ctctcgatct ttggctacaa atcccaactt gatgacctcc	1320
aacgcaytgt mtctaccatc aaygctgtat tccgtgatgc tgagaccaa caggagctca	1380
ctcatgaagc acarcatthg ctcgaggaac tcaaggatgc tgtctttgaa gcagatgatc	1440
tgttcgacga gtttgtcact cttgccgagc agaagcaact tgtagaggct ggtggcagtc	1500
tttccaaaaa gatgcgcaa ttcttttctg attccaaccc ctttggcaty gcttatarga	1560
tgtcacragg ggttaagaag atcaagaaga agttggatgy tatygcttac aatcatcaat	1620
ttagctttta gattgatctt gagcctataa aagagagaag gctcgagact ggttctgtcg	1680
tgaacgcagg tgatatcatt ggaagagagg atgacttgga gaagatcgta ggtttgttkc	1740
ttgattctaa catccagcgt gatgtgtctt tccttackat wgtgggaatg ggagggttg	1800
gtaaaactgc tcttgcccaa ctggtgtaca atgatccaag ggtcagaact gcttttccat	1860
tgagatgttg gaattgtsnn tctgatcaag atcaaaakma actagatgtg aaagaaattt	1920
tgggtaagat tctgtctaca gctactggta agaatacayra gggttcaacc atggatcakg	1980
tgcaaaccya actacrrgaa caactatgtg gcaagagata cttgcttggt ttggatgatg	2040
tatggaatga gaatcctaata caattgcgrt wyytkgkwra attcttcatg ggagggtcaa	2100
ggggaaattg gattstggta actacgcgtt cgaygagac arcgagaatt ataagagatg	2160
gtccattgca caagctscas ggtttgtctg arrnaaaact yttggcgttt atytgtaagg	2220
tggaccttcg gatcagtgca accaaaattc cctaataact ttgtcatgat tgcacgagat	2280
atagtygaca aatgtgctcg aaaccctytg gctataagag tggtaggaag tcttttgtgt	2340
ggtcaagaca agagtaagt gctttcatth catgagatmt gtttagccaa cattagaaag	2400
agycataatg atatcatgcy aatactgaac ctaagttacc atcatcttga acctccaatc	2460
akgagatgct ttagttattg tgcartgttt ccaaaggatt tccttatagg gaagaagacg	2520
ttgataaacc tttggatggc acaaggttat attgttccat tagacaaaga tcaaagcata	2580
gatgaygcta gtgaggaata catatcaatt ttgytgcaga gatgtttttt cgaaaacatc	2640
ggaacagaaa aagatkatgt tattaagata catgatctca tgcagatgat tgctcaaaat	2700
gtcatgggga aggagctttg tacgacaaaa aacattagtg gcagcttgga taaaaatgtt	2760
cgccatctat ctcttgccag aactagtttt gcaagatact ctttyaatgc aactcatatt	2820
cgctccyatt tctrtgctgg ctactggtgt caggawktg agataamcca gtttgcagtt	2880
gaggcattag taccaaaytg tttgtgccta agggcattgk acctsgcttg gtcgaagata	2940
aaaagtktac cagactcrat tgggtggattg ttgcatttga ggtacttaga tctttcrtat	3000
aasgaagaty tggaagtact tccgaactca attgcyaaac tatataatct rcaaacctta	3060
caattgaagg gttgcaagag attggaaggg ttaycaaaac atttgagcag gctgggttaag	3120



cttcaaactt trgatatata tggttgcaay aatgtaactt atatgccc aa aggcattgggt 3180  
 aagatgactt gccttcacac tctcagtaag tttatagtgg gtggagaagg garttgttca 3240  
 agttggaagm aayggtttga tgggcwggaa gatctaaagg ctctcaacaa cctaaagggt 3300  
 catctggraa tccaaatcag gtggcccga aatactacag atgctgtcaa ggaagatgtt 3360  
 aagagggaag gattatacyt gaatcataag gaacatctca atcacattgt ggttgatttc 3420  
 agatgtgagg aggggtggtg aagaatggat gatgaggaag caagaagatt gatggaagag 3480  
 ytgccggcac atccttatct tgaaaatttg gctgtgaaag cataytatgg tgygaaaayg 3540  
 cctgrttggg yaacccttct yccaaatctt acagagcttt wtcttcttga ttgtggggaa 3600  
 yyggagwrcc ttccatgcm tggaaacttg gwtgdtctra amgtctccg rctttcgcatt 3660  
 ttggcraaat tggagtayat tgragaagat agcwcattcag ctmwtttcag ktktaggcct 3720  
 ggaccrgaaa gtgcaggact atcattatac tccccctccc ttgaackcct tgagttgaag 3780  
 crtttgyrya agttaaaagg atggaggaga rgggaagggt taggagatga tcaccagcct 3840  
 tttaatgaaa gcagcagcaa taagtcattg agaatagaaa gatgcccatt gctgacattt 3900  
 atgccgctgt gtcccaagac agaaaaacd g catttagttg tatttaata ga aygactccgg 3960  
 atagtgcata ctaaaggaga tgagaatttc tatgctccat tacattcatc atcatctgat 4020  
 cctgaaaacc cgaggagcac tattcccatt cccatgttaa gagaggata cataaacaat 4080  
 gtggcatggc taaattcgct gcctatggag gcttttaggt gtctcactca tatgacaata 4140  
 aaaaacgaca aggtagagag tttgggagaa gttggggagg tgtttcggag ctrctcatct 4200  
 tctttg'cgat ccttgaatat cacaggttgc tccaacttaa gaagtgtt'c tggagggctg 4260  
 gagcatctca ctrctttgga gatkttagaa atatacgaca cccataagct gagtctwtca 4320  
 gaagacccag aagggtgtgt gccatggaaa tcccttcac actccctcag ctacttgmaa 4380  
 ttgatgaatc tcccwcagct ggtcaacctg cctgattcga tgcagttctt ggyctccctc 4440  
 caaaccttt caatggtgca ttgcagtaaa ctggaatcag tgccagattg gatgcccmga 4500  
 ctactttcyc tcaggaagct tatggtttca ttctgttccg cacatctgga gagaagatgy 4560  
 caaatccaa ctgggggtgga ctggcctaac attcaacaca tccccscat tgatgtcacc 4620  
 tctagccgct ctaagttttt agtgttgccg tatgaatag 4659

<210> 7  
 <211> 434  
 <212> DNA  
 <213> Beta vulgaris

<400> 7  
 gaactcaagg atgctgtctt tgaagcagat gatctgttcc acgagtttgt cactcttgcc 60  
 gagcagaagc aacttgtaga ggctggtggc agtctttcca aaaagatgcg ccaattcttt 120  
 tctgattcca acccccttgg cattgcttat aggatgtcac gaggggttaa gaagatcaag 180  
 aagaagttgg atgctatcgc ttacaatcat caatttagct ttaagattga tcttgagcct 240

atgaaagaga gaaggctaga gactggttct gtcgtgaacg caggtgatat cattggaaga	300
gaggacgact tggagaagat tgtaggtttg ttgcttgatt ctaacatcca acgtgatgtg	360
tctttcctta ctattgtggg aatgggaggg ttgggtaaaa ctgctcttgc ccaactcgtg	420
tacaatgatc caag	434

#### ФОРМУЛА ВИНАХОДУ

- 5 1. Молекула нуклеїнової кислоти, що кодує поліпептид, здатний зумовлювати резистентність до вірусу некротичного пожовтіння жилок буряка (BNYVV) у рослини роду буряків, в якій експресується вказаний поліпептид, яка **відрізняється** тим, що молекула нуклеїнової кислоти містить нуклеотидну послідовність, вибрану з групи, що складається з:
  - а) нуклеотидної послідовності, що кодує поліпептид, який включає послідовність амінокислот
  - 10 відповідно до SEQ ID NO: 2 або SEQ ID NO: 3,
  - б) нуклеотидної послідовності, що включає послідовність, яка кодує поліпептид, відповідно до SEQ ID NO: 1,
  - в) нуклеотидної послідовності, що гібридизується з комплементарною нуклеотидною послідовністю відповідно до а) або б) у жорстких умовах, та
  - 15 г) нуклеотидної послідовності, що кодує поліпептид, який містить заміну, делецію та/або приєднання однієї або більше амінокислот в порівнянні з SEQ ID NO: 2 або SEQ ID NO: 3 або з поліпептидом, кодованим SEQ ID NO: 1, причому нуклеотидною послідовністю є нуклеотидна послідовність, що кодує принаймні один домензв'язуючий нуклеотид (NBS) відповідно до позицій амінокислот 168-227 у SEQ ID NO: 2 або 182-241 у SEQ ID NO: 3, принаймні один
  - 20 лейцинзбагачений домен (LRR), що відповідає позиціям амінокислот 591-613 у SEQ ID NO: 2 або 605-627 у SEQ ID NO: 3, та/або принаймні один внутрішній повторюваний домен (IR), що відповідає позиціям амінокислот 1013-1072 у SEQ ID NO: 2 або 1027-1086 у SEQ ID NO: 3.
2. Вектор, що містить молекулу нуклеїнової кислоти за п. 1.
3. Клітина-хазяїн, що містить молекулу нуклеїнової кислоти за п. 1, в якій вищезгадана молекула
- 25 нуклеїнової кислоти є гетерологічною до вищезгаданої клітини-хазяїна.
4. Клітина-хазяїн, що містить вектор за п. 2.
5. Поліпептид, що здатний зумовлювати резистентність до вірусу некротичного пожовтіння жилок буряка (BNYVV) у рослини роду буряків, в якій експресується вказаний поліпептид, що кодується молекулою нуклеїнової кислоти за п. 1.
- 30 6. Рослинна клітина рослини роду буряків, що містить гетерологічну молекулу нуклеїнової кислоти за п. 1.
7. Рослинна клітина рослини роду буряків, що містить вектор за п. 3.
8. Рослина роду буряків або її частина, що включає рослинну клітину за п. 7.
9. Рослина або її частина за п. 8, де рослина є рослиною роду буряків, за виключенням рослин,
- 35 що належать до *Beta vulgaris* subsp. *maritima*.
10. Рослина роду буряків або її частина, що містить рослинну клітину за п. 7.
11. Насіння рослини роду буряків, що містить молекулу нуклеїнової кислоти за п. 1 як трансген.
12. Насіння рослини за п. 11, де рослина є рослиною роду буряків, за виключенням рослин, що належать до *Beta vulgaris* subsp. *maritima*.
- 40 13. Насіння рослини роду буряків, де насіння містить молекулу нуклеїнової кислоти за п. 1 у вигляді інтрогредованої області.
14. Насіння рослини за п. 13, де рослина є рослиною роду буряків, за виключенням рослин, що належать до *Beta vulgaris* subsp. *maritima*.
15. Спосіб одержання трансгенної рослинної клітини рослини роду буряків, що включає
- 45 гетерологічну молекулу нуклеїнової кислоти за п. 1, який **відрізняється** тим, що містить стадію включення молекули нуклеїнової кислоти за п. 1 до рослинної клітини.
16. Спосіб одержання трансгенної рослини роду буряків, який **відрізняється** тим, що він включає наступні стадії:
  - а) включення молекули нуклеїнової кислоти за п. 1 у рослинну клітину та
  - 50 б) регенерацію трансгенної рослини із трансгенної рослинної клітини, одержаної на стадії а).
17. Спосіб ідентифікації молекули нуклеїнової кислоти, що кодує білок, здатний зумовлювати резистентність до вірусу некротичного пожовтіння жилок буряка (BNYVV) у рослинах роду буряків, в яких експресується вказаний білок, який **відрізняється** тим, що він включає наступні стадії:



- 5 i) визначення відсутності інсерції в кодуючій нуклеотидній послідовності молекули нуклеїнової кислоти за п. 1, або  
ii) визначення принаймні одного поліморфізму відповідно до Фігур 1, 2 та/або 3 у нуклеотидній послідовності, яка кодує поліпептид, що має послідовність амінокислот згідно з SEQ ID NO: 2 або SEQ ID NO: 3, за допомогою молекулярних маркерів, що визначають поліморфізм.

		1	70
consensus-sensitive	(1)	CAAACTCTTCGGCATCAATGGCGGTGTTGCCGTTTCATCAWTTTAAACATCAATGGAGGTAAGAGTCATGTT	
resistant sequence	(1)	CAAACTCTTCGGCATCAATGGCGGTGTTGCCGTTTCATCA <u>ATT</u> TAAACATCAATGGAGGTAAGAGTCATGTT	
		71	140
consensus-sensitive	(71)	TTTTCAACAATATAAACTTATA-----TGATTTTCTGTTTTTCCCC-----	
resistant sequence	(71)	TTTTCAACAATATAAACTTATACTTCCTCTGTTCTGTTTTAAATGAACGTTTGTTTTCTACGCAACC	
		141	210
consensus-sensitive	(113)	-----	
resistant sequence	(141)	CAACCCACITTTTAAATAATAAATATTTTAGTTGTTGTCACGTAATAAATATAAAAAAGTTATAATTG	
		211	280
consensus-sensitive	(114)	---GTATCTTGTTTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAATATTCGAAAGTCTACGTAATC	
resistant sequence	(211)	ATAAGTATCTTGTTTGAGATTGTGATTATTAAGAGAGTCAAGTCTCACAATATTCGAAAGTCTACGTAATC	
		281	350
consensus-sensitive	(181)	CACCTCAAATTGACGAAGAAAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAACCA <sup>1</sup> YTAGAATTAA <sup>2</sup> TTT	
resistant sequence	(281)	CACCTCAAATTGACGAAGAAAAACAAGCAGGAAAGGATTAAGTAAGTTCGTGGAACCACTAGAATTGATT	
		351	420
consensus-sensitive	(251)	TCAAATATAGCTCTACCTAATATATGGCCTACTTTTAATTTTAAATAAGGAAGGTAATGTGATTAGAA	
resistant sequence	(351)	TCAAATATAGCTCTACCTAATATATGGCCTACTTTTAATTTTAAATAAGGAGAAGGTAATGTGATTAGAA	
		421	490
consensus-sensitive	(321)	ACAAATTGGTCTTAAATTATTCA <sup>1</sup> TAAAGCTTAA <sup>2</sup> TAATGTCATAAACTTATCAAGTGCTATCTTCTTTTCA	
resistant sequence	(421)	ACAAATTGGTCTTAAATTATTCA <sup>1</sup> TAAAGCTTAA <sup>2</sup> TAATGTATAACATATCAAGTGCTATCTTCTTTTCA	
		491	560
consensus-sensitive	(391)	GGGCCGTCTTGAAGATTTTGGGKCCCRGTTCTATTATGAAAATTG <sup>1</sup> GGCCCTAAATTTATASAAAAATAAA	
resistant sequence	(491)	GGGCCGTCTTGAAGATTTTGGGGCCCGGTTCTATTATGAAAATTGGGCCCTAAATTTATAGAAAATAAA	

Fig. 1A

		561		700
consensus-sensitive	(461)	NNNNNNNNNNNNNNNNNNNNNNNNNNNGATGGAAGGTTAGAGYYCTAAAGATAGAAAAGTTGAAAATCTAAMT		
resistant sequence	(561)	-----GATGGAAGGTTAGAGTTCTAAAGATAGAAAAGTTGAAAATCTAAAT		
		631		700
consensus-sensitive	(531)	ATAAATCATTGCACAAATTTATTAAAGGGTGAGAAACAAGGGTGTTTTCTCAAATATGAAGCAAAATTTTC		
resistant sequence	(606)	ATAAATCATTGCACAAATTTATTAAAGGGTGAGAAACAAGGGTGTTTTCTCAAATATGAAGCAAAATTTTC		
		701		770
consensus-sensitive	(601)	AAAAAANNNNNNNNNNNNNNNNNNNN-----		
resistant sequence	(676)	AAAAATAACTTCCTCCGTTCTAAATAAGTGCAACATTGTCATAATGTTTACTATTACAGTTTAAACT		
		771		840
consensus-sensitive	(610)	-----		
resistant sequence	(746)	TTAATTAGCTTTGGTGATTACATTTTAGGAAAAAACATAGTCATGTGGGATCTTGTTAGATTGCTCTGA		
		841		910
consensus-sensitive	(648)	-----		
resistant sequence	(816)	ATGTGAATTTTATAATATCAACTTTTTATAATTTTACTTATTGATAATTGAAGATATTAATGGTTAAA		
		911		980
consensus-sensitive	(691)	-----NNNNNNNN		
resistant sequence	(886)	ATAATGCATTGGCAAACGTGAAACAAGAAGTGTGCACTATTAGAAACGAGGAAGTATATTTTGGG		
		981		1050
consensus-sensitive	(691)	NNNNNTATAATTTTGGNNCTTCNNNNNTTT		
resistant sequence	(956)	CCTTTATAATTTTGGAGACCCTGCGCTGTTGGGCTCCTTGACACCCITCATCTACCCCCTGCTTC---TTT		
		1051		1120
consensus-sensitive	(709)	GATACAATTTTTSAGCGACATGATTGTCGATTGATGCTTATATTATTGTATACTCGATCCATATTGTTTA		
resistant sequence	(1026)	GATACAATTTTTCAGCGACATGATTGTCGATTGATGCATATATTATTGTATACTCGATCCATATTGTTT		

Fig. 1B

consensus-sensitive resistant sequence	(779) (1096)	1121 AGATGAATTGTTTGTCTTTGATGGTCTCCAATGCATATTTTGTATWMTTAGGAATTYTAATTATGTACTA	1190
		AGATGAATTGTTTGTCTTTGATGGTCTCCAATGCATATTTTGTATACCTTAGGAATTCTAATTATGTACTA	
consensus-sensitive resistant sequence	(849) (1166)	1191 TTAGTASAYATYGAGAYGAATACARAATYGCCATAATGAAGTATGATTATTTTARTATATACTTTCTCC	1260
		TTAGTAGACATTGAGATGAATACATAAATGCCATAATGAAGTATGATTATTTTATATATACTTTCTCC	
consensus-sensitive resistant sequence	(919) (1236)	1261 GTTCCAAATATATAAATGTAACACTTGTGTACTTTATGCGTACTAATGCATAAAYACGTGCACTCTCTYG	1330
		ATTCCAAATATATAAATGTAACACTTGTGTACTTTATGCGTACTAATGCATAACAACGTGCACTCTCTTG	
consensus-sensitive resistant sequence	(989) (1306)	1331 TGTTTAATTATATACTTTTGTAGAGAAGTGTTACATTGGGGACCATGGGACTGTGTATAATTGACCGCA	1400
		TGTTTAATTATATACTTTTGTAGAGAAGTGTTACATTGGGGACCATGGGACTGTGTATAATTGACCGCA	
consensus-sensitive resistant sequence	(1059) (1376)	1401 AAATYGAAGTGTGCACTTTGATTGAAAAAYGGAGARRGTAGTATATAGRTGGAACACNNNNNNNNNTGCTG	1470
		AAATCGAAGTGTGCACTTTGATTGAAAAAYGGAGAGTAGTATATAGRTGGAACACAGCAGAGACTGCTG	
consensus-sensitive resistant sequence	(1129) (1446)	1471 RTCATCTTTGGCCAAACAAACCAAATTTGATATTAATCCYYTWWTYMRGKYWNNTTTCATCTTTTGTACA	1540
		GTCATCTTTGGCCAAACAAACCAAATTTGATATTAATCCCTTATTCAAGGTCA--TTTCATCTTTTGTACA	
consensus-sensitive resistant sequence	(1199) (1514)	Start2 CAAAATGGATGTTGTAGGCWCTGCGCTATCTGCTGCCAATCTCTGTTTGACGCCCTGCAAAGTTCTGAG	1610
		CAAAATGGATGTTGTAGGCWCTGCGCTATCTGCTGCCAATCTCTGTTTGACGCCCTGCAAAGTTCTGAG	
consensus-sensitive resistant sequence	(1269) (1584)	1611 CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCCACTTGATGACCTCCAACGCAYTGTMTCTACCA	1680
		CTCAAAGAGATCCTCTCGATCTTTGGCTACAAATCCCACTTGATGACCTCCAACGCATGTCTCTACCA	

## Φir. 1C

consensus-sensitive resistant sequence	(1339) (1654)	1681 TCAAYGCTGTATTCCGTGATGCTGAGACCAACAGGAGCTCACTCATGAAGCACARCATTGGCTCGAGGA	1750
		TCAACGCTGTATTCCGTGATGCTGAGACCAACAGGAGCTCACTCATGAAGCACAGCATTGGCTCGAGGA	
consensus-sensitive resistant sequence	(1409) (1724)	1751 ACTCAAGGATGCTGTCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAGCAA	1820
		ACTCAAGGATGCTGTCTTTGAAGCAGATGATCTGTTGACGAGTTTGTCACTCTTGCCGAGCAGAAGCAA	
consensus-sensitive resistant sequence	(1479) (1794)	1821 CTTGATAGAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCAATTCTTTCTGATTCCAACCCCTTTGGCA	1890
		CTTGATAGAGGCTGGTGGCAGTCTTTCCAAAAAGATGCGCCAATTCTTTCTGATTCCAACCCCTTTGGCA	
consensus-sensitive resistant sequence	(1549) (1864)	1891 TYGCTTATARGATGTCACRAGGGGTTAAGAAGATCAAGAAGAAGTTGGATGYTATYGCTTACAATCATCA	1960
		TTGCTTATAGGATGTCACGAGGGGTTAAGAAGATCAAGAAGAAGTTGGATGCTATCGCTTACAATCATCA	
consensus-sensitive resistant sequence	(1619) (1934)	1961 ATTTAGCTTTAAGATTGATCTTGAGCCTATTAAGAGAGAAGGCTGAGACTGGTTCTGTCGTGAACGCA	2030
		ATTTAGCTTTAAGATTGATCTTGAGCCTATTAAGAGAGAAGGCTGAGACTGGTTCTGTCGTGAACGCA	
consensus-sensitive resistant sequence	(1689) (2004)	2031 GGTGATATCATTGGAAGAGAGGATGACTTGGAGAAGATCGTAGGTTTGTTCCTTGATTCTAACATCCAGC	2100
		GGTGATATCATTGGAAGAGAGGATGACTTGGAGAAGATCGTAGGTTTGTTCCTTGATTCTAACATCCAGC	
consensus-sensitive resistant sequence	(1759) (2074)	2101 GTGATGTGTCTTTCTTACKATWGTGGGAATGGGAGGGTTGGGTAAAAGTCTCTTGCCCAACTCGTGTA	2170
		GTGATGTGTCTTTCTTACTATTGTGGGAATGGGAGGGTTGGGTAAAAGTCTCTTGCCCAACTCGTGTA	
consensus-sensitive resistant sequence	(1829) (2144)	2171 CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGTNNCTCTGATCAAGATCAAAAG	2240
		CAATGATCCAAGGGTCAGAACTGCTTTTCCATTGAGATGTTGGAATTGTGCTCTGATCAAGATCAAAAG	

## Φir. 1D

		2241		2310
consensus-sensitive	(1899)	MAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTGCTACAGCTACTGGTAAGAATCAYRAGGGTTCAA		
resistant sequence	(2214)	CAACTAGATGTGAAAGAAATTTTGGGTAAGATTCTGCTACAGCTACTGGTAAGAATCATGAGGGTTCAA		
		2311		2380
consensus-sensitive	(1969)	CCATGGATCAKGTGCAAAACCYAACTACRRGAACAACATATGTGGCAAGAGATACTTGCTTGTGTTTGGATGA		
resistant sequence	(2284)	CCATGGATCAGGTGCAAAACCCAACTACGAGAACAACTATGTGGCAAGAGATACTTGCTTGTGTTTGGATGA		
		2381		2450
consensus-sensitive	(2039)	TGTATGGAATGAGAATCCTAATCAATTGCGTDWYYTKGKWRRAATTCCTCATGGGAGGTCGAAGGGGAAAT		
resistant sequence	(2354)	TGTATGGAATGAGAATCCTAATCAATTGCGTGATCTGGTAGAATTCCTCATGGGAGGTCGAAGCAGAAAT		
		2451		2520
consensus-sensitive	(2109)	TGGATTSTGGTAACACGCGTTCGCAYGAGACARCGAGAATTATAAGAGATGGTCCATTGCACAGCTSC		
resistant sequence	(2424)	TGGATTGTGGTAACACGCGTTCGCACGAGACAGCGAGAATTATAAGAGATGGTCCATTGCACAGCTCC		
		2521		2590
consensus-sensitive	(2179)	AAGGTTTGTCTGARRNAAAACCTTTTGGCGTTTATYTGTAGGTGGACCTTCGGATCAGTGCAACCAAAAT		
resistant sequence	(2494)	AAGGTTTGTCTGAGG-AAAACCTTTTGGCGTTTATTTGTAGGTGGACCTTCGGATCAGTGCAACCAAAAT		
		2591		2660
consensus-sensitive	(2249)	TCCCTAATGACTTTGTCATGATTGCACGAGATATAGTYGACAAATGTGCTCGAAACCCCTTGGCTATAAG		
resistant sequence	(2563)	TCCCTAATGACTTTATCATGATTGCACGAGATATAGTTGACAAATGTGCTCGAAACCCCTTGGCTATAAG		
		2661		2730
consensus-sensitive	(2319)	AKTGGTAGGAAGTCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTTCATTTTCATGAGATMTGTTTAGCC		
resistant sequence	(2633)	AGTGGTAGGAAGTCTTTTGTGTGGTCAAGACAAGAGTAAGTGGCTTTTCATTTTCATGAGATCGATTAGGC		
		2731		2800
consensus-sensitive	(2389)	AACATTAGAAAGAGYCATAATGATATCATGCGYAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA		
resistant sequence	(2703)	AACATTAGAAAGAGCCATAATGATATCATGCGCAATACTGAACCTAAGTTACCATCATCTTGAACCTCCAA		

## Φir. 1E

		2801		2870
consensus-sensitive	(2459)	TCAAGAGATGCTTTAGTTATTGTGCACTGTTTCCAAAGGATTTCTTATAGGGAAGAGACGTTGATAAA		
resistant sequence	(2773)	TCAAGAGATGCTTTAGTTATTGTGCACTGTTTCCAAAGGATTTCTTATAGGGAAGAGACGTTGATAAA		
		2871		2940
consensus-sensitive	(2529)	CCTTTGGATGGCACAAGGTTATATTGTTCCATTAGACAAAGATCAAAGCATAGATGAGTCTAGTGAGGAA		
resistant sequence	(2843)	CCTTTGGATGGCACAAGGTTATATTGTTCCATTAGACAAAGATCAAAGCATAGATGAGTCTAGTGAGGAA		
		2941		3010
consensus-sensitive	(2599)	TACATATCAATTTTGTGCGAGAGATGTTTTTCGAAAACATCGGAAACAGAAAAAGATGATGTTATTAAGA		
resistant sequence	(2913)	TACATATCAATTTTGTGCGAGAGATGTTTTTCGAAAATGTCGGAGCAGAAAAAGATGATGTTATTAAGA		
		3011		3080
consensus-sensitive	(2669)	TACATGATCTCATGATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGACAAAAACATTAG		
resistant sequence	(2983)	TACATGATCTCATGATGATATTGCTCAAAATGTCATGGGGAAGGAGCTTTGTACGACAAAAACATTAG		
		3081		3150
consensus-sensitive	(2739)	TGGCAGCTTGGATAAAATGTTTCGCCATCTATCTCTTCCAGAACTAGTTTTGCAAGATACTCTTTTAAAT		
resistant sequence	(3053)	TGGCAGCTTGGATAAAATGTTTCGCCATCTATCTCTTCCAGAACTAGTTTTGCAAGATACTCTTTTCAAT		
		3151		3220
consensus-sensitive	(2809)	GCAACTCATATTGCTCCYATTTCTRTGCTGGCTACTGGTGTGAGGAWKCTGAGATAAMCCAGTTTTCAG		
resistant sequence	(3123)	GCAACTCATATTGCTCCCTATTCTGCTGGCTACTGGTGTGAGGATGCTGAGATAAACCCAGTTTTCAG		
		3221		3290
consensus-sensitive	(2879)	TTGAGGCATTAGTACCAAATYGTGTTGTCCTAAGGGCAATGKACCTSGCTTGGTCGAAGATAAAAAAGTKT		
resistant sequence	(3193)	TTGAGGCATTAGTACCAAATCTGTTGTCCTAAGGGCAATGACCTCGCTTGGTCGAAGATAAAAAAGTTT		
		3291		3360
consensus-sensitive	(2949)	ACCAGACTCATTGGTGGATTGTTGCATTGAGGTACTTAGATCTTTCTATAASGAAGATYTGGAAGTA		
resistant sequence	(3263)	ACCAGACTCGATTGGTGGATTGTTGCATTGAGGTACTTAGATCTTTCTATAACGAAGATCTGGAAGTA		

## Φir. 1F



		3361		3430
consensus-sensitive	(3019)	CTTCCAACTCAATTGCGYAAACTATATAATCTRCAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG		
resistant sequence	(3333)	CTTCCAACTCAATTGCTAAACTATATAATCTCAAACCTTACAATTGAAGGGTTGCAAGAGATTGGAAG		
		3431		3500
consensus-sensitive	(3089)	G GTTAYCAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTTGATATAATGGTTGCAAYATGTAAC		
resistant sequence	(3403)	G GTTACCAAAACATTTGAGCAGGCTGGTTAAGCTTCAAACCTTTGATATAATGGTTGCAACATGTAAC		
		3501		3570
consensus-sensitive	(3159)	TTATATGCCCCAAAGGCATGGGTAAGTGGACTTGCCTTCACACTCTCAGTAAGTTTATAGTGGGTGGAGAA		
resistant sequence	(3473)	TTATATGCCCCAAAGGCATGGGTAAGTGGACTTGCCTTCACACTCTCAGTAAGTTTATAGTGGGTGGAGAA		
		3571		3640
consensus-sensitive	(3229)	G GGARTTGTTC AAGTTGGAAGMAAYGTTTGATGGGTGGGAAGATCTAAAGGCTCTCAAACCTAAAGG		
resistant sequence	(3543)	G GGAGTTGTTC AAGTTGGAAGCAATGTTTGATGGGTGGGAAGATCTAAAGGCTCTCAAACCTAAAGG		
		3641		3710
consensus-sensitive	(3299)	GTCATCTGGRAATCCAAATCAGGTGGCCCAAAATACTACAGATGCTGTCAAAGAAGATGTTAGAGGGA		
resistant sequence	(3613)	GTCATCTGGAAATCCAAATCAGGTGGCCCAAAATACTACAGATGCTGTCAAAGAAGATGTTAGAGGGA		
		3711		3780
consensus-sensitive	(3369)	AGGATTATACYTGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTAGATGTGAGGAGGGTGGT		
resistant sequence	(3683)	AGGATTATACCTGAATCATAAGGAACATCTCAATCACATTGTGGTTGATTAGATGTGAGGAGGGTGGT		
		3781		3850
consensus-sensitive	(3439)	GGAAGAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAGTTCGGGCCACATCCTTATCTTGAAAATT		
resistant sequence	(3753)	GGAAGAATGGATGATGAGGAAGCAAGAAGATTGATGGAAGAGTTCGGGCCACATCCTTATCTTGAAAATT		
		3851		3920
consensus-sensitive	(3509)	TGGCTGTGAAAGCATATATGGTGYGAAAYGCCTGRTTGGGYAACCTTCTYCCAAATCTTACAGAGCT		
resistant sequence	(3823)	TGGCTGTGAAAGCATATATGGTGTGAAATGCCTGGTGGGCAACCTTCTCCAAATCTTACAGAGCT		

## Φir. 1G

		3921		3990
consensus-sensitive	(3579)	TTWTCTTTTGTGATTGTGGGGAAYGGAGWRCCTTCCATGCMTGGGAACTTGGWTTDTCTRAAMGTCTC		
resistant sequence	(3893)	TTTCTTTTGTGATTGTGGGGAATGGAGAACCTTCCATGCTGGGAACTTGGATCATCTAAAAGTCTC		
		3991		4060
consensus-sensitive	(3649)	CGRCTTTTCGCTTTTGGCRAAATTGGAGTAYATTGRAGAAGATAGCWCATCAGCTMTTTCAGKTTTAGGC		
resistant sequence	(3963)	CGACTTTTCGCTTTTGGCAAAATTGGAGTACATTGAAGAAGATAGCTCATCAGCTAATTTACAGGTGTAGGC		
		4061		4130
consensus-sensitive	(3719)	CTGGACCRGAAAGTGCAGGACTATCATTATACTTCCCTCCCTTGAACKCCTTGAGTTGAAGCRTTTGYR		
resistant sequence	(4033)	CTGGACCRGAAAGTGCAGGACTATCATTATACTTCCCTCCCTTGAACCCCTTGAGTTGAAGCCTTTGTG		
		4131		4200
consensus-sensitive	(3789)	Y AAGTTAAAAGGATGGAGGAGARGGGAAGGGTTAGGAGATGATCACCAGCCTTTTAAATGAAAGCAGCAGC		
resistant sequence	(4103)	T AAGTTAAAAGGATGGAGGAGARGGGAAGGGTTAGGAGATGATCACCAGCCTTTTAAATGAAAGCAGCAGC		
		▼Retrotransposon Insertion		4270
consensus-sensitive	(3859)	AAT-----AAGTCATTGAGAATAGAAAGATGCCATTGC		
resistant sequence	(4173)	AATACACAAGTCCAATTACAATTATGTCTTCTCAATTGAAGTCATTGAGAATAGAAAGATGCCATTGC		
		4271		4340
consensus-sensitive	(3893)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAACDGCATTTAGTTGTATTTAATGAAYGACTCCGGAT		
resistant sequence	(4243)	TGACATTTATGCCGCTGTGTCCCAAGACAGAAAAACDGCATTTAGTTGTATTTAATGAACGACTCCGGAT		
		4341		4410
consensus-sensitive	(3963)	AGTGCATCTAAGGAGATGAGAATTTCTATGCTCCATTACATTCATCATCTGATCCTGAAAACCCG		
resistant sequence	(4313)	AGTGCATCTAAGGAGATGAGAATTTCTATGCTCCATTACATTCATCATCTGATCCTGAAAACCCG		
		4411		4480
consensus-sensitive	(4033)	AGGACACTATTCCCATTCCCATGTTAGAGAGGTATACATAAACAATGTGGCTGGCTAAATTCGCTGC		
resistant sequence	(4383)	AGGACACTATTCCCATTCCCATGTTAGAGAGGTATACATAAACAATGTGGCTGGCTAAATTCGCTGC		

## Φir. 1H

	4481		4550
consensus-sensitive	(4103)	CTATGGAGGCTTTTAGGTGTCTCACTCATATGACAATAAAAAACGAC	AGGTAGAGAGTTTGGGAGAAGT
resistant sequence	(4453)	CTATGGAGGCTTTTAGGTGTCTCACTCATATGACAATAAAAAACGAC	AGGTAGAGAGTTTGGGAGAAGT
	4551		4620
consensus-sensitive	(4173)	TGGGAGGTGTTTCGGAGCTCTCATCTTCTTTGCGATCCTTGAATATCACAGGTGCTCCAACTTAAGA	
resistant sequence	(4523)	TGGGAGGTGTTTCGGAGCTCTCATCTTCTTTGCGATCCTTGAATATCACAGGTGCTCCAACTTAAGA	
	4621		4690
consensus-sensitive	(4243)	AGTGTTTCTGGAGGGCTGGAGCATCTCACTCTTTGGAGATKTTAGAAATATACGACACCCATAAGCTGA	
resistant sequence	(4593)	AGTGTTTCTGGAGGGCTGGAGCATCTCACTCTTTGGAGATCTTAGAAATATACGACACCCATAAGCTGA	
	4691		4760
consensus-sensitive	(4313)	GTCTWTCAGAAGACCCAGAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGMAAT	
resistant sequence	(4663)	GTCTATCAGAAGACCCAGAAGGTGTTGTGCCATGGAAATCCCTTCATCACTCCCTCAGCTACTTGCAAT	
	4761		4830
consensus-sensitive	(4383)	GATGAATCTCCCGCAGCTGGTCAACCTGCCTGATTGCGATGCAGTCTCTGGG	CCCTCCAACTTTTCA
resistant sequence	(4733)	GATGAATCTCCCGCAGCTGGTCAACCTGCCTGATTGCGATGCAGTCTCTGGG	CCCTCCAACTTTTCA
	4831		4900
consensus-sensitive	(4453)	ATGTGCATTGCAATAACTGTAATCAGTGCCAGATTGGATGCCCMGACTCACTTCYCTCAGGAAGCTTA	
resistant sequence	(4803)	ATGTGCATTGCAATAACTGTAATCAGTGCCAGATTGGATGCCCMGACTCACTTCCTCAGGAAGCTTA	
	4901		4970
consensus-sensitive	(4523)	TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGYCAAAATCCAACTGGGGTGGACTGGCCTAACAT	
resistant sequence	(4873)	TGGTTTCATTCTGTTCCGCACATCTGGAGAGAAGATGCCAAATCCAACTGGGGTGGACTGGCCTAACAT	
	4971		5037
consensus-sensitive	(4593)	TCAACACATCCCTCATTGATGTACCTCTAGCC	TCCTAAGTTTTTAGTGTGCCGTATGAATAG
resistant sequence	(4943)	TCAACACATCCCTCATTGATGTACCTCTAGCC	TCCTAAGTTTTTAGTGTGCCGTATGAATAG

Фиг. 11

	1		70
Res. sequence	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence2	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence3	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence4	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence5	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence6	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence7	(1)	-----	STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence8	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence9	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence10	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence11	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence12	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence13	(1)	-----	STINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence14	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence15	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence16	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence17	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence18	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence19	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence20	(1)	MDVVGTAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence21	(1)	MDVVGSAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL
Sen. sequence22	(1)	MDVVGSAALSAAQSLFAALQSSSELKEILSIFGYKS	LDLQRTVSTINAVFRDAETKQELTHEAQHWLEEL

Фиг. 2A







Fig. 2E

Fig. 2E

Fig. 2G





```

701
Res. sequence (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVDFRCEEGGGRMDDEEARRIMEELRPHPYLENLA
Sen. sequence1 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVDFRCEEGGGRMDDEEARRIMEELRPHPYLENLA
Sen. sequence2 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence3 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence4 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence5 (701) LEIQIRWPEITTTDAVKEDVREGLYLNH-----
Sen. sequence6 (701) LEIQIRWPEITTTDAVKEDVREGLYLNH-----
Sen. sequence8 (701) LEIQIRWPEITTTDAVKEDVREGLYLNH-----
Sen. sequence9 (701) LEIQIRWPEITTTDAVKEDVREGLYLNH-----
Sen. sequence10 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVDFRCEEGGGRMDDEEARRIMEELRPHPYLENLA
Sen. sequence11 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence12 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence14 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence15 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence16 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence17 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence18 (700) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence19 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence20 (701) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----
Sen. sequence21 (644) LEIQIRWPEITTTDAVKEDVREGLYLNH-----
Sen. sequence22 (669) LEIQIRWPEITTTDAVKEDVREGLYLNHKEHLNHIVVD-----

```

## Φir. 2K

```

771
Res. sequence (771) VKAYYGVRMPGWATLLPNLTFLSDCGELENLPCGLNLDHLKVLRLSHLAKLEYIEEDSSSANFRCPRG
Sen. sequence1 (771) VKAYYGAKMPGWATLLPNLTFLSDCGESECLPCMGNDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG
Sen. sequence10 (771) VKAYYGAKMPGWATLLPNLTFLSDCGESECLPCMGNDCLKVLRLSHLAKLEYIEEDSTSANFSFRPG

```

```

841
Res. sequence (841) PESAGLSLYFSPSLERLELKRLEKLGWRRREGGLGDDHQPFNESSS>Ret.Ins<NTQVQLQLCLPQLKSLRIERCPLLT
Sen. sequence1 (841) PESAGLSLYFSPSLERLELKRLEKLGWRRREGGLGDDHQPFNESSS>Ret.Ins<-----
Sen. sequence10 (841) PESAGLSLYFSPSLERLELKRLEKLGWRRREGGLGDDHQPFNESSS>Ret.Ins<-----

```

```

911
Res. sequence (911) FMPLCPKTEKHLVVFNERLRIVHAKRDENFYAPLHSSSSDPENPRNTIPIPMFREVYINNVAWLNSLPM

```

```

981
Res. sequence (981) EAFRCCLTHMTIKNDEVESLGEVGEVFRSCSSSLRSLNITGCSNLRSVSGGLEHLTALEMLEIYDTHKLSL

```

```

1051
Res. sequence (1051) SEDPEGVVPWKSLSHLSYQLMNLPLQLVNLPSMQFLAALRTLSIVHCTKLQSVFDWMPRLTSLRKLMV

```

```

1121
Res. sequence (1121) SFCSAHLERRCQNFTGVDWPNIQHIPSIDVTSSSLPKFLVLPYE

```

## Φir. 2L







```

281
Res. sequence (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLREQLCGKRYILLVLDDV
Sen. sequence1 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence2 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence3 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence4 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence5 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence6 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence7 (168) DPRVRTAFPLRCWNC*-----
Sen. sequence8 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence9 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence10 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence11 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence12 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence13 (168) DPRVRTAFPLRCWNC*-----
Sen. sequence14 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence15 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence16 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence17 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence18 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLRE-LCGKRYILLVLDDV
Sen. sequence19 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence20 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDHVQTQLREQLCGKRYILLVLDDV
Sen. sequence21 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLREQLCGKRYILLVLDDV
Sen. sequence22 (211) DPRVRTAFPLRCWNCVSDQDQKQLDVKEILGKILSTATGKNHEGSTMDQVQTQLREQLCGKRYILLVLDDV

```

## Fig. 3D

[illegible]

**Fig. 3E**

	351	420
Res. sequence	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence1	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence2	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence3	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence4	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence5	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence6	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence8	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence9	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence10	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence11	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence12	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence14	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence15	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence16	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence17	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence18	(350) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence19	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence20	(351) NDFVMIARDIVDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	
Sen. sequence21	(333) XX	
Sen. sequence22	(331) XXXXXXXXXXXXXDKARNPLAIRVVGSLGCGQDKSKWLSFHEICLANIRKSHNDIMPILNLSYHHLEPPIR	

## Fig. 3E





631 700

Res. sequence (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence1 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence2 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence3 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence4 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence5 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence6 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence8 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence9 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence10 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence11 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence12 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence14 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence15 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence16 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence17 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence18 (630) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence19 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence20 (631) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence21 (574) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH  
 Sen. sequence22 (599) PKHLSRLVKLQTLDIYGCNNVTYMPKGMGRKCLHTLSKFTIVGGEGSCSSWKQFDGLEDLKALNNLKGH

## Φir. 3J

701 770

Res. sequence (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVVDFRCEEGGGRMDDEEARRLMPEELRPHPYLENLA  
 Sen. sequence1 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVVDFRCEEGGGRMDDEEARRLMPEELRPHPYLENLA  
 Sen. sequence2 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence3 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence4 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence5 (701) LEIQIRWENTTDAVKEDVREGLYLNH-----  
 Sen. sequence6 (701) LEIQIRWENTTDAVKEDVREGLYLNH-----  
 Sen. sequence8 (701) LEIQIRWENTTDAVKEDVREGLYLNH-----  
 Sen. sequence9 (701) LEIQIRWENTTDAVKEDVREGLYLNH-----  
 Sen. sequence10 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVVDFRCEEGGGRMDDEEARRLMPEELRPHPYLENLA  
 Sen. sequence11 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence12 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence14 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence15 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence16 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence17 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence18 (700) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence19 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence20 (701) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----  
 Sen. sequence21 (644) LEIQIRWENTTDAVKEDVREGLYLNH-----  
 Sen. sequence22 (669) LEIQIRWENTTDAVKEDVREGLYLNHKEHLNHIIVD-----

## Φir. 3K

771 840

Res. sequence (771) VKAYYGKMPGWATLLPNLTFLSDCGELENLPCLGNDHLKVLRLSHLAKLEYI EEDSSSANF3FRPG  
 Sen. sequence1 (771) VKAYYGKMPGWATLLPNLTFLSDCGELENLPCMGNDCLKVLRLSHLAKLEYI EEDSTSANF3FRPG  
 Sen. sequence10 (771) VKAYYGKMPGWATLLPNLTFLSDCGELENLPCMGNDCLKVLRLSHLAKLEYI EEDSTSANF3FRPG

841 910

Res. sequence (841) PESAGLSLYFPSLERLELKRLLKLGWRRREGGLDDHQPNESSS>Ret. Ins<NTQVQLQLCLPQLKSLRIERCPLLT  
 Sen. sequence1 (841) PESAGLSLYFPSLELLELKRLLKLGWRRREGGLDDHQPNESSS>Ret. Ins<-----  
 Sen. sequence10 (841) PESAGLSLYFPSLELLELKRLLKLGWRRREGGLDDHQPNESSS>Ret. Ins<-----

911 980

Res. sequence (911) FMPLCPKTEKHLVVFNERLRIVHAKRDENFYAPLHSSSSDPENPRNTIPIPMFREVIYINNVAVLNSLPM

981 1050

Res. sequence (981) EAFRCILHMTIRNDEVESLGEVGEVFRSCSSSLRSNLITGCSNLSVSGGLEHLTALEMLEIYDTHKLSL

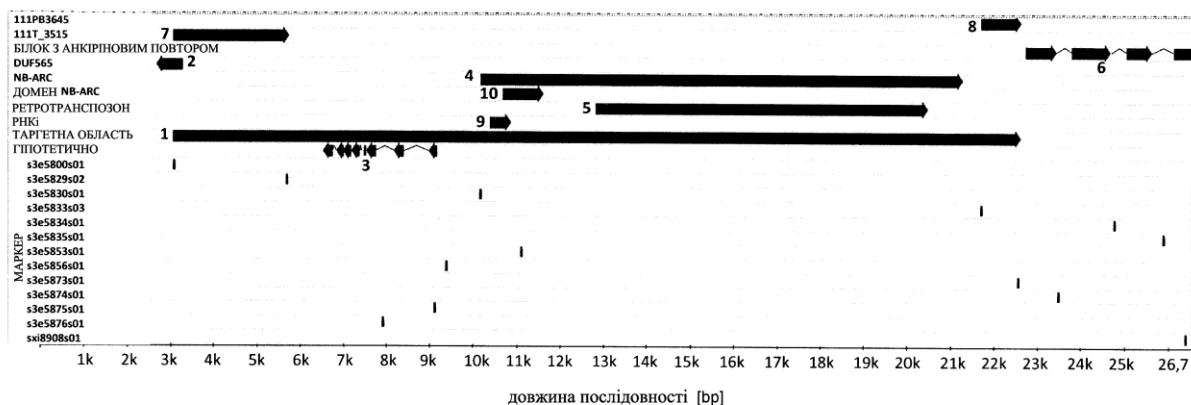
1051 1120

Res. sequence (1051) SEDPEGVVPWKSLSLSYLQIMNLPQLVNLPSMQFLAALRTLSIVHCTKLQSVDPWMPRLTSLRLKLMV

1121 1163

Res. sequence (1121) SPCSALHERRCQNPFGVDWPNIQHIP3IDVTSSLPKFLVLPYE

## Φir. 3L



**Fig. 4**

[illegible]

1051

Fig. 5



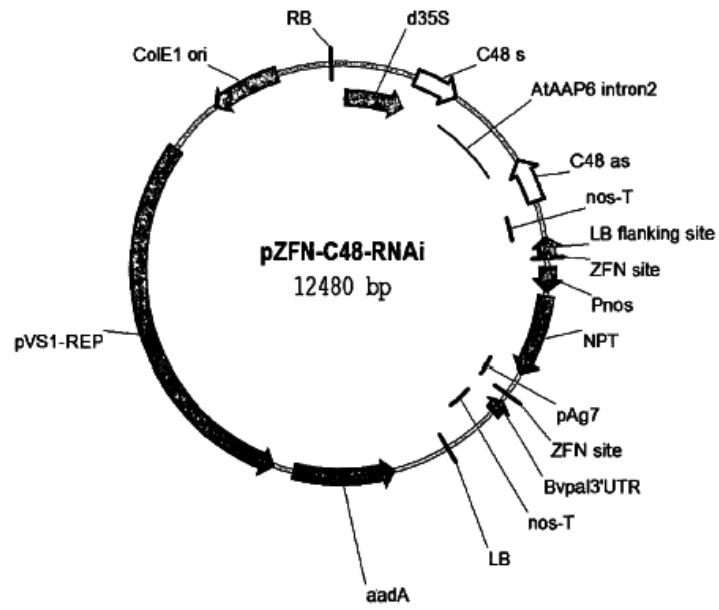


Fig. 6

Комп'ютерна верстка Л. Ціхановська

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

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