



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11)
CỤC SỞ HỮU TRÍ TUỆ



1-0031579

(51)⁷ A01N 63/02; C12N 15/32; C12N 15/31; (13) B
A61K 35/742; C07K 14/325

(21) 1-2018-00222

(22) 23/06/2016

(86) PCT/US2016/038947 23/06/2016

(87) WO2017/003811 05/01/2017

(30) 62/187,468 01/07/2015 US

(45) 25/04/2022 409

(43) 27/08/2018 365A

(73) SYNGENTA PARTICIPATIONS AG (CH)

Schwarzwaldallee 215, CH-4058 Basel, Switzerland

(72) BRAMLETT, Matthew Richard (US); SEGUIN, Katherine (US); ROSE, Mark Scott (US).

(74) Công ty TNHH Ban Ca (BANCA)

(54) PHƯƠNG PHÁP KIỂM SOÁT CÔN TRÙNG CÁNH VÂY KHÁNG CRY1AB VÀ
PHƯƠNG PHÁP NGĂN NGỪA SỰ PHÁT TRIỂN SỨC KHÁNG TRONG QUẦN
THỂ SÂU ĐỤC THÂN MÍA ĐỐI VỚI PROTEIN CRY1AB

(57) Sáng chế đề cập đến phương pháp kiểm soát côn trùng cánh vẩy kháng Cry1Ab và
phương pháp ngăn ngừa sự phát triển sức kháng trong quần thể sâu đục thân mía (*Diatraea
saccharalis*) đối với protein Cry1Ab.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp kiểm soát côn trùng cánh vảy kháng Cry1Ab và phương pháp ngăn ngừa sự phát triển sức kháng trong quần thể sâu đục thân mía (*Diatraea saccharalis*) đối với protein Cry1Ab.

Tình trạng kỹ thuật của sáng chế

Bacillus thuringiensis (Bt) là bào tử gam dương tạo thành vi khuẩn đất khác biệt bởi khả năng của nó tạo ra thể vùi tinh thể mà đặc biệt độc đối với một số bộ và loài gây hại cây trồng, bao gồm côn trùng, nhưng vô hại đối với cây trồng và các sinh vật khác không phải đích. Vì lý do này, chế phẩm chứa chủng *Bacillus thuringiensis* hoặc protein diệt côn trùng của chúng có thể được dùng làm chất diệt côn trùng chấp nhận được cho môi trường để kiểm soát các loài gây hại côn trùng nông nghiệp hoặc vector côn trùng của nhiều bệnh ở người hoặc động vật khác nhau.

Protein Crystal (Cry) từ *Bacillus thuringiensis* có hoạt tính diệt côn trùng tiềm năng đối với phần lớn côn trùng gây hại bộ cánh vảy, bộ hai cánh, và bộ cánh cứng. Các protein này còn thể hiện hoạt tính đối với loài gây hại thuộc bộ gây hại Hymenoptera, Homoptera, Phthiraptera, Mallophaga, và Acari, cũng như các bộ không xương sống khác như Nematelminthes, Platyhelminthes, và Sarcomastigophora (Feitelson, J. 1993. The *Bacillus Thuringiensis* family tree. In *Advanced Engineered Pesticides*. Marcel Dekker, Inc., New York, N.Y.). Các protein này ban đầu được phân loại là CryI đến CryVI chủ yếu dựa trên hoạt tính diệt côn trùng của chúng. Các lớp chính là đặc hiệu Bộ Cánh vảy (I), đặc hiệu Bộ Cánh vảy và Bộ Hai cánh (II), đặc hiệu Bộ Cánh cứng (III), đặc hiệu Bộ Hai Cánh (IV), và đặc hiệu giun tròn (V) và (VI). Các protein còn được phân loại thành các phân họ; các protein liên quan cao hơn trong mỗi họ được đặt các chữ cái phân chia như CryIA, CryIB, CryIC, v.v.. Các protein liên quan gần hơn nữa trong mỗi nhóm được đặt tên như CryIC(a), CryIC(b), v.v.. Các thuật ngữ “độc tố

Cry” và “nội độc tố denta” được sử dụng thay thế với thuật ngữ “protein Cry.” Danh pháp hiện nay cho protein và gen Cry dựa trên sự tương đồng trình tự axit amin hơn là tính đặc hiệu côn trùng đích (Crickmore và cộng sự 1998) *Microbiol. Mol. Biol. Rev.* 62:807-813). Theo phân loại được chấp nhận hơn này, mỗi độc tố được đặt tên duy nhất kết hợp với hàng chính (chữ số Ả rập), hàng thứ hai (ký tự viết hoa), hàng thứ ba (ký tự viết thường), và hàng thứ tư (chữ số Ả rập khác). Theo phân loại hiện nay, các chữ số La Mã được đổi cho các chữ số Ả rập ở hàng thứ nhất. Ví dụ, “CryIA(a)” theo danh pháp cũ nay là “Cry1Aa” theo danh pháp hiện nay. Theo Ibrahim và cộng sự (2010, *Bioeng. Bugs*, 1:31-50), độc tố Cry có thể còn được chia thành sáu lớp chính theo tính đặc hiệu côn trùng chủ của chúng và bao gồm: Nhóm 1—bộ cánh vảy ví dụ, Cry1, Cry9 và Cry15); nhóm 2—bộ cánh vảy và bộ hai cánh (ví dụ, Cry2); nhóm 3—bộ cánh cứng (Cry3, Cry7 và Cry8); nhóm 4— bộ hai cánh (Cry4, Cry10, Cry11, Cry16, Cry17, Cry19 và Cry20); nhóm 5— bộ cánh vảy và bộ cánh cứng (Cry1I); và nhóm 6—giun tròn (Cry6). Độc tố Cry1I, Cry2, Cry3, Cry10 và Cry11 (73–82 kDa) là duy nhất vì chúng xuất hiện để cắt đứt tự nhiên các protein Cry1 và Cry4 lớn hơn (130–140 kDa).

Các protein Cry là các phân tử protein hình cầu mà tích tụ lại thành tiền độc tố ở dạng tinh thể trong giai đoạn hình thành bào tử của *Bt*. Sau khi loài gây hại ăn vào bụng, tinh thể thường hòa tan để giải phóng tiền độc tố, mà có thể nằm trong khoảng kích thước, ví dụ, từ 130-140 kDa đối với nhiều trong số các protein Cry hoạt tính bộ cánh vảy, như Cry1 và Cry9, và 60-80 kDa đối với protein Cry3 hoạt tính bộ côn trùng cánh cứng và protein Cry2 hoạt tính bộ cánh vảy/bộ hai cánh. Sau khi tinh thể được hòa tan bởi côn trùng nhạy cảm giải phóng tiền độc tố được xử lý bởi proteaza trong ruột côn trùng, ví dụ trypsin và chymotrypsin, để sản xuất ra độc tố protein Cry lõi kháng proteaza. Quá trình xử lý phân giải protein này liên quan đến việc loại bỏ axit amin từ các vùng khác nhau của nhiều tiền độc tố Cry khác nhau. Ví dụ, tiền độc tố Cry mà là 130-140 kDa thường hoạt hóa qua sự loại bỏ phân giải protein của peptid đầu N có 25-30 axit amin và xấp xỉ một nửa protein còn lại từ đầu C tạo ra xấp xỉ 60-70 kDa độc tố Cry chín muồi. Tiền độc tố mà là 60-80 kDa, ví dụ Cry2 và Cry3, cũng được xử lý nhưng không cùng phạm vi như tiền độc tố lớn hơn. Tiền độc tố nhỏ hơn thường có bằng hoặc nhiều axit amin bị loại bỏ khỏi đầu N hơn so với tiền độc tố lớn hơn nhưng ít axit amin bị loại bỏ khỏi đầu C. Ví dụ, việc hoạt hóa phân giải protein của các thành viên trong họ Cry2 thường liên quan đến việc loại bỏ khoảng 40-50 axit amin đầu N. Nhiều loại trong

số protein Cry khá độc đối với côn trùng đích cụ thể, nhưng nhiều loại có phổ hoạt tính hẹp.

Các protein Cry thường có năm vùng trình tự bảo toàn, và ba vùng cấu trúc bảo toàn (xem, ví dụ, de Maagd và cộng sự (2001) Trends Genetics 17:193-199). Vùng cấu trúc bảo toàn thứ nhất, gọi là Vùng I, thường gồm có bảy cấu trúc xoắn anpha và liên quan đến chèn màng và tạo lỗ rỗng. Vùng II thường gồm có ba nếp gấp beta được bố trí trong cấu hình chìa khóa Greek, và vùng III thường gồm có hai nếp gấp beta đối song trong cấu tạo 'cuộn lỏng' (de Maagd và cộng sự, 2001, ở trên). Các vùng II và III liên quan đến nhận diện thụ thể và liên kết, và do đó được coi là các phần xác định đặc trưng độc tố.

Nhiều cây trồng có giá trị thương mại, bao gồm các cây nông nghiệp thông thường, dễ bị tấn công bởi các loài gây hại cây trồng bao gồm các loài gây hại côn trùng và giun tròn, gây giảm đáng kể sản lượng và chất lượng cây trồng. Ví dụ, loài gây hại cây trồng là yếu tố chính trong sự thiệt hại về cây nông nghiệp quan trọng của thế giới. Khoảng 8 tỷ đô la thiệt hại mỗi năm ở Mỹ chỉ do sự tàn phá của loài gây hại động vật thân mềm bao gồm côn trùng. Loài gây hại côn trùng cũng là gánh nặng đối với người trồng rau và hoa quả, người sản xuất hoa cảnh, và người làm vườn gia đình.

Các loài gây hại côn trùng chủ yếu được kiểm soát nhờ áp dụng chuyên sâu các chất diệt loài gây hại hoá học, có hoạt tính thông qua việc ức chế sự tăng trưởng của côn trùng, ngăn ngừa sự lớn lên hoặc sinh sôi nảy nở của côn trùng, hoặc gây chết. Các tác nhân kiểm soát loài gây hại sinh học, như chủng *Bacillus thuringiensis* biểu hiện độc tố diệt loài gây hại như protein Cry, cũng được áp dụng lên cây trồng với hiệu quả tốt, tạo ra sự thay thế hoặc phần thưởng cho thuốc diệt loài gây hại hóa học. Các gen mã hóa đối với một số trong các protein Cry này đã được phân lập và biểu hiện của chúng ở vật chủ khác loại như cây chuyển gen thể hiện là tạo ra công cụ khác để kiểm soát các loài gây hại côn trùng quan trọng về kinh tế.

Do đó, có thể đạt được kiểm soát côn trùng tốt, nhưng một số chất hóa học nhất định đôi khi có thể còn ảnh hưởng đến côn trùng có ích không phải đích và một số chất sinh học nhất định có phổ hoạt tính rất hẹp. Ngoài ra, việc sử dụng tiếp một số phương pháp kiểm soát hóa học và sinh học nhất định làm tăng khả năng cho loài gây hại côn trùng phát triển khả năng kháng các phương pháp kiểm soát này. Điều này được giảm

phần nào nhờ các thực tiễn quản lý độ kháng khác nhau, nhưng vẫn cần phải phát triển các tác nhân kiểm soát loài gây hại mới và hiệu quả mà tạo ra lợi ích kinh tế đối với nông dân và chấp nhận được cho môi trường. Đặc biệt có nhu cầu là các tác nhân kiểm soát mà có thể nhắm đích phổ rộng hơn loài gây hại côn trùng quan trọng về kinh tế và kiểm soát hiệu quả các chủng côn trùng mà kháng hoặc có thể trở nên kháng các tác nhân kiểm soát côn trùng hiện có.

Bản chất kỹ thuật của sáng chế

Trên cơ sở xem xét nhu cầu này, mục đích của sáng chế là đề xuất tác nhân kiểm soát loài gây hại mới bằng cách tạo ra các phân lập *Bacillus thuringiensis* (*Bt*) mới cũng như gen và protein diệt loài gây hại mới mà có thể được sử dụng để kiểm soát nhiều loài gây hại cây trồng khác nhau.

Sáng chế đề xuất các chế phẩm và phương pháp tạo ra hoạt tính diệt loài gây hại lên vi khuẩn, cây, tế bào, mô và hạt của cây. Cụ thể là, gen khám bao gồm polynucleotit mới mà mã hóa protein Cry được phân lập từ *Bt* và trình tự cơ bản tương đồng với chúng, mà biểu hiện của chúng tạo ra protein với độc tố đối với loài gây hại côn trùng quan trọng về kinh tế, cụ thể là loài gây hại côn trùng mà tàn phá cây trồng, được đề xuất. Sáng chế còn đề xuất protein Cry mới tạo ra từ việc biểu hiện polynucleotit, và chế phẩm và chế phẩm bào chế chứa protein Cry, mà độc đối với côn trùng bằng cách ức chế khả năng của loài gây hại côn trùng để sống sót, sinh trưởng và tái sinh, hoặc hạn chế thiệt hại hoặc mất mát liên quan đến côn trùng cho cây trồng. Các protein Cry theo sáng chế bao gồm protein Cry tự nhiên và protein Cry biến thể hoặc đột biến mà có một hoặc nhiều sự thay thế, thêm hoặc loại bỏ axit amin. Các ví dụ của protein Cry đột biến bao gồm không giới hạn các loại được đột biến có phổ hoạt tính rộng hơn hoặc hoạt tính đặc hiệu cao hơn bản sao protein Cry tự nhiên của chúng, các loại được đột biến để đưa vào epitop để tạo ra kháng thể mà nhận diện một cách khác protein đột biến từ protein tự nhiên hoặc các loại được đột biến để điều biến biểu hiện trong sinh vật chuyển gen. Các protein Cry mới theo sáng chế có độc cao đối với loài gây hại côn trùng. Ví dụ, protein Cry theo sáng chế có thể được dùng để kiểm soát một hoặc nhiều loài gây hại côn trùng quan trọng về kinh tế như sâu xám (*Agrotis ipsilon*), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu xanh hại ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá velvetbean (*Anticarsia gemmatilis*), sâu đo hại đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam

(*Diatraea grandiosella*), sâu đục trái bắp (*Richia albicosta*), tuyến trùng ăn rễ cây thuốc lá (*Heliothis virescens*), sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc nâu (*Chilo suppressalis*), sâu đục thân mình hồng (*Sesamia calamistis*), sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*), và tương tự.

Sáng chế còn đề xuất polynucleotit tổng hợp mà mã hóa protein Cry theo sáng chế mà có một hoặc nhiều bộ ba mã hóa được tối ưu để biểu hiện trong sinh vật chuyển gen như vi khuẩn chuyển gen hoặc cây chuyển gen.

Sáng chế còn đề xuất tổ hợp biểu hiện và vectơ tái tổ hợp chứa polynucleotit mà mã hóa protein Cry theo sáng chế. Sáng chế còn đề xuất vi khuẩn, cây, tế bào cây, mô, và hạt được biến nạp bao gồm gen khảm, hoặc cat-xet biểu hiện hoặc vectơ tái tổ hợp mà hữu dụng trong biểu hiện protein Cry theo sáng chế ở vi khuẩn, cây, tế bào cây, mô, và hạt được biến nạp.

Sáng chế còn đề xuất chủng *Bacillus thuringiensis* (*Bt*) được phân lập mà tạo ra protein Cry theo sáng chế. Chủng *Bt* này có thể là phân lập xảy ra trong tự nhiên hoặc chủng *Bt* chuyển gen mà tạo ra một hoặc nhiều trong số các protein Cry theo sáng chế.

Sáng chế còn đề xuất phương pháp sử dụng polynucleotit, ví dụ trong cấu trúc ADN hoặc gen khảm hoặc cat-xet biểu hiện hoặc vectơ tái tổ hợp để biến nạp và biểu hiện trong sinh vật, bao gồm cây và vi sinh vật, như vi khuẩn. Trình tự nucleotit hoặc axit amin có thể là trình tự tự nhiên hoặc tổng hợp mà được thiết kế để biểu hiện trong sinh vật như cây hoặc vi khuẩn hoặc tạo ra độc tố Cry lai với hoạt tính diệt loài gây hại tăng cường. Sáng chế còn đề xuất phương pháp tạo ra protein Cry và phương pháp sử dụng trình tự polynucleotit và protein Cry, ví dụ trong vi sinh vật để kiểm soát côn trùng hoặc trong cây chuyển gen để tạo ra bảo vệ khỏi mối hại côn trùng.

Các khía cạnh khác của sáng chế bao gồm chế phẩm và chế phẩm bào chế diệt côn trùng bao gồm protein Cry hoặc chủng *Bacillus thuringiensis* theo sáng chế, và phương pháp sử dụng chế phẩm hoặc chế phẩm bào chế để kiểm soát quần thể côn trùng, ví dụ bằng cách áp dụng chế phẩm hoặc chế phẩm bào chế lên vùng bị tàn phá bởi côn trùng, hoặc để xử lý phòng ngừa vùng hoặc cây dễ bị côn trùng tàn phá để tạo ra sự bảo vệ đối với loài gây hại côn trùng. Tùy ý, chế phẩm hoặc chế phẩm bào chế theo sáng chế có thể, ngoài protein Cry hoặc chủng *Bt* theo sáng chế, bao gồm các tác nhân diệt loài gây

hại khác như chất diệt loài gây hại hóa học nhằm tăng hoặc tăng cường khả năng kiểm soát côn trùng của chế phẩm hoặc chế phẩm bào chế.

Chế phẩm và phương pháp theo sáng chế hữu dụng để kiểm soát loài gây hại côn trùng mà tấn công cây, đặc biệt là cây trồng. Chế phẩm theo sáng chế còn hữu dụng để tạo ra protein Cry biến đổi hoặc cải tiến mà có hoạt tính diệt loài gây hại, hoặc để phát hiện sự có mặt của protein Cry hoặc axit nucleic trong sản phẩm bán trên thị trường hoặc sinh vật chuyển gen.

Các dấu hiệu, khía cạnh, và lợi ích này và khác của sáng chế sẽ được hiểu rõ hơn khi viện dẫn đến phần mô tả chi tiết và các yêu cầu bảo hộ sau đây.

Mô tả vắn tắt các trình tự trong danh mục trình tự

SEQ ID NO:1 là trình tự nucleotit mã hóa protein BT-0001.

SEQ ID NO:2 là trình tự nucleotit mã hóa protein BT-0003.

SEQ ID NO:3 là trình tự nucleotit mã hóa protein BT-0020.

SEQ ID NO:4 là trình tự nucleotit mã hóa protein BT-0022.

SEQ ID NO:5 là trình tự nucleotit mã hóa protein BT-0027.

SEQ ID NO:6 là trình tự nucleotit mã hóa protein BT-0029.

SEQ ID NO:7 là trình tự nucleotit mã hóa protein BT-0030.

SEQ ID NO:8 là trình tự nucleotit mã hóa protein BT-0031.

SEQ ID NO:9 là trình tự nucleotit mã hóa protein BT-0201 .

SEQ ID NO:10 là trình tự nucleotit mã hóa protein BT-0202.

SEQ ID NO:11 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0001.

SEQ ID NO:12 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0003.

SEQ ID NO:13 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0020.

SEQ ID NO:14 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0022.

SEQ ID NO:15 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0027.

SEQ ID NO:16 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0029.

SEQ ID NO:17 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0030.

SEQ ID NO:18 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0031.

SEQ ID NO:19 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0201 .

SEQ ID NO:20 là trình tự tối ưu bộ ba mã hóa mã hóa protein BT-0202.

SEQ ID NO:21 là trình tự nucleotit mã hóa protein BT-0001 đột biến.

SEQ ID NO:22 là trình tự nucleotit mã hóa protein BT-0003 đột biến.

SEQ ID NO:23 là trình tự nucleotit mã hóa protein BT-0020 đột biến.

SEQ ID NO:24 là trình tự nucleotit mã hóa protein BT-0022 đột biến.

SEQ ID NO:25 là trình tự nucleotit mã hóa protein BT-0027 đột biến.

SEQ ID NO:26 là trình tự nucleotit mã hóa protein BT-0029 đột biến.

SEQ ID NO:27 là trình tự nucleotit mã hóa protein BT-0030 đột biến.

SEQ ID NO:28 là trình tự nucleotit mã hóa protein BT-0031 đột biến.

SEQ ID NO:29 là trình tự nucleotit mã hóa protein BT-0201 đột biến.

SEQ ID NO:30 là trình tự nucleotit mã hóa protein BT-0202 đột biến.

SEQ ID NO:31 là trình tự axit amin của protein BT-0001.

SEQ ID NO:32 là trình tự axit amin của protein BT-0003.

SEQ ID NO:33 là trình tự axit amin của protein BT-0020.

SEQ ID NO:34 là trình tự axit amin của protein BT-0022.

SEQ ID NO:35 là trình tự axit amin của protein BT-0027.

SEQ ID NO:36 là trình tự axit amin của protein BT-0029.

SEQ ID NO:37 là trình tự axit amin của protein BT-0030.

SEQ ID NO:38 là trình tự axit amin của protein BT-0031.

SEQ ID NO:39 là trình tự axit amin của protein BT-0201 .

SEQ ID NO:40 là trình tự axit amin của protein BT-0202.

SEQ ID NO:41 là trình tự axit amin của protein BT-0001 đột biến.

SEQ ID NO:42 là trình tự axit amin của protein BT-0003 đột biến.

SEQ ID NO:43 là trình tự axit amin của protein BT-0020 đột biến.

SEQ ID NO:44 là trình tự axit amin của protein BT-0022 đột biến

SEQ ID NO:45 là trình tự axit amin của protein BT-0027 đột biến.

SEQ ID NO:46 là trình tự axit amin của protein BT-0029 đột biến.

SEQ ID NO:47 là trình tự axit amin của protein BT-0030 đột biến.

SEQ ID NO:48 là trình tự axit amin của protein BT-0031 đột biến.

SEQ ID NO:49 là trình tự axit amin của protein BT-0201 đột biến

SEQ ID NO:50 là trình tự axit amin của protein BT-0202 đột biến.

Mô tả chi tiết sáng chế

Mô tả không nhằm làm danh mục cụ thể của tất cả các biến đổi khác trong đó sáng chế có thể được thực hiện, hoặc tất cả các dấu hiệu có thể được thêm vào sáng chế. Ví dụ, các dấu hiệu được minh họa theo một phương án có thể được kết hợp với các phương án khác, và các dấu hiệu được minh họa theo phương án cụ thể có thể được xóa bỏ khỏi phương án này. Do đó, sáng chế bao hàm theo một số phương án theo sáng chế, dấu hiệu bất kỳ hoặc kết hợp các dấu hiệu được đề cập ở đây có thể được loại bỏ hoặc bỏ qua. Ngoài ra, nhiều biến thể và bổ sung cho các phương án khác nhau được khuyến khích ở đây sẽ được thực hiện bởi chuyên gia trong lĩnh vực theo sáng chế, mà không vượt ra khỏi phạm vi sáng chế. Vì vậy, phần mô tả sau đây nhằm minh họa một số phương án cụ thể của sáng chế, và không xác định chi tiết tất cả các sắp xếp, dạng kết hợp và biến thể của chúng.

Trừ khi có chỉ định khác, toàn bộ các thuật ngữ kỹ thuật và khoa học được sử dụng ở đây có cùng nghĩa như cách hiểu thông thường của chuyên gia trong lĩnh vực đối mà sáng chế thuộc về. Thuật ngữ được sử dụng trong phần mô tả của sáng chế ở đây là nhằm mục đích mô tả chỉ các phương án cụ thể và không nhằm giới hạn sáng chế.

Định nghĩa

Như được sử dụng ở đây và ở yêu cầu bảo hộ kèm theo, các dạng mạo từ số ít bao gồm viện dẫn số nhiều trừ khi ngữ cảnh quy định rõ ràng khác. Do đó, ví dụ, viện dẫn đến "cây trồng" là viện dẫn đến một hoặc nhiều cây trồng và bao gồm các dạng tương đương của chúng đã biết đối với chuyên gia trong lĩnh vực, và v.v. Như được sử dụng ở

đây, từ "hoặc" nghĩa là một thành phần bất kỳ của danh sách cụ thể và còn bao gồm kết hợp bất kỳ của các thành phần của danh sách đó (tức là, còn bao gồm "và").

Thuật ngữ "khoảng" được sử dụng ở đây nghĩa là xấp xỉ, gần, xung quanh, hoặc nằm trong vùng đó. Khi thuật ngữ "khoảng" được dùng kết hợp với khoảng số, nó thay đổi khoảng này bằng cách mở rộng giới hạn trên và dưới giá trị số được cho. Thông thường, thuật ngữ "khoảng" được sử dụng ở đây để thay đổi giá trị số trên và dưới giá trị được cho với phương sai bằng 20 phần trăm, tốt hơn là 10 phần trăm lên hoặc xuống (cao hơn hoặc thấp hơn). Đối với thuật ngữ nhiệt độ "khoảng" nghĩa là $\pm 1^{\circ}\text{C}$, tốt hơn là $\pm 0,5^{\circ}\text{C}$. Khi thuật ngữ "khoảng" được dùng trong ngữ cảnh của sáng chế (ví dụ, kết hợp với các giá trị nhiệt độ hoặc trọng lượng phân tử) giá trị chính xác (tức là, không có "khoảng") được ưu tiên.

Thuật ngữ "hoạt tính" của protein Cry độc theo sáng chế nghĩa là các chức năng protein độc làm tác nhân kiểm soát côn trùng hoạt tính qua đường miệng, có hiệu quả độc, hoặc có khả năng làm gián đoạn hoặc ngăn cản sự phá hoại của côn trùng, mà có thể hoặc không thể dẫn đến gây chết côn trùng. Khi protein độc theo sáng chế được phân phối lên côn trùng, kết quả thường là gây chết côn trùng, hoặc côn trùng không gây hại đối với nguồn mà tạo ra protein độc cho côn trùng.

Như được sử dụng ở đây, thuật ngữ "khuếch đại" nghĩa là việc xây dựng nhiều bản sao của phân tử axit nucleic hoặc nhiều bản sao bổ sung cho phân tử axit nucleic bằng cách sử dụng ít nhất một trong số các phân tử axit nucleic làm mẫu. Hệ thống khuếch đại bao gồm hệ phản ứng chuỗi polymeraza (PCR), hệ phản ứng chuỗi ligaza (LCR), khuếch đại trên cơ sở trình tự axit nucleic (NASBA, Cengage, Mississauga, Ontario), hệ lặp Q-Beta Replicaza, hệ khuếch đại trên cơ sở phiên mã (TAS), và khuếch đại đôi chỗ sợi (SDA). Xem, ví dụ, Diagnostic Molecular Microbiology: Principles và Applications, PERSING và cộng sự, Ed., American Society for Microbiology, Washington, D.C. (1993). Sản phẩm của khuếch đại được gọi là "amplicon."

Thuật ngữ "cấu trúc khảm" hoặc "gen khảm" hoặc "polynucleotit khảm" hoặc "axit nucleic khảm" (hoặc thuật ngữ tương tự) như được sử dụng ở đây đề cập đến cấu trúc hoặc phân tử bao gồm hai hoặc nhiều polynucleotit có nguồn gốc khác nhau được ghép thành phân tử axit nucleic đơn. Thuật ngữ "cấu trúc khảm", "gen khảm", "polynucleotit khảm" hoặc "axit nucleic khảm" đề cập đến bất kỳ cấu trúc hoặc phân tử mà bao gồm,

không giới hạn, (1) polynucleotit (ví dụ, ADN) , chứa polynucleotit điều chỉnh và mã hóa mà không thấy cùng nhau trong tự nhiên(tức là, ít nhất một trong số các polynucleotit trong cấu trúc là khác loại với ít nhất một trong số các polynucleotit khác của nó), hoặc (2) các phần mã hóa polynucleotit của protein không tiếp giáp trong tự nhiên, hoặc (3) các phần của vùng khởi động mà không tiếp giáp trong tự nhiên. Hơn nữa, cấu trúc khảm, gen khảm, polynucleotit khảm hoặc axit nucleic khảm có thể chứa polynucleotit điều chỉnh và polynucleotit mã hóa mà có nguồn gốc từ các nguồn khác nhau, hoặc chứa polynucleotit điều chỉnh và polynucleotit mã hóa có nguồn gốc từ nguồn giống nhau, nhưng được sắp xếp theo cách khác với cách tìm thấy trong tự nhiên. Theo một số phương án của sáng chế, cấu trúc khảm, gen khảm, polynucleotit khảm hoặc axit nucleic khảm bao gồm tổ hợp biểu hiện chứa polynucleotit theo sáng chế dưới sự kiểm soát của polynucleotit điều chỉnh, cụ thể là dưới sự kiểm soát của polynucleotit điều chỉnh chức năng trong cây trồng hoặc vi khuẩn.

"Trình tự mã hóa" là trình tự axit nucleic mà được phiên mã thành ARN như mARN, rARN, tARN, snARN, ARN có nghĩa hoặc ARN đối nghĩa. Tốt hơn là sau đó ARN được dịch mã vào sinh vật để tạo ra protein.

Như được sử dụng ở đây, trình tự "tối ưu bộ ba mã" nghĩa là trình tự nucleotit của polynucleotit tái tổ hợp, chuyển gen, hoặc tổng hợp trong đó bộ ba mã được chọn để phản ánh độ lệch bộ ba mã cụ thể mà tế bào chủ hoặc sinh vật có thể có. Điều này thường được thực hiện theo cách này để bảo quản trình tự axit amin của polypeptit được mã hóa bởi trình tự nucleotit tối ưu bộ ba mã. Theo một số phương án nhất định, trình tự ADN của cấu trúc ADN tái tổ hợp bao gồm trình tự mà đã là bộ ba mã được tối ưu cho tế bào (ví dụ, tế bào động vật, thực vật, hoặc nấm) trong đó cấu trúc là để được biểu hiện. Ví dụ, cấu trúc được biểu hiện trong tế bào của cây có thể có tất cả hoặc các phần của trình tự của nó (ví dụ, yếu tố ức chế gen thứ nhất hoặc yếu tố biểu hiện gen) bộ ba mã được tối ưu để biểu hiện ở cây trồng. Xem, ví dụ, Sáng chế Mỹ số 6,121,014, được kết hợp ở đây bằng cách viện dẫn.

"Kiểm soát" côn trùng nghĩa là ức chế, thông qua tác dụng độc tố, khả năng sống sót, phát triển, ăn, hoặc sinh sôi nảy nở của loài gây hại côn trùng hoặc hạn chế thiệt hại hoặc mất mát liên quan đến côn trùng ở cây trồng hoặc bảo vệ tiềm năng sản lượng của cây trồng khi phát triển với sự có mặt của loài gây hại côn trùng. "Kiểm soát" côn trùng

có thể có hoặc có thể không có nghĩa là giết côn trùng, mặc dù tốt hơn nên có nghĩa là giết côn trùng.

Thuật ngữ "bao gồm" hoặc "gồm," khi được sử dụng trong bản mô tả này, chỉ rõ sự có mặt của các dấu hiệu đã định, số nguyên, bước, thao tác, yếu tố, hoặc thành phần, nhưng không loại trừ sự có mặt hoặc bổ sung của một hoặc nhiều dấu hiệu, số nguyên, bước, thao tác, yếu tố, thành phần, hoặc nhóm khác của chúng.

Như được sử dụng ở đây, câu chuyển tiếp "chủ yếu gồm có" (và các biến thể ngữ pháp) nghĩa là phạm vi của yêu cầu bảo hộ được hiểu là bao hàm các vật liệu hoặc các bước được nêu rõ hoặc đề cập trong yêu cầu bảo hộ" và những thứ mà không thay thế cơ bản (các) đặc trưng cơ bản và mới" của sáng chế xin cấp bằng. Do đó, thuật ngữ "chủ yếu gồm có" khi được sử dụng trong yêu cầu bảo hộ của sáng chế không dự tính được hiểu là tương đương với "chứa."

Trong bối cảnh của sáng chế, "sự tương ứng với" hoặc "tương ứng với" nghĩa là khi trình tự axit amin của protein Cry biến thể hoặc đồng đẳng được sắp thẳng hàng với nhau, axit amin mà "tương ứng với" các vị trí được đánh số nhất định trong protein biến thể hoặc đồng đẳng là các loại mà thẳng hàng với các vị trí này trong protein đối chiếu nhưng không nhất thiết là ở các vị trí số chính xác này so với trình tự axit amin đối chiếu cụ thể theo sáng chế. Ví dụ, nếu SEQ ID NO:31 là trình tự đối chiếu và được sắp thẳng hàng với SEQ ID NO:32, Ser629 nêu ở SEQ ID NO:32 "tương ứng với" Ser630 nêu ở SEQ ID NO:31.

Như được sử dụng ở đây, thuật ngữ "Cry protein" nghĩa là protein diệt côn trùng của loại nội độc tố delta tinh thể *Bacillus thuringiensis*. Thuật ngữ "Cry protein" có thể đề cập đến dạng tiền độc tố hoặc bất kỳ đoạn hoạt tính diệt côn trùng hoặc độc tố của chúng.

"Phân phối" chế phẩm hoặc protein độc nghĩa là chế phẩm hoặc protein độc được tiếp xúc với côn trùng, mà tạo điều kiện cho việc ăn vào bụng qua đường miệng chế phẩm hoặc protein độc, tạo ra hiệu quả độc và kiểm soát côn trùng. Chế phẩm hoặc protein độc có thể được phân phối theo nhiều cách đã biết, bao gồm nhưng không giới hạn ở, biểu hiện cây chuyển gen, (các) chế phẩm protein phối chế, (các) chế phẩm protein phun được, ma trận bả, hoặc bất kỳ hệ phân phối protein đã biết khác trong lĩnh vực.

Thuật ngữ "vùng" đề cập đến bộ của các axit amin được bảo toàn ở các vị trí cụ thể dọc theo sự sắp hàng của các trình tự của các protein liên quan đến tiến hóa. Trong khi các axit amin ở các vị trí khác có thể thay đổi giữa các tương đồng, các axit amin được bảo toàn cao ở các vị trí cụ thể cho thấy các axit amin mà có vẻ thiết yếu trong cấu trúc, sự ổn định hoặc chức năng của protein. Được xác định bởi mức độ bảo toàn cao của chúng trong trình tự được sắp thẳng hàng của họ chất tương đồng protein, chúng có thể được dùng làm vật xác định để xác định xem polypeptit yêu cầu bất kỳ thuộc nhóm polypeptit được xác định trước đó hay không.

"Lượng có hiệu quả kiểm soát côn trùng" nghĩa là nồng độ của protein độc mà ức chế, thông qua tác dụng độc, khả năng sống sót, phát triển, ăn hoặc tái sinh của côn trùng, hoặc hạn chế thiệt hại hoặc mất mát liên quan đến côn trùng ở cây trồng hoặc bảo vệ tiềm năng năng suất của cây khi phát triển với sự có mặt của loài gây hại côn trùng. "Lượng có hiệu quả kiểm soát côn trùng" có thể có hoặc có thể không có nghĩa là giết côn trùng, mặc dù tốt hơn nên có nghĩa là giết côn trùng.

"Cat-xet biểu hiện" như được sử dụng ở đây nghĩa là phân tử axit nucleic có khả năng định hướng biểu hiện của ít nhất một polynucleotit có lợi, như polynucleotit mã hóa protein Cry theo sáng chế, trong tế bào chủ thích hợp, bao gồm vùng khởi động liên kết hoạt động với polynucleotit có lợi mà liên kết hoạt động với tín hiệu kết thúc. "Tổ hợp biểu hiện" thường còn chứa polynucleotit bổ sung cần để phiên mã đúng polynucleotit được mong muốn. Tổ hợp biểu hiện có thể còn chứa polynucleotit khác không cần trong biểu hiện trực tiếp của polynucleotit được mong muốn nhưng có mặt do vị trí giới hạn thích hợp để loại bỏ tổ hợp khởi vectơ biểu hiện. Cat-xet biểu hiện bao gồm (các) polynucleotit được mong muốn có thể là khảm, có nghĩa là ít nhất một trong số các thành phần của nó là khác loại với ít nhất một trong số các thành phần khác của nó. Cat-xet biểu hiện có thể còn là tổ hợp mà xuất hiện trong tự nhiên nhưng có thể thu được ở dạng tái tổ hợp hữu dụng để biểu hiện khác loài. Tuy nhiên, thông thường, tổ hợp biểu hiện là khác loại với vật chủ, tức là polynucleotit được mong muốn trong tổ hợp biểu hiện không xảy ra tự nhiên trong tế bào chủ và phải được đưa vào tế bào chủ hoặc tổ tiên của tế bào chủ bằng quá trình biến nạp hoặc quá trình gây giống. Biểu hiện của (các) polynucleotit được mong muốn trong tổ hợp biểu hiện thường là dưới sự kiểm soát của vùng khởi động. Trong trường hợp sinh vật nhiều ngăn, như cây trồng, vùng khởi động có thể còn là đặc biệt hoặc ưu tiên cho mô, hoặc cơ quan, hoặc giai đoạn phát

triển cụ thể. Cat-xet biểu hiện, hoặc đoạn của chúng, còn có thể được gọi là "polynucleotit được chèn" hoặc "polynucleotit chèn vào" khi được biến nạp vào cây trồng.

"Gen" được xác định ở đây là đơn vị di truyền bao gồm một hoặc nhiều polynucleotit mà chiếm vị trí cụ thể trên nhiễm sắc thể hoặc plasmid và mà chứa chỉ thị gen cho đặc trưng cụ thể hoặc đặc điểm trong sinh vật.

"Proteaza của ruột" là proteaza được tìm thấy tự nhiên trong đường tiêu hóa của côn trùng. Proteaza này thường tham gia vào việc tiêu hóa protein được ăn vào. Các ví dụ của proteaza ruột bao gồm trypsin, mà thường phân cắt peptit ở bên đầu C của gốc lysin (K) hoặc arginin (R), và chymotrypsin, mà thường phân cắt peptit ở bên đầu C của phenylalanin (F), tryptophan (W) hoặc tyrosin (Y).

Thuật ngữ "khác loại" khi được sử dụng viện dẫn đến gen hoặc polynucleotit hoặc polypeptit đề cập đến gen hoặc polynucleotit hoặc polypeptit mà là hoặc chứa một phần của chúng không có trong môi trường tự nhiên của nó (tức là, đã được thay đổi bởi bàn tay con người). Ví dụ, gen khác loại có thể bao gồm polynucleotit từ một loài được đưa vào loài khác. Gen khác loại còn có thể bao gồm polynucleotit tự nhiên với sinh vật mà được thay đổi theo một số cách (ví dụ, đột biến, thêm nhiều bản sao, liên kết với vùng khởi động không tự nhiên hoặc polynucleotit đoạn tăng cường, v.v.). Gen khác loại còn có thể chứa polynucleotit của gen thực vật mà bao gồm các dạng cADN của gen thực vật; cADN có thể được biểu hiện hoặc là theo hướng có nghĩa (để tạo ra mARN) hoặc theo hướng đối nghĩa (để tạo ra phiên mã ARN đối nghĩa mà bổ sung cho phiên mã mARN). Theo một khía cạnh của sáng chế, gen khác loại phân biệt với gen thực vật nội sinh ở chỗ polynucleotit của gen khác loại thường nối với polynucleotit bao gồm yếu tố điều chỉnh như vùng khởi động mà không thấy kết hợp tự nhiên với gen cho protein được mã hóa bởi gen khác loại hoặc với polynucleotit của gen thực vật trong nhiễm sắc thể, hoặc kết hợp với các phần của nhiễm sắc thể không thấy trong tự nhiên (ví dụ, gen được biểu hiện ở các vị trí mà gen không được biểu hiện tự nhiên). Hơn nữa, polynucleotit "khác loại" đề cập đến polynucleotit không kết hợp tự nhiên với tế bào chủ mà nó được đưa vào, bao gồm nhiều bản sao không xảy ra tự nhiên của polynucleotit có trong tự nhiên.

"Tái tổ hợp cùng loại" là sự trao đổi ("lai chéo") của các đoạn ADN giữa hai phân tử ADN hoặc nhiễm sắc tử của các nhiễm sắc thể bắt cặp trong vùng polynucleotit tương đồng. "Sự kiện tái tổ hợp" ở đây được hiểu có nghĩa là lai chéo meiotic.

Trình tự axit nucleic "đăng mã hóa" với trình tự axit nucleic quy chiếu khi trình tự axit nucleic mã hóa polypeptit có trình tự axit amin giống như polypeptit được mã hóa bởi trình tự axit nucleic quy chiếu. Ví dụ, SEQ ID NO:11 là đăng mã hóa với SEQ ID NO: 1 vì chúng đều mã hóa trình tự axit amin được thể hiện bởi SEQ ID NO:31.

Thuật ngữ phân tử axit nucleic, polynucleotit hoặc protein "được phân lập" là phân tử axit nucleic, polynucleotit hoặc protein mà không còn tồn tại trong môi trường tự nhiên của nó. Phân tử axit nucleic, polynucleotit hoặc protein được phân lập theo sáng chế có thể tồn tại dưới dạng tinh chế hoặc có thể tồn tại trong vật chủ tái tổ hợp như trong vi khuẩn chuyển gen hoặc cây chuyển gen.

"Phân tử axit nucleic" là ADN hoặc ARN mạch đơn hoặc kép mà có thể được phân lập từ nguồn bất kỳ. Trong ngữ cảnh của sáng chế, phân tử axit nucleic tốt hơn là đoạn của ADN.

"Liên kết hoạt động" đề cập đến sự liên kết của polynucleotit trên đoạn axit nucleic đơn sao cho chức năng của cái này ảnh hưởng đến chức năng của cái kia. Ví dụ, vùng khởi động liên kết hoạt động với polynucleotit mã hóa hoặc ARN chức năng khi nó có thể ảnh hưởng đến biểu hiện của polynucleotit mã hóa hoặc ARN chức năng đó (tức là, polynucleotit mã hóa hoặc ARN chức năng đó nằm dưới sự kiểm soát phiên mã của vùng khởi động). Polynucleotit mã hóa theo hướng có nghĩa hoặc đối nghĩa có thể liên kết hoạt động với polynucleotit điều chỉnh.

Như được sử dụng ở đây "diệt loài gây hại," diệt côn trùng," và tương tự, đề cập đến khả năng của protein Cry theo sáng chế để kiểm soát sinh vật gây hại hoặc lượng của protein Cry mà có thể kiểm soát sinh vật gây hại như được định nghĩa ở đây. Do đó, protein Cry diệt loài gây hại có thể giết chết hoặc ức chế khả năng (ví dụ, loài gây hại côn trùng) sống sót, phát triển, ăn, hoặc tái sinh của sinh vật gây hại.

"Cây trồng" là bất kỳ cây trồng ở bất kỳ giai đoạn phát triển, cụ thể là thực vật có hạt.

"Tế bào của cây" là đơn vị cấu trúc và sinh lý của cây trồng, bao gồm tế bào trần và thành tế bào. Tế bào của cây có thể dưới dạng tế bào đơn được phân lập hoặc tế bào được nuôi cấy, hoặc là một phần của đơn vị được tổ chức cao hơn như, ví dụ, mô thực vật, cơ quan của cây, hoặc toàn bộ cây.

"Mẻ nuôi cấy tế bào của cây" nghĩa là mẻ nuôi cấy các đơn vị của cây như, ví dụ, tế bào trần, tế bào nuôi cấy tế bào, tế bào trong mô thực vật, phấn hoa, ống phấn, noãn, túi phôi, hợp tử và phôi ở giai đoạn phát triển khác nhau.

"Vật liệu cây" đề cập đến lá, thân, rễ, hoa hoặc các bộ phận của hoa, quả, phấn hoa, tế bào trứng, hợp tử, hạt, cành giâm, mẻ nuôi cấy tế bào hoặc mô, hoặc bất kỳ bộ phận hoặc sản phẩm khác của cây trồng.

"Cơ quan của cây" là phần được cấu trúc và phân biệt rõ ràng và nhìn thấy được của cây trồng như rễ, thân, lá, nụ hoa, hoặc phôi.

"Mô thực vật" như được sử dụng trong bản mô tả này nghĩa là nhóm tế bào của cây được tổ chức thành một đơn vị cấu trúc và chức năng. Mô thực vật bất kỳ trong thử nghiệm ở thực vật hoặc trong nuôi cấy đều bao gồm. Thuật ngữ này bao gồm, nhưng không giới hạn ở, tổng thể cây, cơ quan của cây, hạt của cây, nuôi cấy mô và nhóm bất kỳ của tế bào cây được tổ chức thành đơn vị cấu trúc hoặc chức năng. Việc sử dụng các thuật ngữ này kết hợp với, hoặc không có mặt, dạng đặc trưng bất kỳ của mô thực vật như được liệt kê ở trên hoặc bao gồm dạng khác theo định nghĩa ngày không nhằm loại trừ bất kỳ dạng nào của mô thực vật.

"Polynucleotit" đề cập đến polyme gồm có nhiều nucleotit monome liên kết cộng hóa trị trong mạch. "Polynucleotit" này bao gồm ADN, ARN, oligo nucleotit được biến đổi (ví dụ, oligonucleotit bao gồm các bazơ mà không cơ bản với ARN hoặc và sinh học, như 2'-O-metyl hóa oligonucleotit), và tương tự. Theo một số phương án, axit nucleic hoặc polynucleotit có thể là sợi đơn, sợi đôi, sợi đa, hoặc các kết hợp của chúng. Trừ khi có chỉ định khác, axit nucleic hoặc polynucleotit cụ thể theo sáng chế tùy ý bao gồm hoặc mã hóa polynucleotit bổ sung, ngoài bất kỳ polynucleotit được chỉ rõ.

"Polynucleotit có lợi" đề cập đến polynucleotit bất kỳ mà, khi được chuyển vào sinh vật, ví dụ, cây, tạo ra cho sinh vật đặc trưng mong muốn như sức kháng côn trùng, sức kháng bệnh, khả năng chịu chất diệt cỏ, sức kháng kháng sinh, giá trị dinh dưỡng

cải thiện, năng suất cải thiện trong quá trình xử lý công nghiệp, sản xuất các enzym hoặc chất chuyển hóa có giá trị kinh tế hoặc khả năng sinh sản thay đổi.

Thuật ngữ "vùng khởi động" đề cập đến polynucleotit, thường ngược dòng (5') với polynucleotit mã hóa của nó, mà kiểm soát biểu hiện của polynucleotit mã hóa bằng cách tạo ra nhận diện cho ARN polymeraza và các yếu tố khác cần cho phiên mã đúng.

"Tế bào trần" là tế bào của cây phân lập mà không có thành tế bào hoặc chỉ có các phần của thành tế bào.

Như được sử dụng ở đây, thuật ngữ "tái tổ hợp" đề cập đến dạng axit nucleic (ví dụ, ADN hoặc ARN) hoặc protein hoặc sinh vật mà thường không thấy trong tự nhiên và do đó được tạo ra bởi sự can thiệp của con người. Như được sử dụng ở đây, "phân tử axit nucleic tái tổ hợp" là phân tử axit nucleic bao gồm kết hợp của polynucleotit mà không xảy ra trong tự nhiên cùng nhau và là kết quả của can thiệp của con người, ví dụ, phân tử axit nucleic mà bao gồm kết hợp của ít nhất hai polynucleotit khác loại với nhau, hoặc phân tử axit nucleic mà được tổng hợp nhân tạo và bao gồm polynucleotit mà lệch khỏi polynucleotit mà có thể tồn tại bình thường trong tự nhiên, hoặc phân tử axit nucleic mà bao gồm gen chuyển được kết hợp nhân tạo vào ADN bộ gen của tế bào chủ và ADN kẹp hai bên được kết hợp của bộ gen của tế bào chủ. Ví dụ của phân tử axit nucleic tái tổ hợp là phân tử ADN tạo ra từ việc chèn gen chuyển vào ADN bộ gen của cây, mà cuối cùng có thể tạo ra biểu hiện của ARN tái tổ hợp hoặc phân tử protein trong sinh vật đó. Như được sử dụng ở đây, "cây tái tổ hợp" là cây mà không tồn tại bình thường trong tự nhiên, là kết quả của can thiệp của con người, và chứa gen chuyển hoặc phân tử axit nucleic khác loại được kết hợp vào bộ gen của nó. Do việc thay đổi bộ gen này, thực vật tái tổ hợp rõ ràng khác với thực vật loại kiểu dại liên quan.

"Các yếu tố điều chỉnh" đề cập đến các trình tự liên quan đến kiểm soát biểu hiện của trình tự nucleotit. Các yếu tố điều chỉnh bao gồm vùng khởi động liên kết hoạt động với trình tự nucleotit được mong muốn và các tín hiệu kết thúc. Chúng cũng thường bao hàm trình tự cần cho sự dịch mã thích hợp của trình tự nucleotit.

Thuật ngữ "sự tương đồng" hoặc "tương đồng" hoặc "cơ bản tương đồng," trong bối cảnh của hai trình tự axit nucleic hoặc axit amin, đề cập đến hai hoặc nhiều trình tự hoặc dưới trình tự mà có độ tương đồng gốc nucleotit hoặc axit amin, ít nhất 60%, tốt hơn là ít nhất 80%, tốt hơn nữa là 90%, còn tốt hơn nữa là 95%, và tốt nhất là ít nhất

99% khi so sánh và sắp thẳng hàng đối với sự tương ứng tối đa, như được đo bằng cách sử dụng một trong các thuật toán so sánh trình tự sau hoặc bằng cách kiểm tra bằng mắt. Tốt hơn là, cơ bản tương đồng tồn tại trên vùng trình tự mà có chiều dài bằng ít nhất khoảng 50 gốc hoặc bazơ, tốt hơn nữa là trên vùng có nhất khoảng 100 gốc hoặc bazơ, và tốt nhất là trình tự cơ bản tương đồng trên ít nhất khoảng 150 gốc hoặc bazơ. Theo phương án được ưu tiên đặc biệt, các trình tự là cơ bản tương đồng trên toàn bộ chiều dài của vùng mã hóa. Ngoài ra, trình tự axit nucleic hoặc axit amin cơ bản tương đồng cơ bản thực hiện cùng chức năng.

Để so sánh trình tự, điển hình là một trình tự đóng vai trò như trình tự quy chiếu mà trình tự thử được so sánh với nó. Khi dùng thuật toán so sánh trình tự, các trình tự thử và quy chiếu được nhập vào máy tính, sau đó các hệ tọa độ được chỉ định nếu cần và các thông số của chương trình thuật toán trình tự được chỉ định. Thuật toán so sánh trình tự sau đó sẽ tính toán phần trăm tương đồng trình tự của (các) trình tự thử so với trình tự quy chiếu, dựa trên các thông số chương trình được chỉ định.

Sự sắp thẳng hàng trình tự tối ưu để so sánh có thể được thực hiện, ví dụ, bằng thuật toán tương đồng cục bộ của Smith & Waterman, Adv. Appl. Math. 2: 482 (1981), bằng thuật toán sắp thẳng hàng tương đồng của Needleman & Wunsch, J. Mol. Biol. 48: 443 (1970), bằng tra cứu phương pháp tương tự của Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85: 2444 (1988), bằng cách thực hiện máy tính hoặc các thuật toán này (GAP, BESTFIT, FASTA, và TFASTA trong gói phần mềm Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, Wis.), hoặc bằng cách kiểm tra bằng mắt (thường xem, Ausubel và cộng sự, dưới đây).

Một ví dụ về thuật toán mà thích hợp để xác định phần trăm tương đồng trình tự và tính tương đồng trình tự là thuật toán BLAST, mà được mô tả trong Altschul và cộng sự, J. Mol. Biol. 215: 403-410 (1990). Phần mềm để thực hiện phân tích BLAST được công bố qua Trung tâm Thông tin Công nghệ sinh học Quốc gia (National Center for Biotechnology Information, U.S. National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894 USA). Thuật toán này liên quan đến việc xác định trước tiên các cặp trình tự có điểm cao (các HSP) bằng cách xác định các từ ngắn của chiều dài W trong trình tự truy vấn, mà hoặc là tương xứng hoặc là thỏa mãn điểm ngưỡng T nào đó có giá trị dương khi được sắp thẳng hàng với từ có cùng chiều dài trong trình tự cơ sở dữ liệu. T được gọi là ngưỡng điểm của từ lân cận (Altschul và cộng sự, 1990). Các

tương hợp từ lân cận khởi đầu này có tác dụng như hạt giống để bắt đầu tìm kiếm các HSP dài hơn có chứa chúng. Các tương hợp từ này sau đó được mở rộng theo cả hai hướng dọc theo mỗi trình tự xa tới mức điểm sắp thẳng hàng tích lũy có thể tăng lên. Các điểm tích lũy được tính bằng cách sử dụng, đối với trình tự nucleotit, thông số M (điểm thưởng đối với cặp các gốc tương hợp; luôn luôn >0) và N (điểm phạt đối với các cặp không tương hợp; luôn luôn <0). Đối với các trình tự axit amin, ma trận điểm được dùng để tính toán điểm tích lũy. Việc mở rộng các tương hợp từ theo mỗi hướng tạm dừng lại khi điểm sắp thẳng hàng tích lũy giảm tới lượng X kể từ giá trị đạt được tối đa của nó, điểm tích lũy tiến tới không hoặc thấp hơn do tích lũy một hoặc nhiều phép sắp thẳng hàng nhóm có điểm âm, hoặc đạt đến điểm cuối của trình tự nào đó. Các thông số W, T, và X của thuật toán BLAST xác định độ nhạy và tốc độ của phép sắp thẳng hàng. Chương trình BLASTN (cho các trình tự nucleotit) dùng mặc định độ dài từ (W) là 11, mức mong đợi (E) là 10, ngưỡng giới hạn là 100, $M=5$, $N=-4$, và so sánh cả hai mạch. Đối với trình tự axit amin, chương trình BLASTP dùng mặc định độ dài từ (W) là 3, mức mong đợi (E) là 10, và ma trận cho điểm BLOSUM62 (xem Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 (1989)).

Ngoài việc tính toán phần trăm tương đồng trình tự, thuật toán BLAST còn thực hiện phân tích thống kê tính tương tự giữa hai trình tự (xem, ví dụ, Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90: 5873-5787 (1993)). Một phép đo tương tự được đưa ra bởi thuật toán BLAST là xác suất tổng nhỏ nhất ($P(N)$), nó cho chỉ dấu về xác suất mà do nó, tương xứng giữa hai trình tự nucleotit hoặc axit amin có thể ngẫu nhiên xảy ra. Ví dụ, trình tự axit nucleic thử nghiệm được coi là tương tự với trình tự quy chiếu nếu tổng xác suất nhỏ nhất khi so sánh trình tự axit nucleic thử nghiệm với trình tự axit nucleic quy chiếu nhỏ hơn khoảng 0,1, tốt hơn nữa là nhỏ hơn khoảng 0,01, và tốt nhất là nhỏ hơn khoảng 0,001.

Dấu hiệu khác cho thấy hai trình tự axit nucleic căn bản là đồng nhất là hai phân tử lai với nhau dưới các điều kiện nghiêm ngặt. Thuật ngữ "lai đặc biệt" đề cập đến liên kết, ghép đôi, hoặc lai phân tử chỉ với trình tự nucleotit cụ thể dưới các điều kiện nghiêm ngặt khi mà trình tự có mặt trong hỗn hợp phức (ví dụ, tế bào tổng) ADN hoặc ARN. "(các) liên kết cơ bản" đề cập đến việc lai bổ sung giữa axit nucleic dò và axit nucleic đích và bao gồm các lỗi bắt cặp nhỏ mà có thể điều tiết bằng cách giảm sự nghiêm ngặt của môi trường lai để thu được việc dò trình tự axit nucleic đích mong muốn.

"Các điều kiện lai nghiêm ngặt" và "các điều kiện rửa lai nghiêm ngặt" trong ngữ cảnh các thử nghiệm lai axit nucleic như lai Southern và Northern là phụ thuộc trình tự, và khác nhau theo các thông số môi trường khác nhau. Các trình tự dài hơn lai đặc biệt ở các nhiệt độ càng cao. Hướng dẫn khái quát về lai các axit nucleic được tìm thấy trong Kỹ thuật thí nghiệm Tijssen (1993) trong Lai Hóa sinh và sinh học phân tử với trình tự dò Axit nucleic phần I chương 2 "Overview of principles of hybridization và the strategy of axit nucleic probe assays" Elsevier, New York. Thông thường, các điều kiện lai và rửa nghiêm ngặt cao được chọn là thấp hơn khoảng 5°C so với điểm nóng chảy nhiệt (T_m) cho trình tự cụ thể ở cường độ ion và pH xác định. Thông thường, dưới "các điều kiện nghiêm ngặt" trình tự dò sẽ lai với trình tự phụ đích, mà không lai với các trình tự khác.

T_m là nhiệt độ (ở cường độ ion và pH xác định) mà ở đó 50% trình tự đích lai với trình tự dò bắt cặp hoàn toàn. Các điều kiện rất nghiêm ngặt được chọn là bằng T_m đối với đoạn dò cụ thể. Ví dụ của các điều kiện lai nghiêm ngặt để lai axit nucleic bổ sung mà có hơn 100 gốc bổ sung trên dụng cụ lọc trong kỹ thuật lai Southern hoặc Northern blot là 50% formamit với 1mg heparin ở 42°C , với việc lai được thực hiện qua đêm. Ví dụ của các điều kiện rửa nghiêm ngặt cao là NaCl 0,15M ở 72°C trong khoảng 15 phút. Ví dụ của các điều kiện rửa nghiêm ngặt là rửa $0,2\times$ SSC ở 65°C trong 15 phút (xem, Sambrook, dưới đây, đối với mô tả đậm SSC). Thông thường, việc rửa có độ nghiêm ngặt cao được thực hiện sau khi rửa nghiêm ngặt thấp để tách dấu hiệu dò của nền. Việc rửa có độ nghiêm ngặt trung bình ví dụ đối với ghép đôi của, ví dụ, hơn 100 nucleotit, là $1\times$ SSC ở 45°C trong 15 phút. Việc rửa có độ nghiêm ngặt thấp ví dụ đối với ghép đôi của, ví dụ, hơn 100 nucleotit, là $4-6\times$ SSC ở 40°C trong 15 phút. Đối với các trình tự dò ngắn (ví dụ, khoảng 10 đến 50 nucleotit), các điều kiện nghiêm ngặt thường liên quan đến các nồng độ muối nhỏ hơn khoảng Na ion 1,0 M, thường nồng độ khoảng 0,01 đến 1,0 M Na ion (hoặc các muối khác) ở pH 7,0 đến 8,3, và nhiệt độ thường ít nhất khoảng 30°C . Các điều kiện nghiêm ngặt còn có thể thu được với điều kiện của các tác nhân gây mất ổn định như formamit. Thông thường, tín hiệu đối với tỷ lệ nhiều hơn gấp $2\times$ (hoặc cao hơn) quan sát được đối với trình tự dò không liên quan trong thử nghiệm lai cụ thể cho thấy phát hiện lai đặc biệt. Các axit nucleic mà không lai với nhau trong điều kiện nghiêm ngặt về cơ bản vẫn giống nhau nếu các protein mà chúng mã hóa cho

về cơ bản giống nhau. Điều này xảy ra, ví dụ, khi bản sao của axit nucleic được tạo ra bằng cách sử dụng sự thoái hóa bộ ba mã hóa lớn nhất cho phép bởi mã di truyền.

Sau đây là các ví dụ của bộ các điều kiện lai/rửa mà có thể được sử dụng để tách dòng trình tự nucleotit cùng loại mà cơ bản tương đồng với trình tự nucleotit quy chiếu theo sáng chế: trình tự nucleotit quy chiếu tốt hơn là lai với trình tự nucleotit quy chiếu trong 7% natri dodecyl sulfat (SDS), 0,5 M NaPO₄, 1 mM EDTA ở 50°C với rửa trong 2× SSC, 0,1% SDS ở 50°C, mong muốn hơn là trong 7% natri dodecyl sulfat (SDS), 0,5 M NaPO₄, 1 mM EDTA ở 50°C với rửa trong 1× SSC, 0,1% SDS ở 50°C, mong muốn hơn nữa là trong 7% natri dodecyl sulfat (SDS), 0,5 M NaPO₄, 1 mM EDTA ở 50°C với rửa trong 0,5× SSC, 0,1% SDS ở 50°C, tốt hơn là trong 7% natri dodecyl sulfat (SDS), 0,5 M NaPO₄, 1 mM EDTA ở 50°C với rửa trong 0,1× SSC, 0,1% SDS ở 50°C, tốt hơn nữa là trong 7% natri dodecyl sulfat (SDS), 0,5 M NaPO₄, 1 mM EDTA ở 50°C với rửa trong 0,1× SSC, 0,1% SDS ở 65°C.

Dấu hiệu khác của việc hai trình tự axit nucleic hoặc protein về cơ bản giống nhau là protein được mã hóa bởi axit nucleic thứ nhất phản ứng chéo về mặt miễn dịch với, hoặc liên kết đặc hiệu với, protein được mã hóa bởi axit nucleic thứ hai. Do đó, protein này thường là về cơ bản giống với protein thứ hai, ví dụ, trong đó hai protein chỉ khác nhau bởi sự thay thế bảo toàn.

"Tổng hợp" đề cập đến trình tự nucleotit bao gồm bazơ hoặc dấu hiệu cấu trúc mà không có mặt trong trình tự tự nhiên. Ví dụ, trình tự nhân tạo mã hóa protein Cry theo sáng chế mà giống gần hơn hàm lượng G+C và phân phối bộ ba mã hóa bình thường của gen của cây hai lá mầm hoặc một lá mầm được nói là tổng hợp.

Như được sử dụng ở đây, protein Cry mà "độc" đối với loài gây hại côn trùng nghĩa là protein Cry thực hiện chức năng như tác nhân kiểm soát côn trùng hoạt tính qua đường miệng để giết loài gây hại côn trùng, hoặc protein Cry có khả năng phá vỡ hoặc cản trở côn trùng ăn, hoặc gây ra ức chế sinh trưởng đối với loài gây hại côn trùng, cả hai trong số chúng có thể hoặc không thể gây ra chết côn trùng. Khi protein Cry theo sáng chế được phân phối đến côn trùng hoặc côn trùng tiếp xúc miệng với protein Cry, kết quả thường là chết côn trùng, hoặc sinh trưởng của côn trùng bị chậm lại, hoặc côn trùng ngừng ăn nguồn mà tạo ra protein Cry độc đối với côn trùng.

"Biến nạp" là quá trình đưa axit nucleic khác loại vào tế bào chủ hoặc sinh vật. Cụ thể là, "biến nạp" nghĩa là tích hợp ổn định phân tử ADN vào bộ gen của sinh vật được mong muốn.

"Được biến nạp / chuyển gen / tái tổ hợp" đề cập đến sinh vật chủ như vi khuẩn hoặc thực vật mà phân tử axit nucleic khác loại được đưa vào đó. Phân tử axit nucleic có thể được tích hợp ổn định vào bộ gen của vật chủ hoặc phân tử axit nucleic cũng có thể hiện diện như một phân tử ngoài nhiễm sắc thể. Phân tử ngoài nhiễm sắc thể này có thể tự sao chép. Các tế bào, mô hoặc cây được biến nạp được hiểu là không chỉ chứa sản phẩm cuối của quá trình biến nạp mà còn chứa cả thể hệ con cái chuyển gen của nó. Vật chủ "không biến nạp", "không chuyển gen", hoặc "không tái tổ hợp" đề cập đến sinh vật kiểu đại, ví dụ, vi khuẩn hoặc thực vật, mà không chứa phân tử axit nucleic khác loại.

Các nucleotit được xác định bởi các bazơ của chúng nhờ các chữ viết tắt chuẩn sau đây: adenin (A), xytosin (C), thymin (T), và guanin (G). Tương tự các axit amin được biểu thị bằng các chữ viết tắt tiêu chuẩn sau: alanin (Ala; A), arginin (Arg; R), asparagin (Asn; N), axit aspartic (Asp; D), xystein (Cys; C), glutamin (Gln; Q), axit glutamic (Glu; E), glyxin (Gly; G), histidin (His; H), isoloxin (Ile; I), lōxin (Leu; L), lysin (Lys; K), methionin (Met; M), phenylalanin (Phe; F), prolin (Pro; P), sorin (Ser; S), threonin (Thr; T), tryptophan (Trp; W), tyrosin (Tyr; Y), và valin (Val; V).

Sáng chế đề xuất các chế phẩm và phương pháp để kiểm soát loài gây hại cây trồng. Cụ thể là, sáng chế đề cập đến protein Cry mà có thể được phân lập từ vi khuẩn, như *Bacillus thuringiensis*, mà độc đối với loài gây hại côn trùng và polynucleotit mà bao gồm trình tự nucleotit mà mã hóa protein Cry, và tạo ra và sử dụng polynucleotit và protein Cry để kiểm soát loài gây hại côn trùng.

Theo một số phương án, sáng chế đề xuất phân tử axit nucleic hoặc tùy ý phân tử axit nucleic được phân lập bao gồm trình tự nucleotit mã hóa protein Cry ở dạng tiền độc tố của nó hoặc đoạn hoạt tính sinh học hoặc đoạn độc tố của chúng, trong đó trình tự nucleotit (a) có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ trong số SEQ ID NO:1-10 hoặc trình tự mã hóa độc tố của chúng; hoặc (b) mã hóa protein bao gồm trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ trong số SEQ ID NO:31-40 hoặc đoạn độc tố của chúng; hoặc (c) là trình tự tổng hợp của (a) hoặc (b) mà có bộ ba mã hóa được tối ưu để biểu hiện trong sinh vật chuyển

gen. Theo các phương án khác, trình tự nucleotit bao gồm bất kỳ trong số SEQ ID NO:1-10 hoặc bất kỳ đoạn mã hóa độc tố nêu ở bất kỳ trong số SEQ ID NO:1-10. Theo các phương án khác, trình tự nucleotit tổng hợp bao gồm bất kỳ trong số SEQ ID NO:11-30 hoặc bất kỳ đoạn mã hóa độc tố nêu ở bất kỳ trong số SEQ ID NO:11-30.

Polynucleotit mà là đoạn của polynucleotit mã hóa tiền độc tố protein Cry cũng thuộc phạm vi của sáng chế. Thuật ngữ "đoạn" để chỉ một phần của trình tự nucleotit mã hóa protein Cry. Đoạn của trình tự nucleotit có thể mã hóa phần hoạt tính sinh học của protein Cry, còn gọi là "đoạn độc tố," hoặc nó có thể là đoạn mà có thể được sử dụng làm mẫu dò lai hoặc đoạn môi PCR sử dụng phương pháp được bộc lộ dưới đây. Phân tử axit nucleic mà là đoạn của protein Cry mã hóa trình tự nucleotit bao gồm ít nhất khoảng 15, 20, 50, 75, 100, 200, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450 nucleotit kề nhau, hoặc đến số lượng nucleotit có mặt trong protein Cry chiều dài đầy đủ mã hóa trình tự nucleotit được bộc lộ ở đây (ví dụ, 3546 nucleotit đối với SEQ ID NO:1) phụ thuộc vào mục đích dự kiến. Thuật ngữ nucleotit "kề nhau" nhằm chỉ gốc nucleotit mà liền ngay cái khác. Một số đoạn của trình tự nucleotit theo sáng chế sẽ mã hóa đoạn độc tố mà giữ hoạt tính sinh học của protein Cry và, do đó, giữ hoạt tính diệt côn trùng. Thuật ngữ "giữ hoạt tính diệt côn trùng" nhằm chỉ đoạn sẽ có ít nhất khoảng 30%, tốt hơn là ít nhất khoảng 50%, tốt hơn nữa là ít nhất khoảng 70%, còn tốt hơn nữa là ít nhất khoảng 80% hoạt tính diệt côn trùng của protein Cry. Phương pháp để đo hoạt tính diệt côn trùng là đã biết trong lĩnh vực. Xem, ví dụ, Czapla và Lang (1990) J. Econ. Entomol. 83:2480-2485; Andrews và cộng sự (1988) Biochem. J. 252:199-206; Marrone và cộng sự (1985) J. of Economic Entomology 78:290-293; và Sáng chế Mỹ No. 5,743,477, toàn bộ trong số chúng được kết hợp ở đây bằng cách viện dẫn toàn bộ.

Đoạn độc tố của protein Cry theo sáng chế sẽ mã hóa ít nhất khoảng 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, và 450 axit amin liền nhau, hoặc lên đến tổng số axit amin có trong protein Cry chiều dài đầy đủ theo sáng chế (ví dụ, 1181 axit amin đối với SEQ ID NO:31).

Theo một số phương án, phân tử axit nucleic theo sáng chế bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mã hóa protein Cry bao gồm trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ trong số SEQ ID NO:31-40 hoặc đoạn độc tố của chúng. Theo một số phương án khác, trình tự axit amin bao gồm,

cơ bản chứa hoặc chứa bất kỳ trong số SEQ ID NO:31-40 hoặc đoạn độc tố của chúng. Do đó, theo một số phương án, protein Cry mà được hoạt hóa bằng cách xử lý phân giải protein, ví dụ, bằng proteaza được điều chế từ ruột của côn trùng, có thể được đặc trưng hóa và axit amin đầu N hoặc đầu C của đoạn độc tố hoạt tính được xác định. Theo khía cạnh này của sáng chế, chuyên gia trong lĩnh vực có thể xác định rằng, ví dụ, đoạn độc tố nêu ở SEQ ID NO:31 hoặc SEQ ID NO:32 hoặc SEQ ID NO:33 có thể bao gồm axit amin từ khoảng 29 đến khoảng 603 hoặc từ khoảng 29 đến khoảng 624 hoặc từ khoảng 29 đến khoảng 635 nêu ở SEQ ID NO:31 hoặc SEQ ID NO:32 hoặc SEQ ID NO:33, hoặc đoạn độc tố nêu ở SEQ ID NO:34 có thể bao gồm axit amin từ khoảng axit amin 21 đến khoảng 606 hoặc từ khoảng 21 đến khoảng 645 hoặc từ khoảng 21 đến khoảng 648 nêu ở SEQ ID NO:34, và vân vân. Protein biến thể Cry được tạo ra bằng cách đưa vào hoặc loại bỏ vị trí xử lý proteaza ở vị trí thích hợp trong trình tự mã hóa cho phép, hoặc loại bỏ, phân cắt phân giải protein của protein biến thể lớn hơn bởi proteaza côn trùng, cây hoặc vi sinh vật cũng thuộc phạm vi của sáng chế. Kết quả cuối cùng của thao tác này được hiệu là tạo thành phân tử đoạn độc tố có hoạt tính giống hoặc tốt hơn so với protein tiền độc tố nguyên vẹn.

Theo một số phương án của sáng chế, gen khảm được đề xuất mà bao gồm vùng khởi động khác loại liên kết hoạt động với polynucleotit bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mà mã hóa protein Cry độc đối với loài gây hại cánh vảy, trong đó trình tự nucleotit (a) có độ tương đồng trình tự ít nhất 80% (ví dụ, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) đến ít nhất 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) với bất kỳ một trong số các SEQ ID NO:1-10, hoặc trình tự mã hóa độc tố của chúng; hoặc (b) mã hóa protein bao gồm trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% (ví dụ, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) đến ít nhất 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) với bất kỳ một trong số các SEQ ID NO:31-40, hoặc đoạn độc tố của chúng; hoặc (c) là trình tự tổng hợp của (a) hoặc (b) mà có bộ ba mã hóa được tối ưu để biểu hiện trong sinh vật chuyển gen.

Theo các phương án khác, vùng khởi động khác loại là vùng khởi động có thể biểu hiện thực vật. Ví dụ, không giới hạn, vùng khởi động có thể biểu hiện ở cây có thể được chọn từ nhóm các vùng khởi động chứa ubiquitin, virus dạ hương vàng, TrpA của ngô, OsMADS 6, histon H3 của ngô, gen 9 5' UTR của vật ăn vi khuẩn T3, sucroza synthetaza 1 của ngô, alcohol dehydrogenaza 1 của ngô, phức thu sáng của ngô, protein sốc nhiệt của ngô, mtl của ngô, RuBP cacboxylaza khối phụ nhỏ của đậu Hà Lan, actin của lúa, xyclophilin của lúa, Mannopin synthaza plasmit Ti, nopalin synthaza plasmit Ti, chalcon isomeraza của dã yên thảo, protein 1 giàu glyxin của đậu, patatin của khoai tây, lectin, vùng khởi động RuBP cacboxylaza khối phụ nhỏ CaMV 35S và S-E9.

Theo các phương án bổ sung, protein được mã hóa bởi gen khả năng là độc đối với một hoặc nhiều loài gây hại cánh vảy được chọn từ nhóm bao gồm sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xám (*Agrotis ipsilon*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu xanh hại ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá velvetbean (*Anticarsia gemmatilis*), sâu đo hại đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục trái bắp (*Richia albicosta*), tuyến trùng ăn rễ cây thuốc lá (*Heliothis virescens*), Sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc nâu (*Chilo suppressalis*), sâu đục thân mình hồng (*Sesamia calamistis*) và sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*).

Theo các phương án khác, polynucleotit bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:1, hoặc trình tự mã hóa độc tố của chúng, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:2, hoặc trình tự mã hóa độc tố của chúng, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:3, hoặc trình tự mã hóa độc tố của chúng, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:4, hoặc trình tự mã hóa độc tố của chúng, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:5, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:6, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:7, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:8, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:9, hoặc có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với SEQ ID NO:10.

Theo các phương án khác, polynucleotit bao gồm, cơ bản chứa hoặc chứa bất kỳ một trong số các SEQ ID NO:1-10, hoặc trình tự mã hóa độc tổ của chúng.

Theo các phương án khác, polynucleotit bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mà mã hóa protein bao gồm, cơ bản chứa hoặc chứa trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ một trong số các SEQ ID NO:31-40, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:31.

Theo các phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:32.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:33, hoặc đoạn độc tổ của chúng.

Theo các phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:34, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:35.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:36.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:37, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:38, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89% ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất

99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:39, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:40.

Theo một số phương án, gen khảm theo sáng chế bao gồm polynucleotit bao gồm, cơ bản chứa hoặc chứa trình tự tổng hợp của trình tự nucleotit mà có ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với bất kỳ trong số SEQ ID NO:11-30, hoặc trình tự mã hóa độc tổ của chúng, trong đó trình tự tổng hợp có bộ ba mã hóa được tối ưu để biểu hiện là sinh vật chuyển gen, như cây chuyển gen hoặc vi khuẩn chuyển gen. Theo các phương án khác, gen khảm theo sáng chế bao gồm polynucleotit bao gồm, cơ bản chứa hoặc chứa trình tự tổng hợp của trình tự nucleotit mà mã hóa protein bao gồm trình tự axit amin mà có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với bất kỳ trong số SEQ ID NO:31-50, hoặc đoạn độc tổ của chúng, trong đó trình tự tổng hợp có bộ ba mã hóa được tối ưu để

biểu hiện là sinh vật chuyển gen. Theo các phương án khác nữa, sinh vật chuyển gen là vi khuẩn chuyển gen hoặc cây chuyển gen.

Theo một số phương án, sáng chế đề xuất polynucleotit tổng hợp bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mà mã hóa protein mà độc đối với loài gây hại cánh vảy, trong đó trình tự nucleotit có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với bất kỳ một trong số các SEQ ID NO:11-30, hoặc trình tự mã hóa độc tố của chúng.

Theo các phương án khác, sáng chế đề xuất polynucleotit tổng hợp bao gồm, cơ bản chứa hoặc chứa trình tự nucleotit mà mã hóa protein mà độc đối với loài gây hại cánh vảy, trong đó trình tự nucleotit mã hóa trình tự axit amin mà có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với bất kỳ một trong số các SEQ ID NO:31-50, hoặc đoạn độc tố của chúng.

Protein Cry theo sáng chế có thể được phân lập từ một số chủng *Bacillus thuringiensis* (*Bt*) nhất định như C0530, C0537, C0651, C0652, C0724, C0801, C1079 và M2776. Cần hiểu rằng protein Cry theo sáng chế còn có thể được phân lập từ các chủng *Bt* khác và các chủng *Bt* này có thể được phân lập bằng các kỹ thuật tiêu chuẩn và được thử nghiệm về độc tố đối với loài gây hại cánh vảy theo sáng chế. Thông thường các chủng *Bt* có thể được phân lập từ bất kỳ mẫu môi trường, bao gồm đất, cây, côn trùng, bụi nân hạt, và vật liệu mẫu khác, v.v., bằng các phương pháp đã biết trong lĩnh vực. Xem, ví dụ, Travers và cộng sự (1987) Appl. Environ. Microbiol. 53:1263-1266; Saleh và cộng sự (1969) Can J. Microbiol. 15:1101-1104; DeLucca và cộng sự (1981)

Can J. Microbiol. 27:865-870; và Norris, và cộng sự (1981) "The genera Bacillus and SporolactoBacillus," In Starr và cộng sự (eds.), The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria, Tập II, Springer-Verlog Berlin Heidelberg. Sau khi phân lập, các chủng *Bt* có thể được thử nghiệm về độc tố đối với loài gây hại cánh vảy và protein Cry bao gồm trong sáng chế có thể được xác định. Do đó, theo một số phương án, sáng chế đề xuất chủng *Bacillus thuringiensis* (*Bt*) phân lập mà tạo ra protein Cry hoặc protein Cry tái tổ hợp bao gồm, cơ bản chứa hoặc chứa trình tự axit amin có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ trong số SEQ ID NO: 31-40. Theo các phương án khác, chủng *Bt* được chọn từ nhóm bao gồm C0530, C0537, C0651, C0652, C0724, C0801, C1079 và M2776. Theo các phương án khác nữa, protein Cry hoặc protein Cry tái tổ hợp bao gồm, cơ bản chứa hoặc chứa nêu ở bất kỳ trong số SEQ ID NO:31-40.

Theo một số phương án, sáng chế đề xuất protein Cry được phân lập tùy ý mà độc đối với loài gây hại cánh vảy, trong đó protein bao gồm, cơ bản chứa hoặc chứa (a) trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% tương đồng trình tự đến ít nhất 99% với trình tự axit amin được thể hiện bởi bất kỳ một trong số các SEQ ID NO:31-50, hoặc đoạn độc tố của chúng; hoặc (b) trình tự axit amin mà được mã hóa bởi trình tự nucleotit mà có ít nhất 80% tương đồng trình tự đến ít nhất 99% tương đồng trình tự với trình tự nucleotit được thể hiện bởi bất kỳ một trong số các SEQ ID NO:11-30, hoặc đoạn mã hóa độc tố của chúng.

Theo các phương án khác, protein Cry được phân lập tùy ý bao gồm, cơ bản chứa hoặc chứa trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ một trong số các SEQ ID NO:31-50, hoặc đoạn độc tố của chúng. Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:31, hoặc đoạn độc tố của chúng.

trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%,

hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:31, hoặc đoạn độc tổ của chúng.

Theo các phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:33, hoặc đoạn độc tổ của chúng.

Theo các phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:34, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:35, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất

99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:36, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:37, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:38, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89% ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:39, hoặc đoạn độc tổ của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89% ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 93%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:40.

Theo một số phương án, trình tự axit amin bao gồm, cơ bản chứa hoặc chứa bất kỳ một trong số các SEQ ID NO:31-50, hoặc đoạn độc tố của chúng. Theo các phương án khác, trình tự axit amin được mã hóa bởi trình tự nucleotit bao gồm, cơ bản chứa hoặc chứa bất kỳ trong số SEQ ID NO:1-30, hoặc trình tự mã hóa độc tố của chúng.

Theo các phương án khác, protein Cry theo sáng chế là độc đối với loài gây hại cánh vảy được chọn từ nhóm bao gồm sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xám (*Agrotis ipsilon*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu xanh hại ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá velvetbean (*Anticarsia gemmatilis*), sâu đo hại đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục trái bắp (*Richia albicosta*), tuyến trùng ăn rễ cây thuốc lá (*Heliothis virescens*), Sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc nâu (*Chilo suppressalis*), sâu đục thân mình hồng (*Sesamia calamistis*) và sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*).

Theo một số phương án, sáng chế bao gồm protein Cry tái tổ hợp mà độc đối với loài gây hại cánh vảy, trong đó protein Cry tái tổ hợp bao gồm, cơ bản chứa hoặc chứa (a) trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với trình tự axit amin được thể hiện bởi bất kỳ trong số SEQ ID NO:41-50, hoặc đoạn độc tố của chúng; hoặc (b) trình tự axit amin mà được mã hóa bởi trình tự nucleotit mà có 80% đến ít nhất 99% tương đồng trình tự với trình tự nucleotit được thể hiện bởi bất kỳ trong số SEQ ID NO:21-30, hoặc trình tự mã hóa độc tố của chúng.

Theo các phương án khác, protein Cry tái tổ hợp bao gồm, cơ bản chứa hoặc chứa trình tự axit amin mà có độ tương đồng trình tự ít nhất 80% đến ít nhất 99% với bất kỳ một trong số các SEQ ID NO:41-50, hoặc đoạn độc tố của chúng. Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:41, hoặc đoạn độc tố của chúng.

Theo các phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với SEQ ID NO:50, hoặc đoạn độc tố của chúng.

Theo các phương án khác nữa, protein Cry tái tổ hợp bao gồm, cơ bản chứa hoặc chứa trình tự axit amin nêu ở bất kỳ trong số SEQ ID NO:41-50, hoặc đoạn độc tố của chúng. Theo các phương án khác, protein Cry tái tổ hợp được mã hóa bởi trình tự nucleotit that bao gồm, cơ bản chứa hoặc chứa nêu ở bất kỳ trong số SEQ ID NO:21-30, hoặc trình tự mã hóa độc tố của chúng.

Các kháng thể tạo thành đáp ứng thách thức miễn dịch bởi BT-0001, BT-0003, BT-0020, BT-0022, BT-0027, BT-0029, BT-0030, BT-0031, BT-0201 và BT-0202 tự nhiên hoặc đột biến và tương tự hoặc protein Cry liên quan cũng bao gồm trong sáng chế. Các kháng thể này có thể được tạo ra sử dụng kỹ thuật miễn dịch tiêu chuẩn để sản xuất huyết thanh miễn dịch đa dòng và, nếu muốn, làm bất diệt tế bào tạo ra kháng thể của vật chủ được chủng ngừa đối với các nguồn sản xuất kháng thể đa dòng. Các kỹ thuật để sản xuất kháng thể cho chất bất kỳ có lợi là đã biết, ví dụ, như trong Harlow và Lane (1988. Antibodies a laboratory manual. pp. 726. Cold Spring Harbor Laboratory) và như trong Goding (Kháng thể đơn dòng: Principles & practice. 1986. Academic Press, Inc., Orlando, FL). Sáng chế bao gồm protein diệt côn trùng mà phản ứng chéo với kháng thể, đặc biệt là kháng thể đơn dòng, tạo ra một hoặc nhiều protein Cry diệt côn trùng theo sáng chế.

Kháng thể được tạo ra theo sáng chế còn hữu dụng trong các thử nghiệm miễn dịch để xác định lượng hoặc sự có mặt của BT-0001, BT-0003, BT-0020, BT-0022, BT-0027, BT-0029, BT-0030, BT-0031, BT-0201 và BT-0202 tự nhiên hoặc đột biến hoặc protein Cry liên quan trong mẫu sinh học. Các thử nghiệm còn hữu dụng trong việc sản xuất kiểm soát chất lượng chế phẩm chứa một hoặc nhiều protein Cry theo sáng chế hoặc protein độc liên quan. Ngoài ra, kháng thể có thể được dùng để đánh giá tính hiệu quả của việc sản xuất tái tổ hợp một hoặc nhiều protein Cry theo sáng chế hoặc protein

liên quan, cũng như để sàng lọc thư viện biểu hiện đối với sự có mặt của trình tự nucleotit mã hóa một hoặc nhiều protein Cry theo sáng chế hoặc trình tự mã hóa protein liên quan. Kháng thể hữu dụng cũng như phối tử ái lực để tinh chế hoặc phân lập bất kỳ một hoặc nhiều protein theo sáng chế và protein liên quan. Protein Cry theo sáng chế và protein chứa epitop kháng nguyên liên quan có thể thu được bằng cách biểu hiện dư chiều dài đầy đủ hoặc một phần của trình tự mã hóa toàn bộ hoặc một phần protein Cry theo sáng chế hoặc protein liên quan trong tế bào chủ được ưu tiên.

Hiểu rằng ADN các trình tự mà mã hóa protein Cry tự nhiên theo sáng chế có thể được thay đổi bằng các phương pháp khác nhau, và các thay đổi này có thể tạo ra các trình tự ADN mã hóa các protein với trình tự axit amin khác với trình tự được mã hóa bởi protein Cry tự nhiên theo sáng chế. Protein Cry có thể được biến đổi theo nhiều cách để tạo ra protein Cry đột biến bao gồm thế, loại, cắt cụt, và chèn axit amin của một hoặc nhiều axit amin, bao gồm lên đến khoảng 2, khoảng 3, khoảng 4, khoảng 5, khoảng 6, khoảng 7, khoảng 8, khoảng 9, khoảng 10, khoảng 15, khoảng 20, khoảng 25, khoảng 30, khoảng 35, khoảng 40, khoảng 45, khoảng 50, khoảng 55, khoảng 60, khoảng 65, khoảng 70, khoảng 75, khoảng 80, khoảng 85, khoảng 90, khoảng 100, khoảng 105, khoảng 110, khoảng 115, khoảng 120, khoảng 125, khoảng 130, khoảng 135, khoảng 140, khoảng 145, khoảng 150, khoảng 155, hoặc nhiều sự thay thế, loại hoặc chèn axit amin. Phương pháp cho các thao tác này thường được biết trong lĩnh vực. Ví dụ, các biến thể trình tự axit amin của protein Cry tự nhiên có thể được điều chế bằng các đột biến trong polynucleotit mà mã hóa protein. Việc này còn có thể được thực hiện bởi một trong số một vài dạng đột biến hoặc trong tiến triển định hướng. Theo một số khía cạnh, các thay đổi được mã hóa trong trình tự axit amin sẽ cơ bản không ảnh hưởng đến chức năng của protein. Các biến thể này sẽ có hoạt tính diệt côn trùng mong muốn. Theo một số phương án theo sáng chế, trình tự nucleotit được thể hiện bởi SEQ ID NO: 1-10 được biến đổi để đưa vào sự thể axit amin trong protein được mã hóa. Theo các phương án khác, protein đột biến tạo thành được mã hóa bởi polynucleotit đột biến tổng hợp bao gồm trình tự nucleotit được thể hiện bởi bất kỳ một trong số các SEQ ID NO: 21-30. Theo các phương án khác, protein đột biến bao gồm, chủ yếu gồm có hoặc gồm có trình tự axit amin được đại diện bởi bất kỳ một trong số các SEQ ID NO: 41-50.

Cần hiểu rằng khả năng của protein diệt côn trùng để tạo ra hoạt tính diệt côn trùng có thể được cải thiện bằng cách sử dụng các kỹ thuật này tùy theo các chế phẩm theo

sáng chế. Ví dụ, nó có thể biểu hiện protein Cry trong tế bào chủ mà thể hiện tỷ lệ cao kết hợp sai lệch bazơ trong quá trình sao chép ADN, như XL-1 Red (Stratagene, La Jolla, CA). Sau khi nhân giống trong các chủng này, có thể phân lập ADN (ví dụ bằng cách điều chế plasmid ADN, hoặc bằng cách khuếch đại bằng kỹ thuật PCR và tách dòng đoạn PCR thu được vào vectơ), nuôi cấy các đột biến protein Cry trong chủng không gây đột biến, và xác định gen được đột biến có hoạt tính diệt côn trùng, ví dụ bằng cách thực hiện thử nghiệm để thử nghiệm hoạt tính diệt côn trùng. Thông thường, protein được trộn và được sử dụng trong các thử nghiệm cho ăn. Xem, ví dụ Marrone và cộng sự (1985) *J. of Economic Entomology* 78:290-293. Thử nghiệm có thể bao gồm việc cho cây tiếp xúc với một hoặc nhiều loài gây hại và xác định khả năng sống sót của cây hoặc khả năng gây ra chết loài gây hại. Các ví dụ của đột biến mà tạo ra độc tố tăng được tìm thấy trong Schnepf và cộng sự (1998) *Microbiol. Mol. Biol. Rev.* 62:775-806.

Tùy ý, thay đổi có thể được thực hiện cho trình tự axit amin theo sáng chế ở đầu amin hoặc cacboxy mà không ảnh hưởng cơ bản đến hoạt tính. Thay đổi này có thể bao gồm chèn, loại bỏ, hoặc các thay đổi được đưa vào bằng các phương pháp phân tử hiện đại, như PCR, bao gồm khuếch đại PCR mà thay thế hoặc mở rộng trình tự mã hóa protein do đưa vào trình tự mã hóa axit amin trong oligonucleotit được sử dụng trong khuếch đại PCR. Tùy ý, các trình tự protein được thêm có thể bao gồm các trình tự mã hóa protein toàn bộ, như các trình tự thường được sử dụng trong lĩnh vực để tạo ra dung hợp protein. Các protein dung hợp này thường được sử dụng để (1) làm tăng biểu hiện của protein được mong muốn (2) đưa vào vùng liên kết, hoạt tính enzym, hoặc epitop để tạo điều kiện cho hoặc là tinh chế protein, phát hiện protein, hoặc các sử dụng thử nghiệm khác đã biết trong lĩnh vực (3) bài tiết dịch hoặc dịch mã protein thành cơ quan dưới mức tế bào, như không gian chu chất của vi khuẩn gam âm, hoặc lưới nội bào của tế bào nhân thực, lưới nội bào thường tạo ra glycosyl hóa protein.

Protein Cry theo sáng chế có thể còn được đột biến để đưa vào một epitop để tạo ra các kháng thể mà nhận diện protein được đột biến. Do đó, theo một số phương án, sáng chế đề xuất protein Cry được đột biến, trong đó thể axit amin trong protein Cry tự nhiên tạo ra protein Cry đột biến có vùng kháng nguyên mà cho phép protein Cry đột biến phân biệt với protein Cry tự nhiên trong thử nghiệm phát hiện protein.

Theo một số phương án, sáng chế đề xuất phương pháp tạo ra kháng thể mà nhận diện một cách khác protein Cry được đột biến từ protein Cry tự nhiên mà protein Cry

đột biến được dẫn xuất từ đó, phương pháp bao gồm các bước thể các axit amin trong cấu trúc vòng kháng nguyên của protein Cry tự nhiên và làm tăng các kháng thể mà đặc biệt nhận diện cấu trúc vòng kháng nguyên đột biến trong protein Cry đột biến và không nhận diện protein Cry tự nhiên. Theo một phương án, cấu trúc vòng kháng nguyên được xác định trong các vùng không bảo toàn bên ngoài vùng I của protein Cry tự nhiên. Theo phương án khác, cấu trúc vòng kháng nguyên không phải là cấu trúc vòng liên quan đến nhận diện thụ thể ruột côn trùng của protein Cry hoặc liên quan đến hoạt hóa proteaza của protein Cry.

Trình tự nucleotit và axit amin biến thể theo sáng chế còn bao hàm trình tự dẫn xuất từ quy trình gây đột biến và sinh tái tổ hợp như đổi chỗ ADN. Với quy trình này, một hoặc nhiều vùng mã hóa protein độc khác nhau có thể được dùng để tạo ra protein độc mới sở hữu các đặc điểm mong muốn. Bằng cách này, các thư viện của polynucleotit tái tổ hợp được tạo ra từ quần thể của polynucleotit trình tự liên quan bao gồm các vùng trình tự mà có tương đồng trình tự cơ bản và có thể được tái tổ hợp cùng loại trong ống nghiệm hoặc trong cơ thể. Ví dụ, sử dụng nghiên cứu này, mô típ trình tự mã hóa vùng được mong muốn có thể được đổi chỗ giữa gen của diệt loài gây hại theo sáng chế và gen của diệt loài gây hại đã biết khác để thu được gen mới mã hóa protein với đặc tính được mong muốn được cải thiện, như tăng hoạt tính diệt côn trùng. Các chiến lược để đổi chỗ ADN này là đã biết trong lĩnh vực. Xem, ví dụ, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer và cộng sự (1997) *Nature Biotech.* 15:436-438; Moore và cộng sự (1997) *J. Mol. Biol.* 272:336-347; Zhang và cộng sự (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer và cộng sự (1998) *Nature* 391:288-291; và Sáng chế Mỹ Số 5,605,793 và 5,837,458.

Sự hoán đổi hoặc đổi chỗ vùng là cơ chế khác để tạo ra các protein Cry thay đổi theo sáng chế. Các vùng có thể được hoán đổi giữa các protein Cry, tạo ra các protein độc lai hoặc khảm với hoạt tính diệt loài gây hại hoặc phổ đích cải thiện. Phương pháp tạo ra protein tái tổ hợp và thử chúng về hoạt tính diệt loài gây hại là đã biết trong lĩnh vực (xem, ví dụ, Naimov và cộng sự (2001) *Appl. Environ. Microbiol.* 67:5328-5330; de Maagd và cộng sự (1996) *Appl. Environ. Microbiol.* 62:1537-1543; Ge và cộng sự (1991) *J. Biol. Chem.* 266:17954-17958; Schnepf và cộng sự (1990) *J. Biol. Chem.* 265:20923-20930; Rang và cộng sự (1999) *Appl. Environ. Microbiol.* 65:2918-2925).

Theo một số phương án, sáng chế đề xuất protein Cry lai bao gồm ở đầu C, axit amin từ protein Cry thứ nhất theo sáng chế và ở đầu N, axit amin từ protein Cry thứ hai theo sáng chế khác protein Cry thứ nhất theo sáng chế.

Theo một số phương án, sáng chế đề xuất vector tái tổ hợp chứa polynucleotit, phân tử axit nucleic, tổ hợp biểu hiện hoặc gen khả thể theo sáng chế. Theo các phương án khác, vector còn được xác định là plasmid, cosmid, phagemid, nhiễm sắc thể nhân tạo, vector thể thực khuẩn hoặc virus. Một số vector dùng trong biến nạp thực vật và các sinh vật khác là đã biết trong lĩnh vực.

Do đó, một số phương án theo sáng chế nhằm đến tổ hợp có thể biểu hiện thiết kế để biểu hiện polynucleotit và phân tử axit nucleic theo sáng chế. Như được sử dụng ở đây, "tổ hợp biểu hiện" nghĩa là phân tử axit nucleic có ít nhất trình tự kiểm soát liên kết hoạt động với trình tự nucleotit được mong muốn. Theo cách này, ví dụ, vùng khởi động của cây liên kết hoạt động với trình tự nucleotit được biểu hiện được đề xuất trong cat-xet biểu hiện để biểu hiện ở cây, bộ phận của cây hoặc tế bào cây.

Cat-xet biểu hiện bao gồm polynucleotit có lợi có thể là khả thể, nghĩa là ít nhất một trong số các thành phần của nó là khác loại với ít nhất một thành phần khác trong số các thành phần khác của nó. Tổ hợp biểu hiện có thể còn là tổ hợp mà xảy ra trong tự nhiên nhưng thu được dưới dạng tái tổ hợp hữu dụng cho biểu hiện khác loại. Thông thường, tuy nhiên, tổ hợp biểu hiện khác loại so với vật chủ, tức là, trình tự axit nucleic cụ thể của tổ hợp biểu hiện không xảy ra tự nhiên trong tế bào chủ và phải đưa vào tế bào chủ hoặc tổ tiên của tế bào chủ bằng sự kiện biến nạp.

Ngoài vùng khởi động liên kết hoạt động với trình tự nucleotit theo sáng chế, tổ hợp biểu hiện theo sáng chế còn có thể bao gồm trình tự điều chỉnh khác. Như được sử dụng ở đây, "trình tự điều chỉnh" nghĩa là trình tự nucleotit được bố trí ngược dòng (trình tự không mã hóa 5'), ở trong hoặc xuôi dòng (trình tự không mã hóa 3') của trình tự mã hóa, và mà ảnh hưởng đến sự phiên mã, sự xử lý hoặc tính ổn định ARN, hoặc dịch mã trình tự mã hóa liên quan. Các trình tự điều chỉnh bao gồm, nhưng không giới hạn ở, gen tăng cường, intron, trình tự dẫn dịch mã, các tín hiệu kết thúc, và trình tự tín hiệu bổ sung poly.

Theo một số phương án, cat-xet biểu hiện theo sáng chế có thể còn bao gồm polynucleotit mà mã hóa các đặc điểm mong muốn khác ngoài protein Cry theo sáng

chế. Cat-xet biểu hiện này bao gồm đặc điểm xếp chồng có thể được dùng để tạo ra cây, bộ phận của cây hoặc tế bào cây có phenotyp mong muốn với đặc điểm xếp chồng (tức là, xếp chồng phân tử). Kết hợp xếp chồng này ở cây còn có thể được tạo ra bằng phương pháp khác bao gồm, nhưng không giới hạn ở, lai giống cây bằng bất kỳ hệ phương pháp thông thường. Nếu được chồng lên bằng các cây biến nạp gen, trình tự nucleotit được mong muốn có thể kết hợp ở thời điểm và trật tự bất kỳ. Ví dụ, cây chuyển gen chứa một hoặc nhiều đặc điểm mong muốn có thể được sử dụng làm đích để đưa vào các đặc điểm khác bằng biến nạp liên tiếp. Các trình tự nucleotit bổ sung có thể được đưa vào đồng thời trong phương pháp đồng biến nạp với trình tự nucleotit, phân tử axit nucleic, cấu trúc axit nucleic, hoặc chế phẩm theo sáng chế, tạo ra bởi kết hợp bất kỳ của cat-xet biểu hiện. Ví dụ, nếu hai trình tự nucleotit sẽ được đưa vào, chúng có thể kết hợp trong các tổ hợp riêng rẽ (*trans*) hoặc có thể kết hợp trong cùng tổ hợp (*cis*). Biểu hiện của polynucleotit có thể dẫn đến nhờ vùng khởi động giống nhau hoặc vùng khởi động khác nhau. Còn thấy rằng polynucleotit có thể chồng lên ở vị trí bộ gen mong muốn sử dụng hệ tái tổ hợp đặc hiệu tại chỗ. *Xem, ví dụ*, Công bố đơn sáng chế quốc tế số WO 99/25821; WO 99/25854; WO 99/25840; WO 99/25855 và WO 99/25853.

Cat-xet biểu hiện có thể còn bao gồm trình tự mã hóa bổ sung cho một hoặc nhiều polypeptit hoặc phân tử ARN chuỗi đôi (dsARN) có lợi cho các đặc điểm nông học mà cơ bản có lợi cho công ty sản xuất hạt, người trồng cây hoặc người xử lý hạt. Polypeptit được mong muốn có thể là bất kỳ polypeptit mã hóa bởi trình tự nucleotit mong muốn. Các ví dụ không giới hạn của polypeptit có lợi mà thích hợp để tạo ra ở cây bao gồm các loại tạo ra các đặc điểm quan trọng về nông học như sức kháng chất diệt cỏ (đôi khi còn được gọi là "khả năng chịu chất diệt cỏ"), sức kháng virus, sức kháng thể sinh bệnh vi khuẩn, sức kháng côn trùng, sức kháng giun tròn, hoặc sức kháng nấm. *Xem, ví dụ*, Sáng chế Mỹ số 5,569,823; 5,304,730; 5,495,071; 6,329,504; và 6,337,431. Polypeptit còn có thể là nhân tố làm tăng sức sống của cây hoặc sản lượng của cây (bao gồm các đặc điểm cho phép cây phát triển ở các nhiệt độ, điều kiện đất và mức ánh sáng và lượng mưa khác nhau), hoặc nhân tố cho phép xác định cây biểu hiện đặc điểm được mong muốn (*ví dụ*, chỉ thị chọn lọc, màu bao hạt, v.v.). Các polypeptit được mong muốn khác nhau, cũng như phương pháp để đưa các polypeptit này vào cây trồng, được mô tả, ví dụ, trong Sáng chế Mỹ số 4,761,373; 4,769,061; 4,810,648; 4,940,835; 4,975,374; 5,013,659; 5,162,602; 5,276,268; 5,304,730; 5,495,071; 5,554,798; 5,561,236;

5,569,823; 5,767,366; 5,879,903, 5,928,937; 6,084,155; 6,329,504 và 6,337,431; cũng như Công bố đơn sáng chế Mỹ số 2001/0016956. *Cũng xem*, trên website tại địa chỉ [lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt /](http://lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/).

Polynucleotit tạo ra đề kháng/khả năng chịu đựng chất diệt cỏ ức chế điểm phát triển hoặc mô phân sinh, như imidazalinon hoặc sulfonilurea có thể thích hợp theo một số phương án theo sáng chế. Polynucleotit ví dụ trong loại này mã hóa đối với enzym ALS và AHAS đột biến như mô tả, ví dụ, trong Sáng chế Mỹ số 5,767,366 và 5,928,937. Sáng chế Mỹ số 4,761,373 và 5,013,659 nhắm đến cây chịu các chất diệt cỏ imidazalinon hoặc sulfonamid. Sáng chế Mỹ số 4,975,374 đề cập đến tế bào của cây và cây chứa axit nucleic mã hóa glutamin synthetaza (GS) đột biến kháng sự ức chế bởi chất diệt cỏ đã biết là ức chế GS, ví dụ, phosphinothrixin và methionin sulfoximin. Sáng chế Mỹ số 5,162,602 bộc lộ cây kháng sự ức chế bởi chất diệt cỏ xyclohexanedion và aryloxyphenoxypropanoic axit. Sức đề kháng được tạo ra bởi axetyl coenzym A cacboxylaza (ACCase) được biến đổi.

Các polypeptit được mã hóa bởi các trình tự nucleotit giúp kháng glyphosat cũng thích hợp với sáng chế. *Xem*, ví dụ, Sáng chế Mỹ số 4,940,835 và Sáng chế Mỹ số 4,769,061. Sáng chế Mỹ số 5,554,798 bộc lộ cây ngô chuyển gen kháng glyphosat, mà sức đề kháng được tạo ra nhờ gen 5-enolpyruvyl-3-phosphoshikimat (EPSP) synthaza biến đổi.

Polynucleotit mã hóa đối với hợp chất kháng phosphono như glufosinat amoni hoặc phosphinothrixin, và pyridinoxy hoặc phenoxy propionic axit và xyclohexon cũng là thích hợp. *Xem*, Công bố đơn châu Âu 0 242 246. *Cũng xem*, Sáng chế Mỹ số 5,879,903, 5,276,268 và 5,561,236.

Polynucleotit thích hợp khác bao gồm các polynucleotit mã hóa đề kháng chất diệt cỏ ức chế quang hợp, như triazin và benzonitril (nitrilaza) *Xem*, Sáng chế Mỹ số 4,810,648. Polynucleotit thích hợp khác mã hóa đề kháng chất diệt cỏ bao gồm các polynucleotit mã hóa đề kháng 2,2-dichloropropionic axit, sethoxydim, haloxyfop, chất diệt cỏ imidazolinon, chất diệt cỏ sulfonilurea, chất diệt cỏ triazolopyrimidin, chất diệt cỏ s-triazin và bromoxynil. Cũng thích hợp là polynucleotit tạo ra khả năng kháng protox enzym, hoặc mã tạo ra khả năng kháng bệnh thực vật tăng cường; sức chịu các điều kiện môi trường bất lợi tăng cường (các ức chế vô sinh) bao gồm nhưng không giới hạn ở

hạn hán, quá lạnh, quá nóng, hoặc độ muối trong đất vượt quá hoặc tính axit hoặc tính kiềm cao; và các thay đổi trong kiến trúc hoặc phát triển của cây trồng, bao gồm các thay đổi về thời gian phát triển. *Xem*, ví dụ, Công bố đơn sáng chế Mỹ số 2001/0016956 và Sáng chế Mỹ số 6,084,155.

Polynucleotit thích hợp bổ sung bao gồm các polynucleotit mã hóa các polypeptit diệt loài gây hại (ví dụ, diệt côn trùng). Các polypeptit có thể được tạo ra với lượng đủ để kiểm soát, ví dụ, các côn trùng gây hại (ví dụ, lượng kiểm soát côn trùng). Cần hiểu rằng lượng tạo thành của polypeptit diệt loài gây hại ở cây cần để kiểm soát côn trùng hoặc loài gây hại khác có thể thay đổi phụ thuộc vào các yếu tố cây trồng, kiểu loài gây hại, yếu tố môi trường và tương tự. Polynucleotit hữu dụng để kháng côn trùng hoặc loài gây hại bổ sung bao gồm, ví dụ, polypeptit mã hóa độc tố được xác định trong sinh vật *Bacillus*. Polynucleotit bao gồm trình tự nucleotit mã hóa protein Cry *Bacillus thuringiensis* (*Bt*) từ một vài phân loài được tách dòng và các dòng vô tính tái tổ hợp được thấy là độc đối với ấu trùng của côn trùng bộ cánh vảy, bộ hai cánh và bộ cánh cứng. Các ví dụ của protein *Bt* diệt côn trùng này bao gồm protein Cry như *Cry1Aa*, *Cry1Ab*, *Cry1Ac*, *Cry1B*, *Cry1C*, *Cry1D*, *Cry1Ea*, *Cry1Fa*, *Cry3A*, *Cry9A*, *Cry9B*, *Cry9C*, và tương tự, cũng như protein diệt côn trùng sinh dưỡng như *Vip1*, *Vip2*, *Vip3*, và tương tự. Danh sách đầy đủ của các protein dẫn xuất từ *Bt* có thể tìm thấy trên trang web toàn cầu ở Cơ sở dữ liệu Danh pháp Độc tố *Bacillus thuringiensis* được duy trì bởi Đại học Sussex (*cũng xem*, Crickmore và cộng sự (1998) *Microbiol. Mol. Biol. Rev.* 62:807-813).

Polypeptit mà thích hợp để sản xuất ở cây còn bao gồm các loại mà cải thiện hoặc tạo điều kiện khác sự chuyển đổi cây được thu hoạch hoặc bộ phận của cây thành sản phẩm hữu dụng về mặt kinh tế, bao gồm, ví dụ, tăng hoặc biến đổi hàm lượng carbohydrat hoặc phân phối, cải thiện đặc điểm lên men, tăng hàm lượng dầu, tăng hàm lượng protein, cải thiện khả năng tiêu hóa, và tăng hàm lượng chế phẩm hỗ trợ dinh dưỡng, ví dụ, tăng hàm lượng phytosterol, tăng hàm lượng tocopherol, tăng hàm lượng stanol hoặc tăng hàm lượng vitamin. Các polypeptit được mong muốn còn bao gồm, ví dụ, polypeptit tạo ra hoặc góp phần làm giảm hàm lượng thành phần không mong muốn trong mùa vụ thu hoạch, ví dụ, axit phytic, hoặc enzym làm giảm đường. Thuật ngữ "tạo ra" hoặc "góp phần" là nhằm chỉ polypeptit được mong muốn có thể đóng góp trực tiếp

hoặc gián tiếp vào sự tồn tại của đặc điểm được mong muốn (ví dụ, làm tăng thoái biến xenluloza bằng cách sử dụng enzyme xenlulaza khác loại).

Theo một số phương án, polypeptit góp phần vào khả năng tiêu hóa cải thiện đối với thức ăn hoặc sự cho ăn. Xylanaza là enzym hemixenluloza cải thiện sự hủy hoại thành tế bào của cây, dẫn đến động vật sử dụng tốt hơn dinh dưỡng thực vật. Điều này làm tăng tốc độ phát triển và chuyển hóa thức ăn. Tương tự, độ nhớt của thức ăn chứa xylan có thể giảm. Sản xuất các xylanaza khác loài trong tế bào của cây còn có thể giúp chuyển hóa lignoxenluloza thành đường lên men được dùng trong sản xuất công nghiệp.

Nhiều xylanaza từ vi sinh vật nấm và vi khuẩn được xác định và đặc trưng (xem, ví dụ, Sáng chế Mỹ số 5,437,992; Coughlin và cộng sự (1993) "Proceedings of the Second TRICEL Symposium on *Trichoderma reesei* Xenluloza và Other Hydrolases" Espoo; Souminen và Reinikainen, eds. (1993) *Foundation for Biotechnical and Industrial Fermentation Research* 8:125-135; Công bố đơn sáng chế Mỹ số 2005/0208178; và Công bố đơn PCT số WO 03/16654). Cụ thể là, ba xylanaza cụ thể (XYL-I, XYL-II, và XYL-III) được xác định trong *T. reesei* (Tenkanen và cộng sự (1992) *Enzyme Microb. Technol.* 14:566; Torronen và cộng sự (1992) *Bio/Technology* 10:1461; và Xu và cộng sự (1998) *Appl. Microbiol. Biotechnol.* 49:718).

Theo các phương án khác, polypeptit hữu dụng cho sáng chế có thể là enzym thoái biến polysaccharit. Cây theo sáng chế tạo ra enzym này có thể hữu dụng cho tạo ra, ví dụ, thức ăn chăn nuôi lên men để xử lý sinh học. Theo một số phương án, enzym hữu dụng cho quá trình lên men bao gồm alpha amylaza, proteaza, pullulanaza, isoamylaza, xenlulaza, hemixenlulaza, xylanaza, xyclodextrin glycotransferaza, lipaza, phytaza, laccaza, oxidaza, esteraza, cutinaza, enzym thủy phân bột nghiền hạt và glucoamylaza khác.

Enzym thoái biến polysaccharit bao gồm: enzym thoái biến tinh bột như α -amylaza (EC 3.2.1.1), glucuronidaza (E.C. 3.2.1.131); exo-1,4- α -D glucanaza như amyloglucosidaza và glucoamylaza (EC 3.2.1.3), β -amylaza (EC 3.2.1.2), α -glucosidaza (EC 3.2.1.20), và exo-amylaza khác; enzym cắt nhánh tinh bột, như a) isoamylaza (EC 3.2.1.68), pullulanaza (EC 3.2.1.41), và tương tự; b) xenluloza như exo-1,4-3-xenlobiohydrolaza (EC 3.2.1.91), exo-1,3- β -D-glucanaza (EC 3.2.1.39), β -glucosidaza (EC 3.2.1.21); c) L-arabinaza, như endo-1,5- α -L-arabinaza (EC 3.2.1.99),

α -arabinosidaza (EC 3.2.1.55) và tương tự; d) galactanaza như endo-1,4- β -D-galactanaza (EC 3.2.1.89), endo-1,3- β -D-galactanaza (EC 3.2.1.90), α -galactosidaza (EC 3.2.1.22), β -galactosidaza (EC 3.2.1.23) và tương tự; e) mannanaza, như endo-1,4- β -D-mannanaza (EC 3.2.1.78), β -mannosidaza (EC 3.2.1.25), α -mannosidaza (EC 3.2.1.24) và tương tự; f) xylanaza, như endo-1,4- β -xylanaza (EC 3.2.1.8), β -D-xylosidaza (EC 3.2.1.37), 1,3- β -D-xylanaza, và tương tự; và g) các enzym khác như α -L-fucosidaza (EC 3.2.1.51), α -L-rhamnosidaza (EC 3.2.1.40), levanaza (EC 3.2.1.65), inulanaza (EC 3.2.1.7), và tương tự. Theo một phương án, α -amylaza là α -amylaza tổng hợp, Amy797E, được mô tả là Sáng chế Mỹ số 8,093,453, được kết hợp vào đây bằng cách viện dẫn toàn bộ.

Các enzym khác mà có thể được sử dụng theo sáng chế bao gồm proteaza, như proteaza nấm và vi khuẩn. Proteaza nấm bao gồm, nhưng không giới hạn ở, loại thu được từ *Aspergillus*, *Trichoderma*, *Mucor* và *Rhizopus*, như *A. niger*, *A. awamori*, *A. oryzae* và *M. miehei*. Theo một số phương án, polypeptit theo sáng chế có thể là enzym xenlobiohydrolaza (CBH) (EC 3.2.1.91). Theo một phương án, xenlobiohydrolaza enzym có thể là CBH1 hoặc CBH2.

Các enzym khác hữu dụng với sáng chế bao gồm, nhưng không giới hạn ở, hemixenlulaza, như mannaza và arabinofuranosidaza (EC 3.2.1.55); ligninaza; lipaza (ví dụ, E.C. 3.1.1.3), glucose oxidaza, pectinaza, xylanaza, transglucosidaza, alpha 1,6 glucosidaza (ví dụ, E.C. 3.2.1.20); esteraza như ferulic axit esteraza (EC 3.1.1.73) và axetyl xylan esteraza (EC 3.1.1.72); và cutinaza (ví dụ E.C. 3.1.1.74).

Phân tử ARN chuỗi đôi hữu dụng với sáng chế bao gồm, nhưng không giới hạn ở các loại mà triệt các gen côn trùng đích. Như được sử dụng ở đây các từ "ức chế gen", khi được sử dụng cùng nhau, nhằm đề cập đến bất kỳ trong số các phương pháp đã biết để giảm mức protein tạo ra do phiên mã gen thành mARN và dịch mã sau đó của mARN. Ức chế gen còn có nghĩa là làm giảm biểu hiện protein từ gen hoặc trình tự mã hóa bao gồm ức chế gen sau phiên mã và ức chế dịch mã. Ức chế gen sau phiên mã được trung gian bởi sự tương đồng giữa tất cả hoặc một phần của mARN được phiên mã từ gen hoặc trình tự mã hóa được nhắm đích để triệt và ARN chuỗi đôi tương ứng được dùng để ức chế, và đề cập đến sự giảm cơ bản và đáng kể lượng của mARN có mặt có trong tế bào để liên kết bởi ribosom. ARN được phiên mã có thể theo hướng có nghĩa để làm

cho có hiệu lực được gọi là đồng ức chế, theo hướng đối nghĩa để làm cho có hiệu lực được gọi là ức chế đối nghĩa, hoặc theo cả hai hướng tạo ra dsARN được làm cho có hiệu lực được gọi là ARN can thiệp (ARNi). Ức chế dịch mã được trung gian bởi sự có mặt trong tế bào của dsARN, tác nhân ức chế gen, thể hiện cơ bản tương đồng trình tự với trình tự ADN vùng khởi động hoặc phần bổ sung của chúng để làm cho có hiệu lực cái gọi là ức chế dịch mã vùng khởi động. Ức chế gen có thể có hiệu lực đối với gen cây tự nhiên kết hợp với đặc điểm, ví dụ, để tạo ra cây với mức giảm của protein được mã hóa bởi gen tự nhiên hoặc với mức tăng hoặc giảm của chất chuyển hóa được tác động. Ức chế gen còn có thể có hiệu lực đối với gen đích ở loài gây hại cho cây mà có thể ăn vào bụng hoặc tiếp xúc vật liệu cây chứa tác nhân ức chế gen, được thiết kế đặc biệt để ức chế hoặc ngăn biểu hiện của một hoặc nhiều trình tự tương đồng hoặc bổ sung trong các tế bào của loài gây hại. Các gen này được nhắm đích để triệt có thể mã hóa protein thiết yếu, chức năng dự báo trong đó được chọn từ nhóm bao gồm hình thành cơ, hình thành hormone thiếu niên, quản lý hormone thiếu niên, quản lý và vận chuyển ion, tổng hợp enzym tiêu hóa, bảo dưỡng tiềm năng màng tế bào, sinh tổng hợp axit amin, thoái hóa axit amin, hình thành tinh dịch, tổng hợp pheromon, cảm biến pheromon, hình thành ăngten, hình thành cánh, hình thành chân, phát triển và phân biệt, hình thành trứng, trưởng thành ấu trùng, hình thành enzym tiêu hóa, tổng hợp haemolymph, bảo dưỡng haemolymph, dẫn truyền thần kinh, phân chia tế bào, chuyển hóa năng lượng, hô hấp, và chết tế bào.

Theo một số phương án, sáng chế đề xuất tế bào chủ chuyển gen không phải của người chứa polynucleotit, phân tử axit nucleic, gen khảm, tổ hợp biểu hiện hoặc vector tái tổ hợp theo sáng chế. Tế bào chủ chuyển gen không phải của người có thể bao gồm, nhưng không giới hạn ở, tế bào của cây, tế bào nấm men, tế bào vi khuẩn hoặc tế bào côn trùng. Theo đó, theo một số phương án, sáng chế đề xuất tế bào vi khuẩn được chọn từ giống *Bacillus*, *BreviBacillus*, *Clostridium*, *Xenorhabdus*, *Photorhabdus*, *Pasteuria*, *EscheRichia*, *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Salmonella*, *Pasteurella*, *Xanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Metylophilus*, *Agrobacterium*, *Axetobacter*, *LactoBacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, hoặc *Alcaligenes*. Do đó, ví dụ, làm tác nhân kiểm soát côn trùng sinh học, protein Cry theo sáng chế có thể được tạo ra bằng cách biểu hiện của gen khảm mã hóa protein Cry

theo sáng chế trong tế bào vi khuẩn. Ví dụ, theo một số phương án, tế bào *Bacillus thuringiensis* bao gồm gen khảm theo sáng chế được đề xuất.

Theo các phương án khác, sáng chế đề xuất tế bào cây chuyển gen mà là tế bào cây hai lá mầm hoặc tế bào cây một lá mầm. Theo các phương án bổ sung, tế bào của cây hai lá mầm được chọn từ nhóm bao gồm tế bào cây đậu tương, tế bào hướng dương, tế bào cây cà chua, tế bào cây cải dầu, tế bào cây bông, tế bào cây củ cải đường và tế bào cây thuốc lá. Theo các phương án khác nữa, tế bào một lá mầm được chọn từ nhóm bao gồm tế bào cây lúa mạch, tế bào cây ngô, tế bào cây yến mạch, tế bào cây lúa, tế bào cây lúa miến, tế bào cây mía đường và tế bào cây lúa mỳ. Theo một số phương án, sáng chế đề xuất nhiều tế bào cây hai lá mầm hoặc tế bào cây một lá mầm biểu hiện protein Cry theo sáng chế mà được mã hóa bởi gen khảm theo sáng chế. Theo các phương án khác nhiều trong số các tế bào được đề cập để tạo ra con đường vô bào và phát triển dưới ánh sáng mặt trời tự nhiên.

Theo các phương án khác theo sáng chế, protein Cry diệt côn trùng theo sáng chế được biểu hiện ở sinh vật cao hơn, ví dụ, cây. Trong trường hợp này, cây chuyển gen biểu hiện lượng có hiệu quả của protein diệt côn trùng bảo vệ chính chúng khỏi loài gây hại cho cây như loài gây hại côn trùng. Khi côn trùng bắt đầu phá hoại trên cây chuyển gen này, chúng ăn vào bụng protein Cry diệt côn trùng được biểu hiện. Điều này có thể ngăn cản côn trùng khỏi hút thêm vào mô thực vật hoặc có thể còn gây hại hoặc giết côn trùng. Polynucleotit theo sáng chế được chèn vào tổ hợp biểu hiện, mà sau đó được tích hợp ổn định trong bộ gen của cây. Theo các phương án khác, polynucleotit được bao gồm trong virus tự sao chép không sinh bệnh. Cây được biến nạp theo sáng chế có thể là cây một lá mầm hoặc cây hai lá mầm và bao gồm, nhưng không giới hạn ở, bắp (ngô), đậu tương, lúa, lúa mỳ, lúa mạch, lúa mạch đen, yến mạch, lúa miến, cây kê, hướng dương, cây rum, củ cải đường, bông, mía đường, cây hạt cải dầu, cỏ linh lăng, thuốc lá, lạc, rau, bao gồm, khoai lang, đậu, đậu Hà Lan, rau diếp xoăn, rau diếp, bắp cải, cải hoa, súp lơ, cây củ cải tròn, cà rốt, cà tím, dưa chuột, củ cải, rau bina, khoai tây, cà chua, măng tây, hành, tỏi, dưa, hạt tiêu, cần tây, bí, bí ngô, bí ngòi, quả, bao gồm, táo, lê, mận, quýt, anh đào, đào, xuân đào, mơ, dâu tây, nho, mâm xôi, dâu, dưa, bơ, đu đủ, xoài, chuối, và cây chuyên ngành, như *Arabidopsis*, và cây gỗ như cây lá kim và cây lá bản. Tốt hơn là, cây trồng theo sáng chế là cây trồng như ngô, lúa miến, lúa mỳ, hướng dương,

cà chua, cây họ cải, hạt tiêu, khoai tây, bông, lúa, đậu tương, củ cải đường, mía đường, thuốc lá, lúa mạch, cây hạt cải dầu, và tương tự.

Một khi polynucleotit mong muốn được biến nạp thành loài thực vật cụ thể, có thể nhân giống loài này hoặc chuyển thành các dạng khác của loài tương tự, cụ thể là bao gồm các dạng kinh doanh được, sử dụng các kỹ thuật sinh sản truyền thống.

Polynucleotit theo sáng chế được biểu hiện ở cây chuyển gen, do đó tạo ra sự sinh tổng hợp của protein Cry tương đương, hoặc là ở tiền độc tố hoặc dạng độc tố chín muồi, ở các cây chuyển gen. Bằng cách này, cây chuyển gen có sự bảo vệ năng suất được tăng cường với sự có mặt của áp lực côn trùng được tạo ra. Đối với biểu hiện của chúng trong cây chuyển gen, trình tự nucleotit theo sáng chế có thể cần biến đổi và tối ưu. Mặc dù trong nhiều trường hợp gen vi trùng có thể được biểu hiện ở cây ở các mức cao mà không cần biến đổi, biểu hiện thấp trong cây chuyển gen có thể tạo ra từ trình tự nucleotit của vi khuẩn có bộ ba mã mà không được ưu tiên trong cây trồng. Đã biết trong lĩnh vực là các sinh vật sống có các ưu tiên cụ thể đối với việc sử dụng bộ ba mã, và bộ ba mã của các trình tự nucleotit được mô tả trong sáng chế có thể được thay đổi để thích hợp với ưu tiên của cây, trong khi vẫn giữ các axit amin được mã hóa nhờ đó. Hơn nữa, biểu hiện cao ở cây, ví dụ cây ngô, tốt nhất thu được từ các trình tự mã hóa mà có ít nhất khoảng 35% hàm lượng GC, hoặc ít nhất khoảng 45%, hoặc ít nhất khoảng 50%, hoặc ít nhất khoảng 60%. Trình tự nucleotit của vi khuẩn mà có hàm lượng GC thấp có thể biểu hiện nghèo nàn ở cây do sự tồn tại của mô típ ATTTA mà có thể làm bất ổn định các thông tin, và các mô típ AATAAA mà có thể tạo ra sự polyadenyl không thích hợp. Mặc dù một số các trình tự gen nhất định có thể được biểu hiện một cách thích hợp trong cả các loài cây một lá mầm và hai lá mầm, các trình tự có thể được biến đổi để tính cho các ưu tiên bộ ba mã cụ thể và ưu tiên hàm lượng GC của cây một lá mầm hoặc cây hai lá mầm như các ưu tiên này đã được thể hiện để phân biệt (Murray và cộng sự Nucl. Acids Res. 17:477-498 (1989)). Ngoài ra, trình tự nucleotit được sàng lọc đối với sự tồn tại của các vị trí nối không hợp lệ mà có thể gây ra cắt cụt thông tin. Tất cả các thay đổi cần được thực hiện trong các trình tự nucleotit như các trình tự được mô tả ở trên được thực hiện bằng cách sử dụng các kỹ thuật đã biết về gây đột biến hướng điểm, PCR, và xây dựng gen tổng hợp bằng cách sử dụng phương pháp được mô tả ví dụ trong Sáng chế Mỹ số 5,625,136; 5,500,365 và 6,013,523.

Theo một số phương án, sáng chế đề xuất trình tự mã hóa tổng hợp hoặc polynucleotit tạo ra theo quy trình được bộc lộ trong Sáng chế Mỹ số 5,625,136, được kết hợp ở đây bằng cách viện dẫn. Trong quy trình này, bộ ba mã được ưu tiên của ngô, tức là, bộ ba mã đơn mà mã hóa thường xuyên nhất axit amin ở ngô, được sử dụng. Bộ ba mã được ưu tiên của ngô cho axit amin cụ thể có thể được dẫn xuất, ví dụ, từ các trình tự gen đã biết từ ngô. Ví dụ, việc sử dụng bộ ba mã của ngô cho 28 gen từ cây ngô được tìm thấy trong Murray và cộng sự, *Axit nucleic Research* 17:477-498 (1989), nội dung của nó được kết hợp ở đây bằng cách viện dẫn. Trình tự tổng hợp được lấy làm ví dụ đặc biệt theo sáng chế được thực hiện với bộ ba mã được tối ưu của ngô được đại diện bởi bất kỳ một trong số các SEQ ID NO: 11-30. Cần hiểu rằng bộ ba mã hóa được tối ưu để biểu hiện trong một loài cây sẽ còn thực hiện chức năng trong loài gây khác nhưng có thể không ở cùng mức như loài gây mà bộ ba mã hóa được tối ưu cho nó. Theo cách này, trình tự nucleotit có thể được tối ưu để biểu hiện ở cây bất kỳ. Hiểu rằng tất cả hoặc một phần bất kỳ của trình tự nucleotit có thể được tối ưu hoặc tổng hợp. Tức là, polynucleotit có thể bao gồm trình tự nucleotit mà là trình tự tự nhiên một phần và trình tự tối ưu bộ ba mã hóa một phần.

Để khởi đầu dịch mã hiệu quả, các trình tự liền kề với methionin có thể cần biến đổi. Ví dụ, chúng có thể được biến đổi bằng cách gộp vào các trình tự đã biết là có hiệu quả trong thực vật. Joshi đã đề xuất sự thống nhất thích hợp đối với cây (NAR 15:6643-6653 (1987)) và Clonotech đề xuất chất khởi đầu dịch mã thống nhất khác (1993/1994 catalog, trang 210). Các liên ứng này thích hợp để dùng với các trình tự nucleotit theo sáng chế. Các trình tự được kết hợp vào các cấu trúc bao gồm trình tự nucleotit, lên đến và bao gồm ATG (trong khi bỏ axit amin thứ hai không được biến đổi), hoặc tùy ý lên đến và bao gồm GTC sau ATG (với khả năng thay đổi axit amin thứ hai của gen chuyển).

Trình tự mã hóa protein Cry mới theo sáng chế, hoặc là trình tự tự nhiên của chúng hoặc là trình tự tổng hợp như được mô tả ở trên, có thể được dung hợp hoạt động với nhiều vùng khởi động để biểu hiện ở cây bao gồm vùng khởi động cấu thành, quy nạp, điều chỉnh tạm thời, điều chỉnh phát triển, điều chỉnh hóa học, ưu tiên mô và đặc hiệu mô để điều chế phân tử ADN tái tổ hợp, tức là, gen khảm. Việc lựa chọn vùng khởi động sẽ thay đổi tùy thuộc vào yêu cầu thời gian và không gian, và còn tùy thuộc vào loài đích. Do đó, biểu hiện của trình tự nucleotit theo sáng chế ở lá, cuống hoặc thân, ở tai, ở cụm hoa (ví dụ cụm hoa, chùy, hạt phỉ, v.v.), ở rễ, hoặc cây giống con được ưu

tiên. Trong nhiều trường hợp, tuy nhiên, đòi hỏi bảo vệ chống lại nhiều hơn một loại côn trùng gây hại, và do đó biểu hiện trong nhiều mô là cần thiết. Mặc dù nhiều vùng khởi động từ các cây hai lá mầm đã thể hiện hoạt động trong các cây một lá mầm và ngược lại, lý tưởng là các vùng khởi động của cây hai lá mầm được chọn để biểu hiện trong cây hai lá mầm và các vùng khởi động của cây một lá mầm để biểu hiện trong cây một lá mầm. Tuy nhiên, không có hạn chế đối với nguồn gốc của vùng khởi động được chọn; chỉ cần chúng hoạt động trong việc dẫn hướng biểu hiện của trình tự nucleotit trong tế bào mong muốn.

Vùng khởi động cấu thành thích hợp bao gồm, ví dụ, vùng khởi động CaMV 35S (SEQ ID NO:1546; Odell và cộng sự, *Nature* 313:810-812, 1985); vùng khởi động *Arabidopsis* At6669 (SEQ ID NO:1652; xem Công bố đơn PCT số W004081173A2); Ubi 1 của ngô (Christensen và cộng sự, *Plant Mol. Biol.* 18:675-689, 1992); actin của lúa (McElroy và cộng sự, *Plant Cell* 2:163-171, 1990); pEMU (Last và cộng sự, *Theor. Appl. Genet.* 81:581-588, 1991); CaMV 19S (Nilsson và cộng sự, *Physiol. Plant* 100:456-462, 1997); GOS2 (de Pater và cộng sự, *Plant J* November; 2(6):837-44, 1992); ubiquitin (Christensen và cộng sự, *Plant Mol. Biol.* 18: 675-689, 1992); Xyclophilin của lúa (Bucholz và cộng sự, *Plant Mol Biol.* 25(5):837-43, 1994); Histon H3 của ngô (Lepetit và cộng sự, *Mol. Gen. Genet.* 231: 276-285, 1992); Actin 2 (An và cộng sự, *Plant J.* 10(1):107-121, 1996), vùng khởi động CT2 cấu thành đầu rễ (SEQ ID NO:1535; cũng xem Đơn PCT số IL/2005/000627) và MAS siêu tổng hợp (Ni và cộng sự, *The Plant Journal* 7: 661-76, 1995). Vùng khởi động cấu thành khác bao gồm các loại trong Sáng chế Mỹ Số 5,659,026, 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; và 5,608,142. *Molec. Biol.*

Vùng khởi động đặc hiệu mô hoặc ưu tiên mô hữu dụng để biểu hiện trình tự mã hóa protein *Cry* mới theo sáng chế ở cây trồng, cụ thể là ngô, là các loại mà định hướng biểu hiện trong rễ, ruột cây, lá hoặc phần hoa. Vùng khởi động đặc hiệu mô thích hợp bao gồm, nhưng không giới hạn ở, vùng khởi động đặc hiệu lá [như được mô tả, ví dụ, bởi Yamamoto và cộng sự, *Plant J.* 12:255-265, 1997; Kwon và cộng sự, *Plant Physiol.* 105:357-67, 1994; Yamamoto và cộng sự, *Plant Cell Physiol.* 35:773-778, 1994; Gotor và cộng sự, *Plant J.* 3:509-18, 1993; Orozco và cộng sự, *Plant Mol. Biol.* 23:1129-1138, 1993; và Matsuoka và cộng sự, *Proc. Natl. Acad. Sci. USA* 90:9586-9590, 1993], vùng khởi động ưu tiên hạt [ví dụ, từ gen đặc hiệu hạt (Simon, và cộng sự, *Plant Mol. Biol.* 5.

191, 1985; Scofield, và cộng sự, J. Biol. Chem. 262: 12202, 1987; Baszczynski, và cộng sự, Plant Mol. Biol. 14: 633, 1990), albumin quả hạch Brazil (Pearson' và cộng sự, Plant Mol. Biol. 18: 235-245, 1992), legumin (Ellis, và cộng sự Plant Mol. Biol. 10: 203-214, 1988), Glutelin (*rice*) (Takaiwa, và cộng sự, Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa, và cộng sự, FEBS Letts. 221: 43-47, 1987), Zein (Matzke và cộng sự, Plant Mol Biol, 143).323-32 1990), napA (Stalberg, và cộng sự, Planta 199: 515-519, 1996), SPA lúa mì (Albanietal, Plant Cell, 9: 171-184, 1997), oleosin hướng dương (Cummins, và cộng sự, Plant Mol. Biol. 19: 873-876, 1992)], vùng khởi động đặc hiệu nội nhũ [ví dụ, LMW và HMW lúa mì, glutenin-1 (Mol Gen Genet 216:81-90, 1989; NAR 17:461-2), gliadina, b và g lúa mì (EMB03:1409-15, 1984), vùng khởi động ltrI lúa mạch, hordein B1, C, D lúa mạch (Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996), DOF lúa mạch (Mena và cộng sự, The Plant Journal, 116(1): 53-62, 1998), Biz2 (EP99106056.7), Vùng khởi động tổng hợp (Vicente-Carbajosa và cộng sự, Plant J. 13: 629-640, 1998), prolamin NRP33 của lúa, -globulin Glb-1 của lúa (Wu và cộng sự, Tế bào cây Physiology 39(8) 885-889, 1998), alpha-globulin REB/OHP-1 của lúa (Nakase và cộng sự Plant Mol. Biol. 33: 513-S22, 1997), ADP-glucoza PP của lúa (Trans Res 6:157-68, 1997), họ gen ESR của ngô (Plant J 12:235-46, 1997), gamma-kafirin lúa miễn (Plant Mol. Biol 32:1029-35, 1996)], vùng khởi động đặc hiệu phôi [ví dụ, rice OSH1 (Sato và cộng sự, Proc. Nati. Acad. Sci. USA, 93: 8117-8122), KNOX (Postma-Haarsma of al, Plant Mol. Biol. 39:257-71, 1999), oleosin của lúa (Wu và cộng sự, J. Biochem., 123:386, 1998)], vùng khởi động đặc hiệu hoa [ví dụ, AtPRP4, chalen synthaza (chsA) (Van der Meer, và cộng sự, Plant Mol. Biol. 15, 95-109, 1990), LAT52 (Twel và cộng sự, Mol. Gen Genet. 217:240-245; 1989), apetala-3, mô tái sinh cây [ví dụ, vùng khởi động OsMADS (Đơn sáng chế Mỹ 2007/0006344)].

Trình tự nucleotit theo sáng chế còn có thể được biểu hiện dưới sự điều biến của vùng khởi động mà là điều biến hóa học. Điều này cho phép các protein Cry theo sáng chế được tổng hợp chỉ khi cây trồng được xử lý bằng các chất hóa học cảm ứng. Các ví dụ của công nghệ cảm ứng hóa học của biểu hiện gen này được mô tả chi tiết trong công bố đơn EP 0 332 104 và Sáng chế Mỹ số 5,614,395. Theo một phương án, vùng khởi động điều biến hóa học là vùng khởi động PR-1a của thuốc lá.

Nhóm vùng khởi động khác hữu dụng theo sáng chế là loại mà cảm ứng cuộn. Nhiều vùng khởi động đã được mô tả mà được biểu hiện tại vị trí cuộn và cả ở các vị trí nhiễm bệnh thực vật. Lý tưởng là, vùng khởi động này chỉ có thể có hoạt tính cục bộ tại vị trí côn trùng xâm chiếm, và bằng cách này protein diệt côn trùng chỉ xếp chồng trong các tế bào cần tổng hợp protein diệt côn trùng để diệt loài gây hại côn trùng xâm chiếm. Các ví dụ của vùng khởi động loại này bao gồm các loại được mô tả bởi Stanford và cộng sự. *Mol. Gen. Genet.* 215:200-208 (1989), Xu và cộng sự. *Plant Molec. Biol.* 22:573-588 (1993), Logemann và cộng sự. *Plant Cell* 1:151-158 (1989), Rohrmeier & Lehle, *Plant Molec. Biol.* 22:783-792 (1993), Firek và cộng sự. *Plant Molec. Biol.* 22:129-142 (1993), và Warner và cộng sự. *Plant J.* 3:191-201 (1993).

Các ví dụ không giới hạn của vùng khởi động mà tạo ra mẫu biểu hiện đặc hiệu mô mà hữu dụng theo sáng chế bao gồm đặc hiệu mô xanh lá, đặc hiệu rễ, đặc hiệu thân, hoặc đặc hiệu hoa. Các vùng khởi động thích hợp để biểu hiện trong mô xanh bao gồm nhiều gen điều chỉnh liên quan đến quang tổng hợp và nhiều vùng khởi động được nhân bản từ cả hai loại cây một lá mầm và cây hai lá mầm. Một vùng khởi động này là vùng khởi động PEPC của ngô từ gen phosphoenol cacboxylaza (Hudspeth & Grula, *Plant Molec. Biol.* 12:579-589 (1989)). Vùng khởi động khác để biểu hiện đặc hiệu rễ là loại được mô tả bởi de Framond (*FEBS* 290:103-106 (1991) hoặc Sáng chế Mỹ số 5,466,785). Vùng khởi động khác hữu dụng theo sáng chế là vùng khởi động đặc hiệu thân được mô tả trong Sáng chế Mỹ số 5,625,136, mà dẫn hướng tự nhiên biểu hiện gen *trpA* của ngô.

Bên cạnh việc lựa chọn vùng khởi động thích hợp, các cấu trúc để biểu hiện độc tố diệt côn trùng ở cây trồng cần trình tự kết thúc phiên mã thích hợp để liên kết hoạt động xuôi dòng của trình tự nucleotit khác loại. Một số trình tự kết thúc là đã có và là đã biết trong lĩnh vực (ví dụ tml từ CaMV, E9 từ rbcS). Trình tự kết thúc bất kỳ hiện có đã biết là có chức năng trong thực vật có thể được sử dụng trong ngữ cảnh của sáng chế.

Nhiều trình tự khác có thể được kết hợp vào tổ hợp có thể biểu hiện mô tả trong sáng chế. Các trình tự này bao gồm các trình tự thể hiện tăng cường biểu hiện như các trình tự intron (ví dụ từ *Adhl* và *bronzel*) và trình tự dẫn dắt vi rút (ví dụ từ TMV, MCMV và AMV).

Có thể ưu tiên là nhắm đích biểu hiện trình tự nucleotit theo sáng chế đến định vị tế bào khác nhau trong cây trồng. Trong một số trường hợp, sự định vị trong xytosol có thể mong muốn, trong khi đó trong các trường hợp khác, sự định vị trong một số hạt cơ quan dưới mức tế bào có thể được ưu tiên. Cơ chế bất kỳ để nhắm đích sản phẩm gen, ví dụ, ở cây, có thể được dùng để thực hiện sáng chế, và các cơ chế này được biết là tồn tại ở cây và các trình tự kiểm soát chức năng của các cơ chế này đã được mô tả cụ thể ở mức nào đó. Trình tự được đặc trưng mà tạo ra nhắm đích sản phẩm gen đến Trình tự amin cuối của ngăn tế bào khác có thể chịu trách nhiệm nhắm đích protein có lợi đến ngăn tế bào bất kỳ, như, không bào, ty thể, peroxisome, thể protein, lưới nội bào, lục lạp, hạt tinh bột, hạt bột, hạt trong hoặc thành tế bào của cây (ví dụ Unger và cộng sự *Plant Molec. Biol.* 13: 411-418 (1989); Rogers và cộng sự (1985) *Proc. Natl. Acad. Sci. USA* 82: 6512-651; Sáng chế Mỹ số 7,102,057; WO 2005/096704, tất cả chúng được kết hợp tại đây bằng cách viện dẫn. Tùy ý, trình tự tín hiệu có thể là trình tự tín hiệu đầu N từ sắn, trình tự tín hiệu đầu N từ gamma-zein, vùng liên kết tinh bột, vùng liên kết tinh bột đầu C, trình tự nhắm đích lục lạp, mà nhập protein chín muối vào lục lạp (Comai và cộng sự (1988) *J. Biol. Chem.* 263: 15104-15109; van den Broeck, và cộng sự (1985) *Nature* 313: 358-363; Sáng chế Mỹ số 5,639,949) hoặc trình tự tín hiệu bài tiết từ tế bào aleuron (Koehler & Ho, *Plant Cell* 2: 769-783 (1990)). Ngoài ra, trình tự amin cuối liên kết với trình tự cacboxy cuối chịu trách nhiệm nhắm đích không bào của sản phẩm gen (Shinshi và cộng sự (1990) *Plant Molec. Biol.* 14: 357-368). Theo một phương án, trình tự tín hiệu được chọn bao gồm điểm phân cắt đã biết, và dung hợp được cấu trúc xét đến các axit amin bất kỳ sau (các) vị trí phân cắt mà cần cho phân cắt. Trong một số trường hợp yêu cầu này có thể được thỏa mãn bằng cách bổ sung số lượng nhỏ các axit amin giữa vị trí phân cắt và ATG gen chuyển hoặc, tùy ý, thay đổi vị trí của một số axit amin trong trình tự gen chuyển. Kỹ thuật cấu trúc này là đã biết trong lĩnh vực và có thể áp dụng tương đương cho các ngăn tế bào bất kỳ.

Sẽ hiểu rằng các cơ chế được mô tả ở trên để nhắm đích tế bào có thể được sử dụng không chỉ cùng với vùng khởi động cùng nguồn của chúng, mà còn cùng với vùng khởi động khác loại sao cho ảnh hưởng đến mục tiêu nhắm đích tế bào cụ thể dưới sự điều biến phiên mã của vùng khởi động mà có mẫu biểu hiện khác của vùng khởi động từ đó tạo ra tín hiệu nhắm đích.

Biến nạp cây

Quy trình để biến nạp cây là đã biết và quen thuộc trong lĩnh vực và được mô tả qua các tài liệu tham khảo. Các ví dụ không giới hạn của phương pháp để biến nạp cây bao gồm biến nạp qua phân phối axit nucleic trung gian vi khuẩn (ví dụ, qua *Agrobacterium*), phân phối axit nucleic trung gian virus, phân phối axit nucleic trung gian sợi tinh thể silic cacbua hoặc axit nucleic, phân phối axit nucleic trung gian liposome, vi tiêm, bắn phá vi hạt, biến nạp trung gian canxi-phosphat, biến nạp trung gian xyclodextrin, điện di, biến nạp trung gian hạt nano, nghiền bằng sóng siêu âm, thẩm lọc, hấp thụ axit nucleic trung gian PEG, cũng như cơ chế điện, hóa, lý (cơ học) hoặc sinh học bất kỳ khác mà tạo ra việc đưa axit nucleic vào tế bào cây, bao gồm kết hợp bất kỳ của chúng. Các hướng dẫn chung về các phương pháp biến nạp cây khác nhau là đã biết trong lĩnh vực này bao gồm Miki và cộng sự. ("Procedures for Introducing Foreign DNA into Plants" trong *Methods in Plant Molecular Biology and Biotechnology*, Glick, B. R. và Thompson, J. E., Eds. (CRC Press, Inc., Boca Raton, 1993), trang 67-88) và Rakowoczy-Trojanowska (*Cell. Mol. Biol. Lett.* 7:849-858 (2002)).

Đối với biến nạp trung gian *Agrobacterium*, các vector hai kỳ hoặc vector mang ít nhất một trình tự viền T-ADN thích hợp, trong khi đó đối với chuyển gen trực tiếp (ví dụ, thả bom hạt và tương tự) vector bất kỳ thích hợp và ADN thẳng chỉ chứa cấu trúc được mong muốn có thể được sử dụng. Trong trường hợp chuyển gen trực tiếp, biến nạp với các loài ADN đơn hoặc đồng biến nạp có thể được sử dụng (Schocher và cộng sự, *Biotechnology* 4:1093- 1096 (1986)). Đối với cả chuyển gen trực tiếp và chuyển trung gian *Agrobacterium*, biến nạp thường (nhưng không nhất thiết) được hấp thụ với dấu chọn lọc có thể là lựa chọn dương (Phosphomannozia Isomerase), tạo ra khả năng kháng kháng sinh (kanamycin, hygromycin hoặc metotrexat) hoặc chất diệt cỏ (glyphosat hoặc glufosinat). Tuy nhiên, việc lựa chọn dấu chọn lọc không thuộc phạm vi sáng chế.

Biến nạp trung gian *Agrobacterium* là phương pháp biến nạp cây thường dùng nhờ tính hiệu quả biến nạp cao của nó và nhờ tính hữu ích rộng của nó với nhiều loài khác nhau. Biến nạp trung gian *Agrobacterium* thường tham gia vào việc vận chuyển vector nhị thể mang ADN bên ngoài có lợi đến chủng *Agrobacterium* thích hợp mà phụ thuộc vào sự bổ sung gen *vir* được mang bởi chủng *Agrobacterium* chủ hoặc là trên plasmid Ti đồng nội trú hoặc nhiễm sắc thể (Uknes và cộng sự (1993) *Plant Cell* 5:159-169). Việc chuyển vector nhị thể tái tổ hợp vào *Agrobacterium* có thể được thực hiện bằng quy trình giao phối ba bên sử dụng *Escherichia coli* mang vector nhị thể tái tổ hợp, chủng *E. coli*

hỗ trợ mang plasmit mà có khả năng di chuyển vectơ nhị thể tái tổ hợp đến chủng *Agrobacterium* đích. Tùy ý, vectơ nhị thể tái tổ hợp có thể được vận chuyển đến *Agrobacterium* bằng cách biến nạp axit nucleic (Höfgen & Willmitzer (1988) *Nucleic Acids Res.* 16:9877).

Cây hai lá mầm cũng như cây một lá mầm có thể được biến nạp bằng cách sử dụng *Agrobacterium*. Phương pháp để biến nạp trung gian *Agrobacterium* của lúa bao gồm phương pháp biến nạp lúa đã biết, như phương pháp được mô tả trong bất kỳ trong số các tài liệu sau đây: Đơn sáng chế châu Âu EP 1198985 A1, Aldemita và Hodges (Planta 199: 612-617, 1996); Chan và cộng sự (Plant Mol Biol 22 (3): 491-506, 1993), Hiei và cộng sự (Plant J 6 (2): 271-282, 1994), nội dung của nó được kết hợp ở đây bằng cách viện dẫn toàn bộ. Trong trường hợp biến nạp ngô, phương pháp được ưu tiên là như được mô tả trong hoặc là Ishida và cộng sự (Nat. Biotechnol 14(6): 745-50, 1996) hoặc Frame và cộng sự (Plant Physiol 129(1): 13-22, 2002), nội dung của nó được kết hợp ở đây bằng cách viện dẫn toàn bộ. Phương pháp này còn được mô tả ví dụ trong B. Jené và cộng sự, Techniques for Gene Transfer, trong: Transgenic Plants, Tập 1, Engineering and Utilization, eds. S. D. Kung và R. Wu, Academic Press (1993) 128-143 và trong Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Axit nucleic hoặc cấu trúc để biểu hiện tốt hơn là được tách dòng thành vectơ, mà thích hợp để biến nạp *Agrobacterium tumefaciens*, ví dụ pBin19 (Bevan và cộng sự, Nucl. Acids Res. 12 (1984) 8711). *Agrobacteria* được biến nạp bằng vectơ này sau đó có thể được dùng theo cách đã biết để biến nạp cây, như cây dùng làm mô hình, như *Arabidopsis* hoặc cây trồng như, ví dụ, cây thuốc lá, ví dụ bằng cách ngâm lá bị thâm hoặc lá bị chặt trong dung dịch *agrobacteria* và sau đó nuôi cấy chúng trong môi trường thích hợp. Biến nạp cây bằng *Agrobacterium tumefaciens* được mô tả, ví dụ, bởi Hagen và Willmitzer trong Nucl. Acid Res. (1988) 16, 9877 hoặc đã biết không kể những cái khác từ F. F. White, Vectors for Gene Transfer in Higher Plants; trong Transgenic Plants, Tập 1, Engineering and Utilization, eds. S. D. Kung và R. Wu, Academic Press, 1993, pp. 15-38.

Biến nạp của cây bằng *Agrobacterium* tái tổ hợp thường liên quan đến nuôi cấy đồng thời *Agrobacterium* với các cấy mô từ cây trồng và theo phương pháp đã biết trong lĩnh vực. Mô biến nạp được tạo ra trong môi trường chọn lọc mang chất kháng sinh hoặc dầu kháng thuốc diệt cỏ giữa các viên T-ADN lặp thể hai kỳ.

Như được bàn đến ở trên, phương pháp khác để biến nạp cây, bộ phận của cây và tế bào của cây liên quan đến việc đẩy các hạt hoạt tính sinh học hoặc trơ ở mô và tế bào cây. *Xem, ví dụ*, Sáng chế Mỹ số 4,945,050; 5,036,006 và 5,100,792. Thông thường, phương pháp liên quan đến việc đẩy các hạt hoạt tính sinh học hoặc trơ ở tế bào của cây dưới các điều kiện hiệu quả để xâm nhập mặt ngoài của tế bào và tạo ra kết hợp bên trong nó. Khi các hạt trơ được sử dụng, vectơ có thể được đưa vào tế bào bằng cách phủ các hạt bằng các vectơ chứa axit nucleic được mong muốn. Tùy ý, tế bào hoặc thể bào có thể bao quanh bởi vectơ mà vectơ mang vào tế bào bằng cách theo hạt. Các hạt hoạt tính sinh học (*ví dụ*, tế bào nấm men khô, vi khuẩn khô hoặc vật ăn vi khuẩn, mỗi hạt chứa một hoặc nhiều axit nucleic cần đưa vào) còn có thể được đẩy vào mô thực vật.

Theo các phương án khác, polynucleotit theo sáng chế có thể được biến nạp trực tiếp vào bộ gen lặp thể. Lợi ích chủ yếu của biến nạp lặp thể là các lặp thể thường có khả năng biểu hiện gen vi khuẩn mà không làm biến đổi cơ bản, và các lặp thể có khả năng biểu hiện khung đọc mở phức dưới kiểm soát của vùng khởi động đơn. Kỹ thuật biến nạp lặp thể được mô tả sâu trong Sáng chế Mỹ số 5,451,513, 5,545,817, và 5,545,818, đơn Đơn PCT số WO 95/16783, và trong McBride và cộng sự (1994) *Proc. Natl. Acad. Sci. USA* 91, 7301-7305. Kỹ thuật cơ sở để biến nạp lặp liên quan đến việc đưa các vùng ADN lặp thể đã tách dòng giới hạn hai đầu một chỉ thị chọn lọc cùng với gen được mong muốn vào mô đích thích hợp, ví dụ, dùng súng bắn gen hoặc biến nạp tế bào trần (như canxi clorua hoặc biến nạp trung gian PEG). Các vùng giới hạn hai đầu 1-1,5kb, được gọi là các trình tự đích, tạo điều kiện cho tái tổ hợp tương đồng với bộ gen lặp thể và do đó cho phép thay thế hoặc sửa đổi các vùng cụ thể của hệ gen lặp thể. Ban đầu, các đột biến điểm trong các gen 16S rRNA và rps12 lặp lại tạo ra sức kháng đối với spectinomycin hoặc streptomycin có thể được sử dụng làm chỉ thị chọn lọc để biến nạp (Svab, Z., Hajdukiewicz, P., và Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., và Maliga, P. (1992) *Plant Cell* 4, 39-45). Sự có mặt các vị trí tách dòng giữa các chỉ thị này cho phép tạo ra vectơ định hướng lặp thể để đưa vào các gen lạ (Staub, J.M., và Maliga, P. (1993) *EMBO J.* 12, 601-606). Sự tăng cơ bản về tần suất biến nạp có thể thu được bằng cách thay thế rARN lặn hoặc gen kháng kháng sinh r-protein với chỉ thị chọn lọc trội, *genaadA* vi khuẩn mã hóa enzym giải độc spectinomycin aminoglycosit- 3'- adenylyltransferaza (Svab, Z., và Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917). Trước đây, chỉ dấu này được sử dụng thành

công để biến nạp tần suất cao của hệ gen lặp thể của tảo xanh *Chlamydomonas reinhardtii* (Goldschmidt-Clermont, M. (1991) *Nucl. Acids Res.* 19:4083-4089). Các chỉ thị chọn lọc khác hữu dụng cho biến nạp lặp thể đã được biết trong lĩnh vực và bao gồm trong phạm vi của sáng chế. Thông thường, cần đến khoảng 15-20 chu kỳ phân chia tế bào sau khi biến nạp để đạt được trạng thái đồng nhất. Biểu hiện lặp thể, mà trong đó các gen được chèn bởi tái tổ hợp tương đồng vào toàn bộ vài nghìn bản sao của bộ gen lặp thể vòng có trong mỗi tế bào của cây, lợi dụng số lượng rất lớn bản sao lợi thế hơn các gen được biểu hiện trong nhân để cho phép các mức biểu hiện có thể dễ dàng vượt quá 10% tổng protein thực vật hoà tan. Theo một phương án, polynucleotit theo sáng chế có thể được chèn vào vectơ định hướng lặp thể và được biến nạp vào bộ gen lặp thể của vật chủ thực vật mong muốn. Do đó, thực vật đồng nhất đối với các bộ gen lặp thể chứa trình tự nucleotit của sáng chế có thể thu được, mà có khả năng biểu hiện cao polynucleotit.

Phương pháp chọn lọc cây được biến nạp, chuyển gen, tế bào cây hoặc nuôi cấy mô cây là đã biết trong lĩnh vực và có thể được dùng trong phương pháp theo sáng chế được đề xuất ở đây. Ví dụ, vectơ tái tổ hợp theo sáng chế có thể còn bao gồm cat-xet biểu hiện bao gồm trình tự nucleotit cho chỉ thị chọn lọc, mà có thể được dùng để chọn lọc cây được biến nạp, bộ phận của cây hoặc tế bào cây. Như được sử dụng ở đây, "chỉ thị chọn lọc" nghĩa là trình tự nucleotit mà khi được biểu hiện tạo ra phenotyp khác biệt cho cây, bộ phận của cây hoặc tế bào cây biểu hiện chỉ dấu và do đó cho phép cây được biến nạp này, bộ phận của cây hoặc tế bào cây phân biệt được với cây không có chỉ dấu. Trình tự nucleotit có thể mã hóa hoặc là chỉ thị chọn lọc hoặc sàng lọc, phụ thuộc và chỉ dấu tạo ra đặc điểm có thể được chọn lọc bằng phương pháp hóa học, như bằng cách sử dụng chất chọn lọc (ví dụ, chất kháng sinh, chất diệt cỏ, hoặc tương tự), hoặc liệu chỉ dấu đơn giản là đặc điểm mà có thể xác định thông qua việc quan sát hoặc thử, như bằng cách sàng lọc (ví dụ, đặc điểm nơi cư trú R). Đương nhiên, nhiều ví dụ về chỉ thị chọn lọc thích hợp là đã biết trong lĩnh vực và có thể được sử dụng trong tổ hợp có thể biểu hiện mô tả ở đây.

Các ví dụ của chỉ thị chọn lọc bao gồm, nhưng không giới hạn ở, trình tự nucleotit mã hóa *neo* hoặc *nptII*, mà tạo ra sức kháng đối với kanamycin, G418, và tương tự (Potrykus và cộng sự (1985) *Mol. Gen. Genet.* 199:183-188); trình tự nucleotit mã hóa *bar*, mà tạo ra sức kháng đối với phosphinothrixin; trình tự nucleotit mã hóa 5-

enolpyruvylshikimat-3-phosphat (EPSP) synthaza biến đổi, mà tạo ra sức kháng đối với glyphosat (Hinchee và cộng sự (1988) *Biotech.* 6:915-922); trình tự nucleotit mã hóa nitrilaza như *bnx* từ *Klebsiella ozaenae* mà tạo ra sức kháng đối với bromoxynil (Stalker và cộng sự (1988) *Science* 242:419-423); trình tự nucleotit mã hóa axetolactat synthaza biến đổi (ALS) mà tạo ra sức kháng đối với imidazolinon, sulfonylurê hoặc chất hóa học ức chế ALS khác (Đơn sáng chế châu Âu số 154204); trình tự nucleotit mã hóa dihydrofolat reductaza kháng metotrexat (DHFR) (Thillet và cộng sự (1988) *J. Biol. Chem.* 263:12500-12508); trình tự nucleotit mã hóa dalapon dehalogenaza mà tạo ra sức kháng đối với dalapon; trình tự nucleotit mã hóa mannoza-6-phosphat isomeraza (còn gọi là phosphomannoza isomeraza (PMI)) mà tạo ra khả năng chuyển hóa mannoza (Sáng chế Mỹ số 5,767,378 và 5,994,629); trình tự nucleotit mã hóa anthranilat synthaza biến đổi mà tạo ra sức kháng đối với 5-metyl tryptophan; hoặc trình tự nucleotit mã hóa *hph* mà tạo ra sức kháng đối với hygromyxin. Chuyên gia trong lĩnh vực có khả năng chọn chỉ thị chọn lọc thích hợp để dùng trong tổ hợp biểu hiện của sáng chế.

Các chỉ thị chọn lọc bổ sung bao gồm, nhưng không giới hạn ở, trình tự nucleotit mã hóa β -glucuronidaza hoặc *uidA* (GUS) mà mã hóa enzyme mà các nền chất sinh màu khác nhau cho chúng là đã biết; trình tự nucleotit điểm R mà mã hóa sản phẩm mà điều biến việc tạo ra chất nhuộm anthoxyanin (màu đỏ) trong mô thực vật (Dellaporta và cộng sự, "Molecular cloning of the maize R-nj allele by transposon-tagging with Ac" 263-282 trong: *Cấu trúc và chức năng nhiễm sắc thể: Impact of New Concepts*, 18th Stadler Genetics Symposium (Gustafson & Appels eds., Plenum Press 1988)); trình tự nucleotit mã hóa β -lactamaza, enzyme mà nhiều chất sinh màu khác nhau cho nó là đã biết (ví dụ, PADAC, a cephalosporin sinh màu) (Sutcliffe (1978) *Proc. Natl. Acad. Sci. USA* 75:3737-3741); trình tự nucleotit mã hóa *xylE* mà mã hóa catechol dioxygenaza (Zukowsky và cộng sự (1983) *Proc. Natl. Acad. Sci. USA* 80:1101-1105); trình tự nucleotit mã hóa tyrosinaza, enzyme có khả năng oxy hóa tyrosin thành DOPA và dopaquinon, mà đến lượt nó làm đặc để tạo thành melanin (Katz và cộng sự (1983) *J. Gen. Microbiol.* 129:2703-2714); trình tự nucleotit mã hóa β -galactosidaza, enzyme mà có chất sinh màu cho nó; trình tự nucleotit mã hóa luciferaza (*lux*) mà cho phép phát hiện phát quang sinh học (Ow và cộng sự (1986) *Science* 234:856-859); trình tự nucleotit mã hóa aequorin mà có thể được sử dụng trong phát hiện phát quang sinh học nhạy canxi (Prasher và cộng sự (1985) *Biochem. Biophys. Res. Comm.* 126:1259-1268); hoặc trình

tự nucleotit mã hóa protein huỳnh quang xanh lá (Niedz và cộng sự (1995) *Plant Cell Reports* 14:403-406). Chuyên gia trong lĩnh vực có khả năng chọn chỉ thị chọn lọc thích hợp để dùng trong tổ hợp biểu hiện của sáng chế.

Ngoài ra, như đã biết trong lĩnh vực, cây chuyển gen nguyên vẹn có thể được tái sinh từ tế bào cây được biến nạp, nuôi cấy mô cây hoặc thể nguyên sinh được nuôi cấy sử dụng bất kỳ trong số các kỹ thuật đã biết. Tái tạo cây từ tế bào cây, nuôi cấy mô cây hoặc thể nguyên sinh được nuôi cấy được mô tả, ví dụ, trong Evans và cộng sự (*Handbook of Plant Cell Cultures*, Tập 1, MacMilan Publishing Co. New York (1983)); và Vasil I. R. (ed.) (*Cell Culture and Somatic Cell Genetics of Plants*, Acad. Press, Orlando, Tập I (1984), và Tập II (1986)).

Ngoài ra, đặc điểm gen được thiết kế gen vào hạt chuyển gen và cây, bộ phận của cây, hoặc tế bào cây theo sáng chế được mô tả ở trên có thể được đi qua trên bằng cách sinh sản hữu tính hoặc tăng trưởng sinh dưỡng và do đó có thể được duy trì và nhân giống ở cây con. Nhìn chung, việc duy trì và nhân giống làm việc sử dụng các phương pháp nông nghiệp đã biết phát triển đến mục đích cụ thể như thu hoạch, gieo hạt hoặc cày bừa.

Do đó polynucleotit có thể được đưa vào cây, bộ phận của cây hoặc tế bào cây theo nhiều cách mà đã biết trong lĩnh vực, như được mô tả ở trên. Do đó, không có phương pháp cụ thể để đưa một hoặc nhiều polynucleotit vào cây, mà có thể sử dụng phương pháp bất kỳ mà cho phép một hoặc nhiều polynucleotit được tích hợp ổn định trong bộ gen của cây. Khi hơn một polynucleotit được đưa vào, polynucleotit tương ứng có thể được tập hợp thành một phần của phân tử axit nucleic đơn, hoặc phân tử axit nucleic riêng biệt, và có thể nằm trên các phân tử axit nucleic giống hoặc khác nhau. Theo đó, polynucleotit có thể được đưa vào tế bào được mong muốn trong sự kiện biến nạp đơn, trong sự kiện biến nạp riêng rẽ, hoặc, ví dụ, trong thực vật, là một phần của yếu tố nhân giống.

Phương pháp bổ sung của sáng chế bao gồm sản phẩm thu hoạch tạo ra từ cây chuyển gen hoặc bộ phận của nó theo sáng chế, cũng như sản phẩm được xử lý tạo ra từ sản phẩm thu hoạch. Các sản phẩm thu hoạch có thể là tổng thể cây hoặc bộ phận bất kỳ của cây, như được mô tả ở đây. Do đó, theo một số phương án, các ví dụ không giới hạn của sản phẩm thu hoạch bao gồm hạt, quả, hoa hoặc bộ phận của chúng (ví dụ, bao phấn,

nhụy, và tương tự), lá, thân, và tương tự. Theo các phương án khác, sản phẩm xử lý bao gồm, nhưng không giới hạn ở, bột mịn, bột xay thô, dầu, tinh bột, ngũ cốc, và tương tự được tạo ra từ hạt được thu hoạch hoặc bộ phận khác của cây theo sáng chế, trong đó hạt này hoặc bộ phận khác của cây bao gồm trình tự phân tử axit nucleic/polynucleotit/nucleotit theo sáng chế.

Theo các phương án khác, sáng chế đề xuất chiết xuất từ hạt chuyển gen hoặc cây chuyển gen theo sáng chế, trong đó chiết xuất bao gồm phân tử axit nucleic, polynucleotit, trình tự nucleotit hoặc protein độc theo sáng chế. Chiết xuất từ cây hoặc bộ phận của cây có thể được làm theo các quy trình đã biết trong lĩnh vực (Xem, de la Torre và cộng sự, *Food, Agric. Environ.* 2(1):84-89 (2004); Guidet, *Nucleic Acids Res.* 22(9): 1772-1773 (1994); Lipton và cộng sự, *Food Agric. Immun.* 12:153-164 (2000)).

Chế phẩm diệt côn trùng

Theo một số phương án, sáng chế đề xuất chế phẩm diệt côn trùng bao gồm protein Cry theo sáng chế trong chất mang nông dụng. Như được sử dụng ở đây "chất mang nông dụng" có thể bao gồm vật liệu tự nhiên hoặc tổng hợp, hữu cơ hoặc vô cơ mà được kết hợp với protein Cry hoạt tính để tạo điều kiện áp dụng nó lên hoặc trong cây, hoặc bộ phận của nó. Các ví dụ của chất mang nông dụng bao gồm, không giới hạn, bột, bụi, viên, hạt, dịch phun, nhũ tương, chất keo, và dung dịch. Chất mang nông dụng còn bao gồm, nhưng không giới hạn ở, inert thành phần trợ, chất phân tán, chất hoạt động bề mặt, tá dược, chất dính, chất dán, chất liên kết, hoặc các kết hợp của chúng, mà có thể được sử dụng trong chế phẩm nông nghiệp. Các chế phẩm này có thể được áp dụng theo cách bất kỳ mà mang protein diệt loài gây hại hoặc tác nhân kiểm soát loài gây hại khác tiếp xúc với loài gây hại. Theo đó, chế phẩm có thể được áp dụng lên bề mặt của cây hoặc bộ phận của cây, bao gồm hạt, lá, hoa, thân, cuống, rễ, và tương tự. Theo các phương án khác, cây tạo ra protein Cry theo sáng chế trong thử nghiệm ở thực vật là chất mang nông nghiệp của protein Cry được biểu hiện.

Theo các phương án khác, chế phẩm diệt côn trùng bao gồm tế bào vi khuẩn hoặc tế bào vi khuẩn chuyển gen theo sáng chế, trong đó tế bào vi khuẩn hoặc tế bào vi khuẩn chuyển gen tạo ra protein Cry theo sáng chế. Chế phẩm diệt côn trùng này có thể được điều chế bằng cách làm khô, làm khô lạnh, đông hóa, chiết, lọc, ly tâm, lắng cặn, hoặc cô nuôi cấy *Bacillus thuringiensis* (Bt). Nuôi cấy Bt này có thể được chọn từ nhóm chủng

Bt chứa C0537, C0651, C0652, C0724, C0801, C1079 và M2776 được mô tả dưới đây trong phần Ví dụ hoặc nuôi cấy *Bt* chuyển gen. Theo các phương án bổ sung, chế phẩm bao gồm từ khoảng 1% đến khoảng 99% theo trọng lượng của protein Cry theo sáng chế.

Protein Cry theo sáng chế còn có thể được sử dụng kết hợp với tác nhân kiểm soát loài gây hại khác để làm tăng khoảng đích loài gây hại hoặc để ngăn ngừa hoặc quản lý sức kháng côn trùng. Do đó, theo một số phương án, sáng chế đề xuất chế phẩm mà kiểm soát một hoặc nhiều loài gây hại cây trồng, trong đó chế phẩm này bao gồm protein Cry thứ nhất theo sáng chế và tác nhân kiểm soát loài gây hại thứ hai khác protein Cry thứ nhất. Theo các phương án khác, chế phẩm là chế phẩm để áp dụng cục bộ cho cây trồng. Theo các phương án khác nữa, chế phẩm là cây chuyển gen. Theo các phương án khác nữa, chế phẩm là kết hợp của chế phẩm được áp dụng cục bộ lên cây chuyển gen. Theo một số phương án, chế phẩm bào chế bao gồm protein Cry thứ nhất theo sáng chế khi cây chuyển gen bao gồm tác nhân kiểm soát loài gây hại thứ hai. Theo các phương án khác, chế phẩm bào chế bao gồm tác nhân kiểm soát loài gây hại thứ hai khi cây chuyển gen bao gồm protein Cry thứ nhất theo sáng chế.

Theo một số phương án, tác nhân kiểm soát loài gây hại thứ hai có thể là tác nhân được chọn từ nhóm bao gồm chất diệt loài gây hại hóa học, như chất diệt côn trùng, protein diệt côn trùng *Bacillus thuringiensis* (*Bt*), protein diệt côn trùng *Xenorhabdus*, protein diệt côn trùng *Photorhabdus*, protein diệt côn trùng *BreviBacillus laterosporus*, protein diệt côn trùng *Bacillus sphaericus*, chất ức chế proteaza (cả loại sorin và xystein), lectin, alpha -amylaza, peroxidaza, cholesterol oxidaza và phân tử ARN chuỗi đôi (dsARN).

Theo các phương án khác, tác nhân kiểm soát loài gây hại thứ hai là chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm pyrethroid, carbamat, neonicotinoid, chất chặn kênh natri nơron, lacton vòng lớn diệt côn trùng, chất đối vận axit gamma-aminobutyric (GABA), urê diệt côn trùng và chất bắt chước hoocmon thiếu niên. Theo các phương án khác, chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm abamectin, axephat, axetamiprid, amidoflomet (S-1955), avermectin, azadirachtin, azinphos-metyl, bifenthrin, binfenazat, buprofezin, carbofuran, chlorfenapyr, chlorfluazuron, chlorpyrifos, chlorpyrifos-metyl, chromafenozit, clothianidin, xyfluthrin, beta-xyfluthrin, xyhalothrin, lambda-xyhalothrin, xypermethrin, xyromazin,

deltamethrin, diafenthiuron, diazinon, diflubenzuron, dimetoat, diofenolan, emamectin, endosulfan, esfenvalerat, ethiprol, fenothicarb, fenoxycarb, fenpropathrin, fenproximat, fenvalerat, fipronil, flonicamid, fluxythrinate, tau-fluvalinate, flufenerim (UR-50701), flufenoxuron, fonophos, halofenozit, hexaflumuron, imidacloprid, indoxacarb, isofenphos, lufenuron, malathion, metaldehyt, methamidophos, methidathion, metomyl, metopren, metoxychlor, monocrotophos, metoxyfenozit, nithiazin, novaluron, noviflumuron (XDE-007), oxamyl, parathion, parathion-metyl, permethrin, phorat, phosalon, phosmet, phosphamidon, pirimicarb, profenofos, pymetrozin, pyridalyl, pyriproxyfen, rotenon, spinosad, spiromesifin (BSN 2060), sulprofos, tebufenozit, teflubenzuron, tefluthrin, terbufos, tetrachlorvinphos, thiacloprid, thiametoxam, thiodicarb, thiosultap-natri, tralomethrin, trichlorfon và triflumuron, aldicarb, oxamyl, fenamiphos, amitraz, chinomethionat, chlorobenzilate, xyhexatin, dicofol, dienochlor, etoxazol, fenazaquin, fenbutatin oxit, fenpropathrin, fenpyroximat, hexythiazox, propargit, pyridaben và tebufenpyrad. Theo các phương án khác nữa, chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm xypermethrin, xyhalothrin, xyfluthrin và beta-xyfluthrin, esfenvalerat, fenvalerat, tralomethrin, fenothicarb, metomyl, oxamyl, thiodicarb, clothianidin, imidacloprid, thiacloprid, indoxacarb, spinosad, abamectin, avermectin, emamectin, endosulfan, ethiprol, fipronil, flufenoxuron, triflumuron, diofenolan, pyriproxyfen, pymetrozine và amitraz.

Theo các phương án bổ sung, tác nhân kiểm soát loài gây hại thứ hai có thể là một hoặc nhiều trong số bất kỳ các protein diệt côn trùng *Bacillus thuringiensis* bao gồm nhưng không giới hạn ở protein Cry, protein diệt côn trùng sinh dưỡng (VIP) và chất bắt chước diệt côn trùng của bất kỳ trong số các protein diệt côn trùng nêu trên. Theo các phương án khác, tác nhân kiểm soát loài gây hại thứ hai là protein Cry được chọn từ nhóm bao gồm Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ad, Cry1Ae, Cry1Af, Cry1Ag, Cry1Ah, Cry1Ai, Cry1Aj, Cry1Ba, Cry1Bb, Cry1Bc, Cry1Bd, Cry1Be, Cry1Bf, Cry1Bg, Cry1Bh, Cry1Bi, Cry1Ca, Cry1Cb, Cry1Da, Cry1Db, Cry1Dc, Cry1Dd, Cry1Ea, Cry1Eb, Cry1Fa, Cry1Fb, Cry1Ga, Cry1Gb, Cry1Gc, Cry1Ha, Cry1Hb, Cry1Hc, Cry1Ia, Cry1Ib, Cry1Ic, Cry1Id, Cry1Ie, Cry1If, Cry1Ig, Cry1Ja, Cry1Jb, Cry1Jc, Cry1Jd, Cry1Ka, Cry1La, Cry1Ma, Cry1Na, Cry1Nb, Cry2Aa, Cry2Ab, Cry2Ac, Cry2Ad, Cry2Ae, Cry2Af, Cry2Ag, Cry2Ah, Cry2Ai, Cry2Aj, Cry2Ak, Cry2Al, Cry2Ba, Cry3Aa, Cry3Ba, Cry3Bb, Cry3Ca, Cry4Aa, Cry4Ba, Cry4Ca, Cry4Cb,

Cry4Cc, Cry5Aa, Cry5Ab, Cry5Ac, Cry5Ad, Cry5Ba, Cry5Ca, Cry5Da, Cry5Ea, Cry6Aa, Cry6Ba, Cry7Aa, Cry7Ab, Cry7Ac, Cry7Ba, Cry7Bb, Cry7Ca, Cry7Cb, Cry7Da, Cry7Ea, Cry7Fa, Cry7Fb, Cry7Ga, Cry7Gb, Cry7Gc, Cry7Gd, Cry7Ha, Cry7Ia, Cry7Ja, Cry7Ka, Cry7Kb, Cry7La, Cry8Aa, Cry8Ab, Cry8Ac, Cry8Ad, Cry8Ba, Cry8Bb, Cry8Bc, Cry8Ca, Cry8Da, Cry8Db, Cry8Ea, Cry8Fa, Cry8Ga, Cry8Ha, Cry8Ia, Cry8Ib, Cry8Ja, Cry8Ka, Cry8Kb, Cry8La, Cry8Ma, Cry8Na, Cry8Pa, Cry8Qa, Cry8Ra, Cry8Sa, Cry8Ta, Cry9Aa, Cry9Ba, Cry9Bb, Cry9Ca, Cry9Da, Cry9Db, Cry9Dc, Cry9Ea, Cry9Eb, Cry9Ec, Cry9Ed, Cry9Ee, Cry9Fa, Cry9Ga, Cry10Aa, Cry11Aa, Cry11Ba, Cry11Bb, Cry12Aa, Cry13Aa, Cry14Aa, Cry14Ab, Cry15Aa, Cry16Aa, Cry17Aa, Cry18Aa, Cry18Ba, Cry18Ca, Cry19Aa, Cry19Ba, Cry19Ca, Cry20Aa, Cry20Ba, Cry21Aa, Cry21Ba, Cry21Ca, Cry21Da, Cry21Ea, Cry21Fa, Cry21Ga, Cry21Ha, Cry22Aa, Cry22Ab, Cry22Ba, Cry22Bb, Cry23Aa, Cry24Aa, Cry24Ba, Cry24Ca, Cry25Aa, Cry26Aa, Cry27Aa, Cry28Aa, Cry29Aa, Cry29Ba, Cry30Aa, Cry30Ba, Cry30Ca, Cry30Da, Cry30Db, Cry30Ea, Cry30Fa, Cry30Ga, Cry31Aa, Cry31Ab, Cry31Ac, Cry31Ad, Cry32Aa, Cry32Ab, Cry32Ba, Cry32Ca, Cry32Cb, Cry32Da, Cry32Ea, Cry32Eb, Cry32Fa, Cry32Ga, Cry32Ha, Cry32Hb, Cry32Ia, Cry32Ja, Cry32Ka, Cry32La, Cry32Ma, Cry32Mb, Cry32Na, Cry32Oa, Cry32Pa, Cry32Qa, Cry32Ra, Cry32Sa, Cry32Ta, Cry32Ua, Cry33Aa, Cry34Aa, Cry34Ab, Cry34Ac, Cry34Ba, Cry35Aa, Cry35Ab, Cry35Ac, Cry35Ba, Cry36Aa, Cry37Aa, Cry38Aa, Cry39Aa, Cry40Aa, Cry40Ba, Cry40Ca, Cry40Da, Cry41Aa, Cry41Ab, Cry41Ba, Cry42Aa, Cry43Aa, Cry43Ba, Cry43Ca, Cry43Cb, Cry43Cc, Cry44Aa, Cry45Aa, Cry46Aa, Cry46Ab, Cry47Aa, Cry48Aa, Cry48Ab, Cry49Aa, Cry49Ab, Cry50Aa, Cry50Ba, Cry51Aa, Cry52Aa, Cry52Ba, Cry53Aa, Cry53Ab, Cry54Aa, Cry54Ab, Cry54Ba, Cry55Aa, Cry56Aa, Cry57Aa, Cry57Ab, Cry58Aa, Cry59Aa, Cry59Ba, Cry60Aa, Cry60Ba, Cry61Aa, Cry62Aa, Cry63Aa, Cry64Aa, Cry65Aa, Cry66Aa, Cry67Aa, Cry68Aa, Cry69Aa, Cry69Ab, Cry70Aa, Cry70Ba, Cry70Bb, Cry71Aa, Cry72Aa và Cry73Aa.

Theo các phương án khác, tác nhân kiểm soát loài gây hại thứ hai là protein diệt côn trùng sinh dưỡng Vip3 được chọn từ nhóm bao gồm Vip3Aa1, Vip3Aa2, Vip3Aa3, Vip3Aa4, Vip3Aa5, Vip3Aa6, Vip3Aa7, Vip3Aa8, Vip3Aa9, Vip3Aa10, Vip3Aa11, Vip3Aa12, Vip3Aa13, Vip3Aa14, Vip3Aa15, Vip3Aa16, Vip3Aa17, Vip3Aa18, Vip3Aa19, Vip3Aa20, Vip3Aa21, Vip3Aa22, Vip3Aa23, Vip3Aa24, Vip3Aa25,

Vip3Aa26, Vip3Aa27, Vip3Aa28, Vip3Aa29, Vip3Aa30, Vip3Aa31, Vip3Aa32, Vip3Aa33, Vip3Aa34, Vip3Aa35, Vip3Aa36, Vip3Aa37, Vip3Aa38, Vip3Aa39, Vip3Aa40, Vip3Aa41, Vip3Aa42, Vip3Aa43, Vip3Aa44, Vip3Ab1, Vip3Ab2, Vip3Ac1, Vip3Ad1, Vip3Ad2, Vip3Ae1, Vip3Af1, Vip3Af2, Vip3Af3, Vip3Ag1, Vip3Ag2, Vip3Ag3, HM117633, Vip3Ag4, Vip3Ag5, Vip3Ah1, Vip3Ba1, Vip3Ba2, Vip3Bb1, Vip3Bb2 và Vip3Bb3.

Theo các phương án khác nữa, protein Cry thứ nhất theo sáng chế và tác nhân kiểm soát loài gây hại thứ hai được đồng biểu hiện trong cây chuyển gen. Việc đồng biểu hiện của nhiều hơn một gốc diệt loài gây hại trong cùng cây chuyển gen có thể thu được bằng cách thiết kế gen của cây để chứa và biểu hiện tất cả các gen cần thiết. Tùy ý, cây trồng, Bố mẹ 1, có thể được thiết kế gen để biểu hiện protein Cry theo sáng chế. Cây thứ hai, Bố mẹ 2, có thể thường được thiết kế gen để biểu hiện tác nhân kiểm soát loài gây hại thứ hai. Bằng cách lai Bố mẹ 1 với Bố mẹ 2, cây thế hệ con cái thu được mà biểu hiện tất cả các gen được đưa vào Bố mẹ 1 và 2.

Theo các phương án khác, sáng chế đề xuất cây chuyển gen xếp chồng kháng sự tàn phá của loài gây hại cây bao gồm trình tự ADN mã hóa dsARN để ức chế gen thiết yếu ở loài gây hại đích và trình tự ADN mã hóa protein Cry theo sáng chế thể hiện hoạt tính sinh học đối với loài gây hại đích. Đã được biết rằng dsARN không có hiệu lực đối với một số loài gây hại cánh vảy nhất định (Rajagopol và cộng sự 2002. J. Biol. Chem. 277:468-494), có thể do độ pH cao của trung tràng mà làm mất ổn định dsARN. Do đó, theo một số phương án trong đó loài gây hại đích là loài gây hại bộ cánh vảy, protein Cry theo sáng chế tác động để làm giảm tạm thời độ pH trung tràng mà giúp ổn định dsARN được ăn đồng thời tạo ra dsARN hiệu quả trong việc làm câm gen đích.

Ngoài việc đề xuất chế phẩm, sáng chế đề xuất phương pháp sản xuất protein Cry độc đối với loài gây hại bộ cánh vảy. Phương pháp này bao gồm, nuôi cấy tế bào chủ chuyển gen không phải người mà bao gồm polynucleotit hoặc gen khảm hoặc phân tử axit nucleic hoặc vector tái tổ hợp theo sáng chế dưới các điều kiện trong đó tế bào chủ tạo ra protein độc đối với loài gây hại cánh vảy. Theo một số phương án, tế bào chủ chuyển gen không phải của người là tế bào của cây. Theo một số phương án khác, tế bào cây là tế bào ngô. Theo các phương án khác, các điều kiện mà tế bào của cây hoặc tế bào cây ngô phát triển bao gồm ánh sáng mặt trời tự nhiên. Theo các phương án khác,

tế bào chủ chuyển gen không phải của người là tế bào vi khuẩn. Theo các phương án khác nữa, tế bào chủ chuyển gen không phải của người là tế bào lên men.

Theo các phương án khác của phương pháp, loài gây hại cánh vảy được chọn từ nhóm bao gồm sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xám (*Agrotis ipsilon*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu xanh hại ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá velvetbean (*Anticarsia gemmatilis*), sâu đo hại đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục trái bắp (*Richia albicosta*), tuyến trùng ăn rễ cây thuốc lá (*Heliothis virescens*), Sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc nâu (*Chilo suppressalis*), sâu đục thân mình hồng (*Sesamia calamistis*), sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*), và kết hợp bất kỳ của chúng.

Theo các phương án khác của phương pháp, gen khám bao gồm bất kỳ trong số SEQ ID NO:1-10. Theo các phương án khác nữa, protein được tạo ra bao gồm trình tự axit amin nêu ở bất kỳ trong số các SEQ ID NO: 31-40.

Theo một số phương án của phương pháp, gen khám bao gồm trình tự nucleotit mà là bộ ba mã hóa được tối ưu để biểu hiện trong cây. Theo các phương án khác, gen khám bao gồm bất kỳ trong số các SEQ ID NO:11-20. Theo các phương án khác nữa, protein được tạo ra bao gồm trình tự axit amin nêu ở bất kỳ trong số các SEQ ID NO:31-50.

Theo các phương án khác, sáng chế đề xuất phương pháp sản xuất cây chuyển gen kháng loài gây hại (ví dụ, kháng côn trùng), bao gồm: đưa vào cây polynucleotit, gen khám, vector tái tổ hợp, cat-xet biểu hiện hoặc phân tử axit nucleic theo sáng chế bao gồm trình tự nucleotit mà mã hóa protein Cry theo sáng chế, trong đó trình tự nucleotit được biểu hiện ở cây, nhờ đó tạo cho cây sức kháng đối với ít nhất loài gây hại côn trùng bộ cánh vảy, và tạo ra cây chuyển gen kháng côn trùng. Theo một số phương án, cây chuyển gen kháng loài gây hại kháng đối với ít nhất sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) hoặc sâu xám (*Agrotis ipsilon*) khi so với cây đối chứng không có polynucleotit, gen khám, vector tái tổ hợp, cat-xet biểu hiện hoặc phân tử axit nucleic theo sáng chế. Theo một số phương án, việc đưa vào đạt được bằng cách biến nạp cây. Theo các phương án khác, việc đưa vào đạt được bằng cách lai cây thứ nhất chứa gen

khảm, vector tái tổ hợp, tổ hợp biểu hiện hoặc phân tử axit nucleic theo sáng chế với cây thứ hai khác.

Theo một số phương án, cây chuyển gen theo sáng chế mà kháng đối với ít nhất sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) hoặc sâu xám (*Agrotis ipsilon*) còn kháng đối với ít nhất một côn trùng bổ sung khác, trong đó côn trùng bổ sung bao gồm, nhưng không giới hạn ở, sâu xanh mùa thu (*Spodoptera frugiperda*), sâu xanh hại ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá velvetbean (*Anticarsia gemmatilis*), sâu đo hại đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục trái bắp (*Richia albicosta*), tuyến trùng ăn rễ cây thuốc lá (*Heliothis virescens*), Sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc nâu (*Chilo suppressalis*), sâu đục thân mình hồng (*Sesamia calamistis*) hoặc sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*), và kết hợp bất kỳ của chúng.

Theo các phương án khác, phương pháp kiểm soát ít nhất loài gây hại côn trùng bộ cánh vảy như sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) hoặc sâu xám (*Agrotis ipsilon*) được đề xuất, phương pháp bao gồm phân phối đến côn trùng lượng có hiệu quả của protein Cry theo sáng chế. Để có hiệu quả, protein Cry đầu tiên được côn trùng ăn vào bụng qua đường miệng. Tuy nhiên, protein Cry được phân phối đến côn trùng theo nhiều cách đã biết. Phương pháp phân phối protein qua đường miệng đến côn trùng bao gồm, nhưng không giới hạn ở, tạo ra protein (1) trong cây chuyển gen, trong đó côn trùng ăn (ăn vào bụng) một hoặc nhiều phần của cây chuyển gen, nhờ đó ăn vào bụng polypeptit mà được biểu hiện ở cây chuyển gen; (2) trong (các) chế phẩm protein được phối chế mà có thể được áp dụng lên hoặc kết hợp vào, ví dụ, môi trường nuôi côn trùng; (3) trong (các) chế phẩm protein mà có thể được áp dụng lên bề mặt, ví dụ, phun, lên bề mặt bộ phận của cây, mà sau đó được côn trùng ăn vào bụng do côn trùng ăn một hoặc nhiều phần cây được phun; (4) ma trận môi; hoặc (5) bất kỳ hệ phân phối protein đã biết khác trong lĩnh vực. Do đó, phương pháp phân phối qua đường miệng bất kỳ đến côn trùng có thể được dùng để phân phối protein Cry độc theo sáng chế. Theo một số phương án cụ thể, protein Cry theo sáng chế được phân phối theo đường miệng đến côn trùng, trong đó côn trùng ăn vào bụng một hoặc nhiều phần của cây chuyển gen.

Theo các phương án khác, protein Cry theo sáng chế được phân phối theo đường miệng đến côn trùng, trong đó côn trùng ăn vào bụng một hoặc nhiều phần của cây được

phun với chế phẩm bao gồm các protein Cry theo sáng chế. Phân phối chế phẩm theo sáng chế lên bề mặt cây có thể được thực hiện bằng cách sử dụng phương pháp bất kỳ đã biết đối với chuyên gia trong lĩnh vực để áp dụng hợp chất, chế phẩm, công thức và tương tự lên bề mặt cây. Một số ví dụ không giới hạn phân phối hoặc cho tiếp xúc cây hoặc bộ phận của nó bao gồm phun, rải bụi, rải, tung rắc, sương mù, tán nhỏ, truyền phát, ngâm, bơm vào đất, kết hợp với đất, làm ướt sũng (ví dụ, rẫy, xử lý đất), nhúng, rót, bao, thấm qua lá hoặc thân, bọc hoặc xử lý hạt, và tương tự, và kết hợp của chúng. Các quy trình này và quy trình khác để cho cây hoặc bộ phận của nó tiếp xúc với (các) hợp chất, (các) chế phẩm hoặc (các) công thức là đã biết với chuyên gia trong lĩnh vực.

Theo các phương án khác, sáng chế đề xuất phương pháp kiểm soát côn trùng cánh vảy kháng Cry1Ab bao gồm phân phối lên côn trùng lượng có hiệu quả của protein diệt côn trùng có độ tương đồng ít nhất 80% với SEQ ID NO:31 hoặc SEQ ID NO:32. Theo các phương án khác, côn trùng kháng Cry1Ab là sâu đục thân mía (*Diatraea saccharalis*). Theo các phương án khác nữa, protein diệt côn trùng bao gồm SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:41 hoặc SEQ ID NO:42.

Theo các phương án khác, sáng chế đề xuất phương pháp ngăn ngừa sự phát triển sức kháng trong quần thể loài gây hại côn trùng bộ cánh vảy đích đối với protein Cry1Ab được biểu hiện ở cây chuyển gen bao gồm xếp chồng trong cây chuyển gen ngoài protein Cry1Ab protein diệt côn trùng thứ hai có độ tương đồng ít nhất 80% với SEQ ID NO:31 hoặc SEQ ID NO:41. Theo một số phương án, loài gây hại côn trùng cánh vảy đích là sâu đục thân mía (*Diatraea saccharalis*). Theo các phương án khác nữa, protein diệt côn trùng thứ hai bao gồm SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:41 hoặc SEQ ID NO:42.

Theo một số phương án, sáng chế bao gồm phương pháp cung cấp cho nông dân bằng cách kiểm soát loài gây hại cánh vảy, phương pháp bao gồm cung cấp hoặc bán cho nông dân vật liệu cây như hạt, vật liệu cây bao gồm polynucleotit, gen khảm, cat-xet biểu hiện hoặc vectơ tái tổ hợp có khả năng biểu hiện protein Cry theo sáng chế ở cây phát triển từ hạt, như được mô tả ở trên.

Các phương án theo sáng chế có thể được hiểu rõ hơn bằng cách tham khảo đến các ví dụ sau đây. Phần mô tả trên và sau đây của các phương án theo sáng chế và các phương án khác nhau không nhằm giới hạn phạm vi bảo hộ, mà chỉ để minh họa cho

yêu cầu bảo hộ. Do đó, cần hiểu rằng điểm không giới hạn ở các chi tiết cụ thể trong các ví dụ này. Chuyên gia trong lĩnh vực hiểu rằng các phương án khác của sáng chế có thể được thực hiện mà không vượt ra khỏi phạm vi và tinh thần của bản mô tả này, phạm vi của sáng chế được xác định trong yêu cầu bảo hộ kèm theo

Ví dụ thực hiện sáng chế

Ví dụ 1. Xác định các chủng Bt hoạt tính

Các phân lập *Bacillus thuringiensis* được nuôi cấy từ các bào tử có mặt trong sự gom hiện tại và được giữ trên các tấm aga T3 + penixillin. Mỗi phân lập được nuôi lớn ưa khí trong các khối sâu 24 giếng trong khoảng 10 ngày ở 28°C đến khi hình thành bào tử, mà được kiểm tra bằng cách nhuộm màu bằng thuốc nhuộm Coomassie blue/axit axetic và quan sát bằng kính hiển vi. Sau khi hình thành bào tử cả các phần tan và không tan được thử hoạt tính kháng các loài bộ cánh vảy được mong muốn. Các đoạn được thử trong sinh thử nghiệm nhiễm xạ bề mặt, khi các đoạn được phủ lên khẩu phần nhân tạo đa loài. Mỗi phân lập được sàng lọc đối với ít nhất bốn loài bộ cánh vảy, bao gồm *Helicoverpa zea* (sâu xanh hại ngô), *Agrotis ipsilon* (sâu xám), *Ostrinia nubilalis* (Sâu đục thân ngô châu Âu), và *Spodoptera frugiperda* (sâu xanh mùa thu) với cùng kích thước của 12 ấu trùng sơ sinh. Thời gian của mỗi thử nghiệm là khoảng 7 ngày ở nhiệt độ phòng; các tấm được chấm điểm đối với tỷ lệ chết cũng như ức chế tăng trưởng ấu trùng. Tỷ lệ chết quan sát được ở mức tăng bằng 30% so với đối chứng âm được coi là có hoạt tính. Dựa trên thử nghiệm côn trùng ban đầu, các chủng C0530, C0537, C0651, C0652, C0724, C0801, C1079 và M2776 được lựa chọn để phân tích thêm.

Ví dụ 2. Ghép và phân tích bộ gen

Các gen *Bt cry* theo sáng chế được xác định từ các chủng được xác định ở Ví dụ 1 sử dụng nguyên tắc sắp trình tự toàn bộ bộ gen. Tóm lại, ADN của *Bacillus* được cắt sử dụng thiết bị siêu âm Covaris S2 (Covaris, Inc., Woburn, MA) với chương trình ADN_400bp được cài đặt ở chu kỳ hoạt động: 10%; cường độ: 4; chu kỳ/nỗ: 200. ADN được xử lý bằng mô hình sửa đầu/nối đuôi dA NEBNext® Ultra™ (New England Biolabs, Inc. Ipswich, MA). Các chuỗi tiếp hợp chỉ số Biooscience 1-57 (1-27 Brazil, 28-57 USA, UK và Switzerland) được buộc lại sử dụng NEB Quick Ligation™ như được mô tả bởi nhà cung cấp (New England Biolabs, Inc. Ipswich, MA). Việc buộc thắt được làm sạch

sử dụng các hạt Agencourt AMPure XP như được mô tả bởi nhà cung cấp (Beckman Coulter, Inc., Indianapolis, IN).

Thư viện có kích thước được chia đoạn như sau: Mẫu 50 uL được trộn với 45 ul 75% hỗn hợp hạt (25% hạt AMPure cộng thêm 75% dung dịch NaCl/PEG TekNova cat # P4136). Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi thu được được chuyển đến giếng mới và 45 ul 50% hỗn hợp hạt (50% Hạt AMPure cộng 50% dung dịch NaCl/PEG TekNova cat # P4136) được thêm. Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi thu được được loại ra và các hạt được rửa bằng 80% etanol. 25 uL đệm đệm tách rửa (EB) được thêm và Hỗn hợp được đặt trên giá từ. Dịch nổi thu được cuối cùng được loại ra và đặt vào ống 1,5 mL. Phương pháp này tạo ra các thư viện trong khoảng kích thước của các cặp bazơ (bp) ADN 525 (chèn thêm chuỗi tiếp hợp).

Thư viện ADN được định kích thước được khuếch đại sử dụng KAPA Biosystem HiFi Hot Start (Kapa Biosystem, Inc., Wilmington, MA) sử dụng các điều kiện chu kỳ sau đây: [98°C, 45s]; 12 x [98°C, 15s, 60°C, 30s, 72°C, 30s]; [72°C, 1 phút]. Mỗi phản ứng chứa: 5 ul thư viện ADN library, 1 uL đoạn môi thông thường Bioscience (25 uM), 18 uL nước vô trùng, 1uL đoạn môi được đánh số thứ tự Bioscience (25 uM), 25 ul 2X KAPA HiFi polymeraza.

Các thư viện được chạy trên thiết bị phân tích sinh học Agilent 2100 (Agilent Technologies, Santa Clara, CA) sử dụng chip độ nhạy cao để xác định khoảng kích thước của thư viện và kích thước chèn trung bình. Tất cả các thư viện được xử lý sắp trình tự các đầu bắt cặp (PE) (100 chu kỳ mỗi lần đọc; 12-24 thư viện mỗi lần) trên hệ sắp trình tự HiSeq 2500 sử dụng nguyên tắc sắp trình tự của nhà sản xuất (Illumina, Inc., San Diego, CA).

Công cụ phân tích điện toán *Bacillus* được phát triển để xác định và đặc trưng gen giống Cry thích hợp được sử dụng để ưu tiên chỉ để thử tiếp trong phòng thí nghiệm.

Ghép bộ gen và phân tích cũng như phân tích thư viện bộ gen được mô tả ở trên dẫn đến xác định năm gen tương tự Cry1 trong chủng *Bacillus thuringiensis* với độc tố đối với ít nhất sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) hoặc sâu xám (*Agrotis ipsilon*). Xác định các đặc điểm của gen tương tự Cry1 và protein được thể hiện trong Bảng 1.

Bảng 1. GenCry/protein được xác định trong chủng *Bacillus thuringiensis*.

Chủng	Protein/Gen Tên	Thành viên họ Cry gần nhất	Trọng lượng phân tử (kD)	Nucleotit SEQ ID NO:	Axit amin SEQ ID NO:
C0530	BT0001	Cry1Ai	132,2	1	31
C0537	BT0003	Cry1Aa	129,7	2	32
C0651	BT0020	Cry1Bg	129,6	3	33
C0652	BT0022	Cry1If	130,8	4	34
C0724	BT0027	Cry1La	130,3	5	35
C0801	BT0029	Cry1Gb	133,1	6	36
C1079	BT0030	Cry1Hb	128,9	7	37
C1079	BT0031	Cry1Ib	81,0	8	38
M2776	BT0201	Cry1Ga	132,0	9	39
M2776	BT0202	Cry1Jd	132,6	10	40

Ví dụ 3. Biểu hiện protein Bt trong tế bào chủ tái tổ hợp

Biểu hiện *Bacillus*. Protein Cry được mô tả trong Ví dụ 2 được biểu hiện ở chủng *Bacillus thuringiensis* (*Bt*) tinh thể âm không có hoạt tính diệt côn trùng nên quan sát được thông qua pCIB5634' có thiết kế vectơ hình thoi, được thiết kế để biểu hiện trong cả *E. coli* và *Bt*. Vectơ pCIB5634' bao gồm vùng khởi động Cry1Ac mà dẫn hướng biểu hiện của gen *Bt* Cry tách dòng và chỉ dấu kháng erythromycin. Cat-xet biểu hiện bao gồm trình tự mã hóa Cry có lợi được biến nạp vào chủng *Bt* chủ thông qua điện di và chủng *Bt* chuyển gen được lựa chọn trên tấm aga chứa erythromycin. Chủng *Bt* chuyển gen được chọn được nuôi đến pha hình thành bào tử trong môi trường T3 ở 28°C trong 4-5 ngày. Các viên tế bào được thu và rửa đi rửa lại trước khi hòa tan trong đệm carbonat độ pH cao (50mM) chứa 2mM DTT.

Biểu hiện *E. coli*. Cry protein được biểu hiện ở chủng *E. coli* sử dụng vectơ pET28a hoặc pET29a (Merck KGaA, Darmstadt, Đức). Các cấu trúc được biến nạp bằng điện di

và dòng vô tính *E. Coli* chuyển gen được lựa chọn trên tấm aga chứa kanamycin. Các chủng *E. Coli* chuyển gen được chọn được nuôi và biểu hiện protein Cry được cảm ứng sử dụng cảm ứng IPTG ở 28°C. Các tế bào được tạo huyền phù lại trong đệm carbonat độ pH cao (50mM) chứa 2mM DTT và sau đó phá vỡ bằng cách sử dụng thiết bị đồng nhất LV-1 vi dòng chảy.

Phân tích biểu hiện. Hợp chất dung giải tế bào tạo thành từ hoặc là chủng *Bt* hoặc *E. coli* chuyển gen sau đó được làm trong qua ly tâm và các mẫu được phân tích độ tinh khiết thông qua SDS-PAGE và điện di đồ sử dụng hệ BioRad Experion (Biorad, Hercules, CA). Tổng nồng độ protein được xác định qua thử nghiệm Bradford hoặc Thermo 660. Protein Cry được tinh chế sau đó được thử trong thử nghiệm sinh học được mô tả dưới đây.

Ví dụ 4. Hoạt tính của protein Cry trong sinh thử nghiệm

Protein Cry được tạo ra trong Ví dụ 3 được thử đối với một hoặc nhiều trong số loài gây hại bộ cánh vảy sau đây sử dụng phương pháp thử nghiệm sinh học khẩu phần nhân tạo đã biết trong lĩnh vực: Sâu đục thân ngô châu Âu (ECB; *Ostrinia nubilalis*), sâu xám (BCW; *Agrotis ipsilon*), sâu đục thân mía (SCB; *Diatraea saccharalis*), sâu xanh mùa thu (FAW; *Spodoptera frugiperda*), sâu xanh hại ngô (CEW; *Helicoverpa zea*), sâu ăn lá velvetbean (VBC; *Anticarsia gemmatilis*), sâu đo hại đậu tương (SBL; *Pseudoplusiabao gommns*), sâu đục thân ngô tây nam (SWCB; *Diatraea grandiosella*), sâu đục trái bắp (WBCW; *Striacostaalbicosta*), tuyến trùng ăn rễ cây thuốc lá (TBW; *Heliothis virescens*), sâu đục thân ngô châu Á (ACB; *Ostrinia furnacalis*), sâu đục quả bông (CBW; *Helicoverpa armigera*), sâu đục thân sọc nâu (SSB; *Chilo suppressalis*), sâu đục thân mình hồng (PSB; *Sesamia inferens*), Sâu cắn lá phượng Đông (OAW; *Mythimna separata*) và sâu cuốn lá nhỏ hại lúa (RLF; *Cnaphalocrocis medinails*).

Lượng bằng nhau của protein trong dung dịch được áp dụng lên bề mặt của các khẩu phần nhân tạo cho côn trùng (Bioserv, Inc., Frenchtown, NJ) trong tấm 24 giếng. Sau khi bề mặt khẩu phần khô, ấu trùng của các loài côn trùng thử được thêm vào mỗi giếng. Các tấm được đóng kín và giữ ở điều kiện môi trường phòng thí nghiệm về nhiệt độ, ánh sáng và độ ẩm tương đối. Nhóm đối chứng dương bao gồm ấu trùng được phơi nhiễm ra chủng *Bacillus* kiểu đại có hoạt tính mạnh và phổ rộng. Nhóm đối chứng âm bao gồm ấu trùng được phơi nhiễm ra khẩu phần côn trùng được xử lý chỉ với dung dịch

đệm và ấu trùng trên khẩu phần côn trùng không được xử lý; tức là chỉ khẩu phần. Tỷ lệ chết được đánh giá sau khoảng 120 giờ và cho điểm so với các đối chứng.

Các kết quả được thể hiện trong Bảng 2, trong đó a "-" nghĩa là không có hoạt tính so với nhóm đối chứng, a "+/-" nghĩa là 0-10% hoạt tính được so sánh để kiểm tra (loại này còn bao gồm 0% tỷ lệ chết với ức chế tăng trưởng ấu trùng mạnh), a "+" nghĩa là 10-25% hoạt tính được so sánh để kiểm tra, a "++" nghĩa là 26-75% hoạt tính được so sánh để kiểm tra, và a "+++" 76-100% hoạt tính được so sánh để kiểm tra. Ký hiệu "nt" trong Bảng 2 nghĩa là protein được chỉ ra không được thử đối với côn trùng cụ thể đó.

Bảng 2. Các kết quả của thử nghiệm sinh học với Protein Cry.

	Protein BT Cry									
Côn trùng	0001	0003	0020	0022	0027	0029	0030	0031	0201	0202
FAW	-	-	-	-	+	+	-	-	-	+/-
CEW	+++	++	-	+	-	+/-	+/-	-	-	-
ECB	+++	+++	+++	+++	+++	-	+/-	-	++	++
BCW	+++	+++	++	+++	-	+	-	+/-	-	+/-
SCB	+++	+++	+++	+++	-	++	nt	-	-	-
VBC	nt	nt	+++	+++	nt	+++	nt	-	+++	nt
SBL	+++	+++	+++	+++	nt	+++	nt	-	-	+++
SWCB	++	++	++	+++	nt	+	nt	nt	+++	+/-
WBC W	++	++	+	-	nt	nt	++	nt	-	nt
TBW	+++	+++	+++	++	nt	nt	nt	nt	-	+/-
ACB	+++	+++	+++			+	+++		++	+
OAW	+++	++	+			-	+/-		+/-	-

Ví dụ 5. Cry Protein hoạt tính đối với Côn trùng kháng Cry1Ab

Mặc dù sử dụng phổ biến cây trồng chuyển gen biểu hiện protein Cry tạo ra kiểm soát chưa từng thấy của một số loài gây hại côn trùng cánh vảy nhất định, nhưng mối đe dọa mà loài gây hại côn trùng được nhắm đích có thể phát triển sức kháng đối với kỹ thuật này là có thực. Đến nay đã có các báo cáo về sức kháng trên cánh đồng mà dẫn đến kiểm soát không thành công hoặc hiệu quả kiểm soát giảm đáng kể do việc sử dụng mạnh cây trồng chuyển gen biểu hiện protein Cry ở ít nhất bốn trường hợp: sâu xanh mùa thu (*Spodoptera frugiperda*) đối với Cry1F ở ngô ở Puerto Rico (Storer và cộng sự 2010. J. Econ. Entomol. 103:1031-1038), Sâu đục thân châu Phi (*Busseolafusca*) đối với ngô Cry1Ab ở Nam Phi (Van Rensburg. 2007. J. Plant Soil. 24:147-151), sâu hồng đục quả (*Pectinophoragossypiella*) đối với bông Cry1Ac ở Ấn Độ (Dhurua & Gujar. 2011. India Pest Manage. Sci. 67:898-903) và sâu cắn rễ ngô phương tây (*Diabroticavirgifera*) đối với ngô Cry3Bb1 ở Mỹ. Loài gây hại côn trùng khác cũng có thể có tiềm năng phát triển mức kháng liên quan về kinh tế đối với protein Cry hiện được sử dụng một cách kinh tế trong cây trồng chuyển gen. Ví dụ, Zhang và cộng sự. (2013. J. Invert. Pathol. 112:267-272) sâu đục thân mía (SCB; *Diatraea saccharalis*) được gom trên cánh đồng được sàng lọc về sức kháng và xác định tám trong số 191 SCB cá thể mà có alen kháng chính đối với cây ngô Cry1Ab. Từ tám cá thể này Zhang và cộng sự thiết lập và chọn lọc thêm quần thể SCB kháng Cry1Ab trong phòng thí nghiệm. Zhang và cộng sự còn xác định rằng SCB kháng Cry1Ab kháng chéo cao đối với các protein Cry khác, bao gồm Cry1Aa và Cry1Ac và kháng chéo vừa phải đối với Cry1F.

Sâu đục thân mía là loài gây hại cánh vảy chính của ngô ở vùng trung nam của Mỹ và ở Nam Mỹ. Do đó, quan trọng là tìm ra protein Cry mà có cách hoạt động khác với protein Cry hienejcos trên thị trường để dùng làm công cụ quản lý sức kháng côn trùng. Đặc biệt hữu dụng là protein mà có hiệu quả cao đối với côn trùng kháng Cry1Ab.

Một số protein Cry nhất định theo sáng chế được thử về tính hiệu quả đối với sâu đục thân mía kháng Cry1Ab được mô tả trong Zhang và cộng sự (*trước đây*). Cả SCB nhạy Cry1Ab (kiểu đại) và kháng Cry1Ab được thử nghiệm cùng thời điểm cơ bản như được mô tả trong Zhang và cộng sự các Kết quả, được thể hiện trong Bảng 3, chỉ ra rằng BT-0001 và BT-0003 hiệu quả cao đối với côn trùng kháng Cry1Ab. Điều này là đặc biệt ngạc nhiên vì BT-0003 và BT-0001 có 93 và 87 phần trăm tương đồng, tương ứng, với Cry1Aa; protein mà Zhang và cộng sự đã chứng minh là không hiệu quả đối với sâu đục thân mía kháng Cry1Ab do kháng chéo.

Bảng 3. Hoạt tính của Protein Cry đối với Côn trùng kháng Cry1Ab.

Xử lý	Phần trăm tỷ lệ chết SCB	
	Nhạy Cry1Ab	Kháng Cry1Ab
BT-0001	100	92*
BT-0003	100	92*
Đối chứng âm 1	12	0
Đối chứng âm 2	8	0

* Tỷ lệ chết 'thực tế', như được xác định bởi Zhang và cộng sự (*trước đây*), là 100%.

Kết quả này đặc biệt ngạc nhiên đối với protein BT-0003 vì hơn 80% protein là trình tự Cry1Aa, chỉ có ba khác biệt axit amin với Cry1Aa1 trong vùng từ vị trí 203-1176 của SEQ ID NO:31 như được thể hiện trong các Bảng 4A và 4B dưới đây. Khác biệt axit amin được thể hiện bằng thay đổi chữ cái. Cùng axit amin được chỉ ra bằng đoạn ("."). Thú vị là, trình tự axit amin BT-0003 tự nhiên là vật lai giữa Cry1Ah từ vị trí axit amin 1-202 và Cry1Aa từ vị trí axit amin 203-1176. Do đó thấy rằng 202 axit amin thứ nhất và/hoặc ba đột biến axit amin trong vùng từ vị trí 203-1176 chịu trách nhiệm đối với tính hiệu quả đối với côn trùng kháng Cry1Ab.

Do đó rõ ràng từ các kết quả này rằng cả BT-0001 và BT-0003 tạo ra trung bình phát triển sức kháng giảm nhẹ đối với cây chuyển gen chứa protein Cry1Ab hoặc trung bình đánh bại côn trùng kháng Cry1Ab, đặc biệt là trong vùng nhiệt đới nơi một số khía cạnh sinh vật học của loài gây hại, như sự hình thành chong lán mỗi mùa sinh trưởng, góp vào sự tiến triển sức kháng.

Bảng 4A. Sắp hàng trình tự của BT-0003 với Cry1Ah1 và Cry1Aa1.

BT-0003pro	1	MEIVNNQNQCVPYNCLNNPEIEIELEGGRISVGNTPIDISLSLTQFLLSEF
Cry1Ah1	1	
Cry1Aa1	1	.DNNP.I.E.I.....S...V.V.G.E..ET.Y.....
BT-0003pro	51	VPGAGFVLGLIDLWGFVGPSQWDAFLAQVEQLINQRIAEAVRNTAIQEL
Cry1Ah1	51	
Cry1Aa1	51	V.I...IF.....PV.I.....E.FA..Q..SR.
BT-0003pro	101	EGMARVYRTYATAFAEWEKAPDDPELREALRTQFTATETYISGRISVLKI
Cry1Ah1	101	
Cry1Aa1	101	..LSNL.QI..ES.R...AD.TN.A...EM.I..NDMNSALT.TA.PL.AV
BT-0003pro	151	QTFEVQLLSVFAQAANLHLSLLRDVVFQGRWGFTTTVNNYYNDLTEGI
Cry1Ah1	151	
Cry1Aa1	151	.NYQ.P....YV.....V....SV.....DAA.I.SR.....RL.
BT-0003pro	201	STYTDYAVRWYNTGLERVWGPDSRDWVRYNQFRRELTLTVLDIVAFSNI
Cry1Ah1	201	P..
Cry1Aa1	201	GN.....
BT-0003pro	251	DSRRYPPIRTVSQLTREIYTNPVLENFDGSFRGMAQRIEQNIRQPHLMDIL
Cry1Ah1	251	S..G..RS..S.....
Cry1Aa1	251	
BT-0003pro	301	NSITIYTDVHRGFNYWSGHQITASPVGFSGPEFAFPLFGNAGNAAPPV-L
Cry1Ah1	301	A...YY.....M.....T...Y.TM.....QQRI
Cry1Aa1	301	-
BT-0003pro	350	VSLTGLGIFRTLSSPLYRRIILGSGPNNQELFVLDGTEFSFASLTTLNLP
Cry1Ah1	351	.AQL.Q.VY.....TF...-PFNI.I...Q.S.....AYGT-SS....
Cry1Aa1	350	
BT-0003pro	400	TIYRQRGTVDSLDVIPPQDNSVPPRAGFSHRLSHVTML-SQAAGAVYTLLR
Cry1Ah1	399	AV..KS.....E....N.N....Q.....S.FR.GSSSS.SII.
Cry1Aa1	400	-
BT-0003pro	449	APTFSWQHRSAEFNNIIPSSQITQIPLTKSTNLGSGTSVVKGPFGFTGGDI
Cry1Ah1	449	..M...I.....A.DS....AV.GNF.FN.-..IS.....L
Cry1Aa1	449	
BT-0003pro	499	LRRTSPGQISTLRVNITAPL-----SQRYRVIRIYASTTNLQFHTSIDGR
Cry1Ah1	498	V.LN.S.NNIQN.GY.EV.IHFPST.T....V...V.PIHLNVNWGNS
Cry1Aa1	499	-----
BT-0003pro	544	PINQGNFSATMSSGSNLQSGSFRTVGFTTPFNFSNGSSVFTLSAHVFNSG
Cry1Ah1	548	S.FSNTVP..AT.LD....SD.GYFESANA.TS.L.NI.---GVRN.SGT
Cry1Aa1	544	
BT-0003pro	594	NEVYIDRIEFVPAEVTFAEYDLERAQKAVNELFTSSNQIGLKTVDVTDYH
Cry1Ah1	595	AG.I...F..I.VTA.L....N.....A....T..L....N....
Cry1Aa1	594	
BT-0003pro	644	IDQVSNLVECLSDDEFCLDEKQELSEKVKHAKRLSDERNLLQDPNFRGINR
Cry1Ah1	645	T.....R.....S..KD...
Cry1Aa1	644	

Bảng 4B. Sắp hàng tiếp của BT-0003 với Cry1Ah1 và Cry1Aa1.

BT-0003pro	694	QLDRGWGRGSTDITIQGGDDVFKENYVTLTGTDFDECYPTYLYQKIDESKLG
Cry1Ah1	695	.PE...G...G.....V.....S.....
Cry1Aa1	694	

BT-0003pro	744	AYTRYQLRGYIEDSQDLEIYLIIRYNAKHETVNVPGTGSWPLSAQSPIGK
Cry1Ah1	745	.F.....V.....L.....AVK....R
Cry1Aa1	744	
BT-0003pro	794	CGEPNRCAPHLEWNPDLDCSCRDGEKCAHHSHHFSLDIDVGCTDLNEDLG
Cry1Ah1	795	-----.....
Cry1Aa1	794	
BT-0003pro	844	VWVIFKIKTQDGHARLGNLEFLEEKPLVGEALARVKIAEKKWRDKREKLE
Cry1Ah1	825	KI.....L.L.....K.....
Cry1Aa1	844	R.....
BT-0003pro	894	WETNIVYKEAKESVDALFVNSQYDQLQADTNIAMIHAAADKRVHSIREAYL
Cry1Ah1	875	D...NR..T.....R.....
Cry1Aa1	894	
BT-0003pro	944	PELSVIPGVNAAI FEELEGRIFTAFSLYDARNVIKNGDFNGLSCWNVKG
Cry1Ah1	925	L.....Y.....
Cry1Aa1	944	
BT-0003pro	994	HVDVEEQNNQRSVLVPEWEAEVSQEVVRVCPGRGYILRV TAYKEGYGEGC
Cry1Ah1	975	H.....I.....
Cry1Aa1	994	L.....
BT-0003pro	1044	VTIHEIENNTDELKFSNCVEEIIYSNNTVTCNDYTVNQEEYGGAYTSRNR
Cry1Ah1	1025	D.....V.P.....AT....E.T.....
Cry1Aa1	1044	P.....
BT-0003pro	1094	GYNEA----PSVPADYASVYEEKSYTDGRRENPCFNRGYRDTPLPVGY
Cry1Ah1	1075	..DG.YESNS.....A....A.....D....S.....A..
Cry1Aa1	1094	----.....
BT-0003pro	1140	VTKELEYFPETDKVWIEIGETEGTFIVDSVELLLMEE
Cry1Ah1	1125	
Cry1Aa1	1140	

Ví dụ 6. Tạo vector của gen để biểu hiện thực vật

Trước khi biểu hiện ở cây, polynucleotit tổng hợp bao gồm trình tự nucleotit mã hóa protein Cry đột biến, mBT-0001 và mBT-0003 (SEQ ID NO:41 và 42, tương ứng) được tổng hợp trên nền tổng hợp gen tự động hóa (Genscript, Inc., Piscataway, NJ). Cho ví dụ, cat-xet biểu hiện thứ nhất được tạo ra bao gồm vùng khởi động ubiquitin của ngô (Ubi1) liên kết hoạt động với protein Cry mã hóa trình tự mà liên kết hoạt động với vùng kết thúc Nos và cat-xet biểu hiện thứ hai được tạo ra bao gồm vùng khởi động Ubi1 liên kết hoạt động với phosphomannoza isomeraza (PMI) mã hóa trình tự mà liên kết hoạt động với trình tự kết thúc số. Biểu hiện của PMI cho phép chọn lọc dương tính cây chuyển gen trên mannoza. Cả tổ hợp biểu hiện được tách dòng thành vector thích hợp để biến nạp ngô trung gian *Agrobacterium*.

Ví dụ 7. Biểu hiện và Hoạt tính của Cry Protein ở Cây Ngô

Biến nạp phôi ngô chưa chín muối được thực hiện cơ bản như được mô tả trong Negrotto và cộng sự, 2000, Tế bào cây Reports 19: 798-803. Tóm lại, chủng *Agrobacterium* LBA4404 (pSB1) bao gồm vector biểu hiện được mô tả trong Ví dụ 5 được nuôi trên môi trường rắn YEP (chiết xuất nấm men (5 g/L), pepton (10g/L), NaCl (5g/L), 15g/l agar, pH 6,8) trong 2-4 ngày ở 28°C. Xấp xỉ $0,8 \times 10^9$ tế bào *Agrobacterium* được tạo huyền phù trong môi trường LS-inf được tạo huyền phù với 100 μ M As. Vi khuẩn được cảm ứng trước trong môi trường này trong khoảng 30-60 phút.

Phôi chưa chín muối từ dòng ngô bẩm sinh được cắt mô từ các tai 8-12 ngày tuổi vào LS-inf + 100 μ M As lỏng. Phôi được rửa một lần với môi trường gây nhiễm mới. Sau đó dung dịch *Agrobacterium* được thêm và phôi được tạo xoáy trong 30 giây và cho xử lý bằng vi khuẩn trong 5 phút. Sau đó phôi được chuyển vẩy nhỏ hướng lên môi trường LSAs và được nuôi cấy trong tối trong hai đến ba ngày. Sau đó, khoảng từ 20 đến 25 phôi mỗi tấm petri được chuyển vào môi trường LSDc được bổ sung xefotaxim (250 mg/l) và bạc nitrat (1,6 mg/l) và được nuôi cấy trong tối ở khoảng 28°C trong 10 ngày.

Các phôi chưa chín, tạo ra các thể sản phôi được chuyển vào môi trường LSD1M0.5S. Mề nuôi cấy được chọn trên môi trường này trong khoảng 6 tuần với bước nuôi cấy phụ ở khoảng 3 tuần. Thể sản sống sót được chuyển sang môi trường Reg1 được bổ sung mannoza. Sau khi nuôi cấy trong sáng (chế độ 16 giờ sáng/ 8 giờ tối), mô xanh lá sau đó được chuyển vào môi trường Reg2 mà không cần chất điều chỉnh tăng trưởng và được ủ trong khoảng 1-2 tuần. Cây non được chuyển sang hộp Magenta GA-7 (Magenta Corp, Chicago Ill.) chứa môi trường Reg3 và nuôi trong sáng. Sau khoảng 2-3 tuần, cây được thử về sự có mặt của gen PMI gen *Bt cry* bằng thử nghiệm PCR. Cây dương tính từ thử nghiệm PCR được chuyển vào nhà kính để đánh giá thêm.

Cây chuyển gen được đánh giá về số sao chép (được xác định bằng phân tích Taqman), mức biểu hiện protein (được xác định bằng ELISA), và tính hiệu quả kháng các loài côn trùng được mong muốn trong các sinh thử nghiệm cắt bỏ lá. Đặc biệt là, mô cây (lá hoặc tơ) được cắt mô từ sự kiện bản sao đơn (giai đoạn V3-V4) và được cho tàn phá với ấu trùng sơ sinh của loài gây hại đích, sau đó ủ ở nhiệt độ phòng trong 5 ngày. Các đĩa lá từ cây chuyển gen biểu hiện BT-0001 được thử đối với sâu đục thân ngô châu

Ấu (*Ostrinia nubilalis*), sâu xanh hại ngô (*Helicoverpa zea*) và sâu xám (*Agrotis ipsilon*). Các đĩa lá từ cây chuyển gen biểu hiện BT-0003 hoạt tính đối với ít nhất Sâu đục thân ngô châu Á (*Ostrinia furnacalis*) và sâu xám (*Agrotis ipsilon*).

Các kết quả của sinh thử nghiệm mô thực vật chuyển gen xác nhận rằng các protein Cry theo sáng chế khi được biểu hiện trong cây chuyển gen là độc đối với côn trùng đích. Ví dụ, BT-0001 được biểu hiện ở ngô được biến nạp ổn định với gen kháng theo sáng chế hoạt tính đối với ít nhất sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xanh hại ngô (*Helicoverpa zea*) và sâu xám (*Agrotis ipsilon*). BT-0003 hoạt tính đối với ít nhất Sâu đục thân ngô châu Á (*Ostrinia furnacalis*) và sâu xám (*Agrotis ipsilon*).

Danh mục trình tự

Bản sao chính chức của danh mục trình tự được nộp điện tử qua trang EFS-Web là danh mục trình tự có định dạng ASCII với tên tập tin là "80829-US-L-ORG-NAT-1_SeqList_ST25.txt", tạo ngày 30 tháng sáu năm 2015, và có kích thước bằng 310 kilobytes và được nộp đồng thời với bản mô tả. Danh mục trình tự chứa trong tài liệu có định dạng ASCII này là một phần của bản mô tả và được kết hợp ở đây bằng cách viện dẫn toàn bộ.

Tham khảo chéo đến đơn liên quan

Đơn này xin hưởng quyền ưu tiên theo Đơn Mỹ Số 62/187,468, nộp ngày 1 tháng bảy năm 2015, được kết hợp ở đây bằng cách viện dẫn toàn bộ cho mọi mục đích.

YÊU CẦU BẢO HỘ

1. Phương pháp kiểm soát côn trùng cánh vảy kháng Cry1Ab bao gồm bước phân phối lượng có hiệu quả của protein diệt côn trùng chứa SEQ ID NO:31 lên côn trùng kháng Cry1Ab.

2. Phương pháp theo điểm 1, trong đó côn trùng kháng Cry1Ab là sâu đục thân mía (*Diatraea saccharalis*).

3. Phương pháp ngăn ngừa sự phát triển sức kháng trong quần thể sâu đục thân mía (*Diatraea saccharalis*) đối với protein Cry1Ab được biểu hiện ở cây chuyển gen bao gồm xếp chồng protein diệt côn trùng thứ hai chứa SEQ ID NO:31 lên cây chuyển gen ngoài protein Cry1Ab.

DANH MỤC TRÌNH TỰ

<110> Syngenta Participations AG
 Bramlett, Matthew Richard
 Seguin, Katherine
 Rose, Mark Scott

<120> PHƯƠNG PHÁP KIỂM SOÁT CÔN TRÙNG CÁNH VÂY KHÁNG CRY1AB VÀ PHƯƠNG
 PHÁP NGĂN NGỪA SỰ PHÁT TRIỂN SỨC KHÁNG TRONG QUẦN THỂ SÂU ĐỤC THÂN MÍA
 ĐỐI VỚI PROTEIN CRY1AB

<130> 80829-US-L-ORG-NAT-1

<160> 50

<170> PatentIn phiên bản 3.5

<210> 1
 <211> 3546
 <212> ADN
 <213> Bacillus thuringiensis

<400> 1
 atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa
 60
 gtagaagtat taggtggaga aagaatagaa actgggttaca cccaatcga tatttccttg
 120
 tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta
 180
 gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt
 240
 gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta
 300
 gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat
 360
 cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc
 420
 cttacaaccg ctattcctct ttttgcagtt caaaattatc aagttcctct tttatcagta
 480
 tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa
 540
 aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt
 600
 ggcaactata cagattatgc tgtgcgctgg tacaatacgg gattagagcg tgtatgggga
 660
 ccggattcta gagattgggt aagggtataat caatttagaa gagagctaac acttactgta
 720

ttagatatcg ttgctctatt ctcaaattat gatagtcgaa ggtatccaat tcgaacagtt
 780
 tcccaattaa caagagaaat ttatacgaac ccagtattag aaaattttga tggtagtttt
 840
 cgtggaatgg ctgagagaat agaacagaat attaggcaac cacatcttat ggatatcctt
 900
 aatagtataa ccattttatac tgatgtgcat agaggcttta attattggtc agggcatcaa
 960
 ataacagctt ctctgtagg gttttcagga ccagaattcg cattcccttt atttggaat
 1020
 gcggggaatg cagctccacc cgtacttgtc tcattaactg gtttggggat ttttagaaca
 1080
 ttatcttcac ctttatatag aagaattata cttgggtcag gcccaaataa tcaggaactg
 1140
 tttgtccttg atggaacgga gttttctttt gcctccctaa cgaccaactt gccttcact
 1200
 atatatagac aaaggggtac agtcgattca ctagatgtaa taccgccaca ggataatagt
 1260
 gtaccacctc gtgogggatt tagccatcga ttgagtcatg ttacaatgct gagccaagca
 1320
 gctggagcag tttacacctt gagagctcca acgttttctt ggcagcatcg cagtgctgaa
 1380
 ttttaataata taattgcac ggatagtatt actcaaatcc ctgcagtgaa gggaaacttt
 1440
 ctttttaatg gttctgtaat ttcaggacca ggatttactg gtggggactt agttagatta
 1500
 aatagtagtg gaaataacat tcagaataga gggatatattg aagttccaat tcacttccca
 1560
 tcgacatcta ccagatatcg agttcgtgta cggtatgctt ctgtaacccc gattcacctc
 1620
 aacgttaatt ggggtaattc atccattttt tccaatacag taccagctac agctacgtca
 1680
 ttagataatc tacaatcaag tgattttggt tattttgaaa gtgccaatgc ttttacatct
 1740
 tcattaggta atatagtagg tgttagaaat tttagtggga ctgcaggagt gataatagac
 1800
 agatttgaat ttattccagt tactgcaaca ctcgaggctg aatataatct ggaaagagcg
 1860

cagaaggcgg tgaatgcgct gtttacgtct acaaaccaac tagggctaaa aacaaatgta
 1920
 acggattatc atattgatca agtgtccaat ttagttacgt gtttatcgga tgaattttgt
 1980
 ctggatgaaa agcgagaatt gtccgagaaa gtcaaacaatg cgaagcgact cagtgatgaa
 2040
 cgcaattttac tccaagattc aaattttcaaa gacattaata ggcaaccaga acgtgggtgg
 2100
 ggcggaagta cagggattac catccaagga ggggatgacg tatttaaaga aaattacgtc
 2160
 acactatcag gtacctttga tgagtgtat ccaacatatt tgtatcaaaa aatcgatgaa
 2220
 tcaaaattaa aagcctttac ccgttatcaa ttaagagggt acatcgaaga tagtcaagat
 2280
 ttagaagttt atttgatccg ttacaatgca aaacacgaaa cgttaaactg gccaggtacg
 2340
 ggttccttat ggccacttgc agttaaaagt ccaattggaa ggtgcggtga accgaatcga
 2400
 tgtgcaccac ggattgagtg gaaacctgat gtagattgtt cctgcagaga cggagaaaaa
 2460
 tgtgcgcac attcccatca tttctccttg gacattgatg taggatgtac agacttaaat
 2520
 gaggatttag gcgtatgggt gatattcaag attaagacac aagatggcca tgcgaaaata
 2580
 ggaaatctag aatttctcga agagaagctt ttattaggag aagcattagc acgtgtgaag
 2640
 aaagcggaga aaaaatggag agacaaacgc gaaaaattgg aatgggaaac aaatattgtt
 2700
 tataaagagg caaaagaatc tgtagatgct ttattcgtag attctcaata taatagatta
 2760
 caaacggata cgaacattgc gatgattcat gcggcagata aacgcgttca tcgaatccga
 2820
 gaagcgtatt tgccagagtt atctgtgatt ccgggtgtca atgcggctat tttcgaagaa
 2880
 ttagaaggtc ttattttcac cgcattctcc ctatatgatg cgagaaatgt cattaaaaac
 2940
 ggagatttca attatggttt atcatgctgg aatgtgaaag ggcattgata tgtagaagaa
 3000

caaaacaacc accgttccgt ccttgttatc ccagaatggg aagcagaagt gtcccaagaa
3060

gttcgtgtct gtccaggctg tggctatatc cttcgtgtta cagcgtacaa agagggatat
3120

ggagagggct gcgtaacgat ccatgagatc gaagacaata cagacgaact gaaattcagc
3180

aactgtgtag aagaggaagt atatccaaac aacacggtaa cgtgtaatga ttatactgcg
3240

actcaagaag aatatgaggg tacgtacact tctcgtaatc gaggatatga cggagcctat
3300

gaaagcaatt cttctgtacc agctgattat gcatcagcct atgaagaaaa agcgtataca
3360

gatggacgaa gagacaatcc ttgtgaatct aacagaggat atagggatta cacaccacta
3420

ccagctggct atgtgacaaa agaattagag tacttcccag aaaccgataa ggtatggatt
3480

gagatcggag aaacggaagg aacattcatt gtggatagcg tggaattact ccttatggag
3540

gaatag
3546

<210> 2
<211> 3531
<212> ADN
<213> *Bacillus thuringiensis*

<400> 2
atggagatag tgaataatca gaatcaatgc gtgccttata attgtttgaa taatcccgaa
60

atcgaaatat tagaaggcgg aagaatatca gttggtaata cccaattga tatttctctt
120

tcgcttactc agtttctttt gagtgaattt gtcccaggctg cggggtttgt attaggatta
180

attgatttaa tatggggatt tgtaggtcct tccaatggg acgcatttct tgctcaagtg
240

gaacagttaa ttaaccaaag aatagcagaa gctgtaagaa atacagcaat tcaggaatta
300

gagggaatgg cacgggttta tagaacctat gctactgctt ttgctgagtg ggaaaaagct
360

cctgatgacc cagagctaag agaagcacta cgtacacaat ttacagcaac tgagacttat
420

ataagtggaa gaatatccgt tttaaaaatt caaacttttg aagtacagct gttatcagtg
 480
 tttgccaag ctgcaaattt acatttatct ttattaagag acgttggtt ttttgggcaa
 540
 agatgggggtt tttcaacgac aaccgtaa atactactaca atgatttaac agaagggatt
 600
 agtacctata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga
 660
 ccggattcta gagattgggt aagggtataat caatttagaa gagagctaac acttactgta
 720
 ttagatatcg ttgctctatt ctcaaattat gatagtagaa gatatccaat tcgaacagtt
 780
 tcccaattaa caagagaaat ttatacgaac ccagtattag aaaattttga tggtagtttt
 840
 cgtggaatgg ctacagagaat agaacagaat attaggcaac cacatcttat ggatatcctt
 900
 aatagtataa ccatttatac tgatgtgcat agaggcttta attattggtc agggcatcaa
 960
 ataacagctt ctctgtagg gttttcagga ccagaattcg cattcccttt atttggaat
 1020
 gcggggaatg cagctccacc cgtacttgtc tcattaactg gtttggggat ttttagaaca
 1080
 ttatcttcac ctttatatag aagaattata cttgggtcag gcccaaataa tcaggaactg
 1140
 tttgtccttg atggaacgga gttttctttt gcctccctaa cgaccaactt gccttcact
 1200
 atatatagac aaaggggtac agtcgattca ctagatgtaa taccgccaca ggataatagt
 1260
 gtaccacctc gtgcgggatt tagccatcga ttgagtcag ttacaatgct gagccaagca
 1320
 gctggagcag tttacacctt gagagcccca acgttttctt ggcagcatcg cagtgtgaa
 1380
 tttaataata taattccttc atcacaaatt acacaaatac ctttaacaaa atctactaat
 1440
 cttggctctg gaacttctgt cgttaaagga ccaggattta caggaggaga tattcttcga
 1500
 agaacttcac ctggccagat ttcaacctta agagtaaata ttactgcacc attatcacia
 1560

agatatcggg taagaattcg ctacgcttct actacaaatt tacaattcca tacatcaatt
 1620
 gacggaagac ctattaatca gggtaatttt tcagcaacta tgagtagtgg gagtaattta
 1680
 cagtccggaa gctttaggac tgtaggtttt actactccgt ttaacttttc aaatggatca
 1740
 agtgtattta cgttaagtgc tcatgtcttc aattcaggca atgaagtta tatagatcga
 1800
 attgaatttg ttccggcaga agtaaccttt gaggcagaat atgatttaga aagagcacia
 1860
 aaggcgggtga atgagctggt tacttcttcc aatcaaactg ggtaaaaaac agatgtgacg
 1920
 gattatcata ttgatcaagt atccaattta gttgagtgtt tatcagatga attttgtttg
 1980
 gatgaaaaac aagaattgtc cgagaaagtc aaacatgcga agcgacttag tgatgagcgg
 2040
 aatttacttc aagatccaaa cttcagaggg atcaatagac aactagaccg tggctggaga
 2100
 ggaagtacgg atattacat ccaaggaggc gatgacgtat ttaaagagaa ttacgttacg
 2160
 ctattgggta cctttgatga gtgctatcca acgtatttat atcaaaaaat agatgagtcg
 2220
 aaattaaag cctatacccg ttatcaatta agagggtata tcgaagatag tcaagactta
 2280
 gaaatctatt taattcgcta caatgcaaaa catgaaacag taaatgtgcc aggtacgggt
 2340
 tccttatggc cgctttcagc ccaaagtcca atcggaagt gtggagagcc gaatcgatgc
 2400
 gcgccacacc ttgaatggaa tcctgactta gattgttcgt gtagggatgg agaaaagtgt
 2460
 gcccatcatt cgcacattt ctcttagac attgatgtag gatgtacaga cttaaagag
 2520
 gacctaggtg tatgggtgat ctttaagatt aagacgcaag atgggcacgc aagactaggg
 2580
 aatctagagt ttctcgaaga gaaaccatta gtaggagaag cgctagctcg tgtgaaaata
 2640
 gcggagaaaa aatggagaga caaacgtgaa aaattggaat gggaacaaa tatcgtttat
 2700

aaagaggcaa aagaatctgt agatgcttta tttgtaaact ctcaatatga tcaattacaa
2760

gcggatacga atattgccat gattcatgcg gcggataaac gtgttcatag cattcgagaa
2820

gcttatctgc ctgagctgtc tgtgattccg ggtgtcaatg cggctatttt tgaagaatta
2880

gaagggcgta ttttactgc attctcccta tatgatgcga gaaatgtcat taaaaatggt
2940

gattttaata atggcttata ctgctggaac gtgaaagggc atgtagatgt agaagaacaa
3000

aacaaccaac gttcggctct tgttggtccg gaatgggaag cagaagtgtc acaagaagtt
3060

cgtgtctgtc cgggtcgtgg ctatatcctt cgtgtcacag cgtacaagga gggatatgga
3120

gaaggttgcg taaccattca tgagatcgag aacaatacag acgaactgaa gtttagcaac
3180

tgcgtagaag aggaaatcta ttcaaataac acggtaacgt gtaatgatta tactgtaaat
3240

caagaagaat acggaggtgc gtacacttct cgtaatcgag gatataacga agctccttcc
3300

gtaccagctg attatgcgtc agtctatgaa gaaaaatcgt atacagatgg acgaagagag
3360

aatccttgty aatttaacag agggatatagg gattacacgc cactaccagt tggttatgtg
3420

acaaaagaat tagaatactt ccagaaacc gataaggtat ggattgagat tggagaaacg
3480

gaaggaacat ttatcgtgga cagcgtggaa ttacttctta tggaggaata g
3531

<210> 3

<211> 3531

<212> ADN

<213> *Bacillus thuringiensis*

<400> 3

atggagatag tgaataatca gaatcaatgc gtgccttata attgtttgaa taatcccgaa
60

atcgaaatat tagaaggcgg aagaatatca gttggtaata cccaattga tatttctctt
120

tcgcttactc agtttctttt gagtgaattt gtcccagggtg cgggggtttgt attaggatta
180

attgatttaa tatggggatt tgtaggtcct tcccaatggg acgcatttct tgctcaagtg
240

gaacagttaa ttaaccaaag aatagcagaa gctgtaagaa atacagcaat tcaggaatta
300

gagggaatgg cacgggttta tagaacctat gctactgctt ttgctgagtg ggaaaaagct
360

cctgatgacc cagagctaag agaagcacta cgtacacaat ttacagcaac tgagacttat
420

ataagtggaa gaatatccgt tttaaaaatt caaacttttg aagtacagct gttatcagtg
480

tttgcccaag ctgcaaattt acatttatct ttattaagag acgttgtggt ttttgggcaa
540

agatgggggtt tttcaacgac aaccgtaaatt aattactaca atgatttaac agaagggatt
600

agtaacctata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga
660

ccggattcta gagattgggt aaggtataat caatttagaa gagagctaac acttactgta
720

ttagatatcg ttgctctatt ctcaaattat gatagtagaa gatatccaat tcgaacagtt
780

tcccaattaa caagagaaat ttatacgaac ccagtattag aaaattttga tggtagtttt
840

cgtggaatgg ctacagagaat agaacagaat attaggcaac cacatcttat ggatatecct
900

aatagtataa ccattttatac tgatgtgcat agaggcttta attattggtc agggcatcaa
960

ataacagctt ctctgtagg gttttcagga ccagaattcg cattcccttt atttgggaat
1020

gcggggaatg cagctccacc cgtacttgct tcattaactg gtttggggat ttttagaaca
1080

ttatcttcac ctttatatag aagaattata cttgggttcag gcccaaataa tcaggaactg
1140

tttgtccttg atggaacgga gttttctttt gcctccctaa cgaccaactt gccttccact
1200

atatatagac aaaggggtac agtcgattca ctagatgtaa taccgccaca ggataatagt
1260

gtaccacctc gtgcgggatt tagccatcga ttgagtcatg ttacaatgct gagccaagca
1320

gctggagcag ttacacctt gagagcccca acgttttctt ggcagcatcg cagtgctgaa
1380

tttaataata taattccttc atcacaaatt acacaaatac ctttaacaaa atctactaat
1440

cttggctctg gaacttctgt cgttaaagga ccaggattta caggaggaga tattcttcga
1500

agaacttcac ctggccagat ttcaacctta agagtaaata ttactgcacc attatcacia
1560

agatatcggg taagaattcg ctacgcttct actacaaatt tacaattcca tacatcaatt
1620

gacggaagac ctattaatca gggtaatttt tcagcaacta tgagtagtgg gagtaattta
1680

cagtccggaa gctttaggac tgtaggtttt actactccgt ttaacttttc aaatggatca
1740

agtgtattta cgttaagtgc tcatgtcttc aattcaggca atgaagtta tatagatcga
1800

attgaatttg ttccggcaga agtaaccttt gaggcagaat atgatttaga aagagcacia
1860

aaggcgtga atgagctggt tacttcttcc aatcaaactc ggttaaaaac agatgtgacg
1920

gattatcata ttgatcaagt atccaattta gttgagtgtt tatcagatga attttgtttg
1980

gatgaaaaac aagaattgtc cgagaaagtc aaacatgcga agcgacttag tgatgagcgg
2040

aatttacttc aagatccaaa cttcagaggg atcaatagac aactagaccg tggctggaga
2100

ggaagtacgg atattaccat ccaaggaggc gatgacgtat ttaaagagaa ttacgttacg
2160

ctattgggta cctttgatga gtgctatcca acgtatttat atcaaaaaat agatgagtcg
2220

aaattaaaag cctatacccg ttatcaatta agaggggtata tcgaagatag tcaagactta
2280

gaaatctatt taattcgcta caatgcaaaa catgaaacag taaatgtgcc aggtacgggt
2340

tccttatggc cgctttcagc ccaaagtcca atcggaaagt gtggagagcc gaatcgatgc
2400

gcgccacacc ttgaatggaa tcctgactta gattgttcgt gtagggatgg agaaaagtgt
2460

gcccattcatt cgcattcattt ctcttagac attgatgtag gatgtacaga cttaaagtga
2520

gacctagggtg tatgggtgat ctttaagatt aagacgcaag atgggcacgc aagactaggg
2580

aatctagagt ttctcgaaga gaaaccatta gtaggagaag cgctagctcg tgtgaaaata
2640

gcggagaaaa aatggagaga caaacgtgaa aaattggaat gggaaacaaa ttcgtttat
2700

aaagaggcaa aagaatctgt agatgcttta ttgttaaact ctcaatatga tcaattacaa
2760

gcggatacga atattgccat gattcatgcg gcggataaac gtgttcatac cattcgagaa
2820

gcttatctgc ctgagctgtc tgtgattccg ggtgtcaatg cggctatctt tgaagaatta
2880

gaagggcgta ttttcaactgc attctcccta tatgatgcga gaaatgtcat taaaaatggt
2940

gattttaata atggcttatc ctgctggaac gtgaaagggc atgtagatgt agaagaacaa
3000

aacaaccaac gtccggtcct tgttggtccg gaatgggaag cagaagtgtc acaagaagtt
3060

cgtgtctgtc cgggtcgtgg ctatatcctt cgtgtcacag cgtacaagga gggatatgga
3120

gaaggttgcg taaccattca tgagatcgag aacaatacag acgaactgaa gtttagcaac
3180

tgcgtagaag aggaaatcta ttcaaataac acggtaacgt gtaatgatta tactgtaaat
3240

caagaagaat acggaggtgc gtacacttct cgtaatcgag gatataacga agctccttcc
3300

gtaccagctg attatgcgtc agtctatgaa gaaaaatcgt atacagatgg acgaagagag
3360

aatccttggtg aatttaacag agggatatagg gattacacgc cactaccagt tggttatgtg
3420

acaaaagaat tagaatactt cccagaaacc gataaggtat ggattgagat tggagaaacg
3480

gaaggaacat ttatcgtgga cagcgtggaa ttacttctta tggaggaata g
3531

<210> 4
<211> 2148
<212> ADN

<213> Bacillus thuringiensis

<400> 4

atgaaatcta agaatcaaaa tatgcatcaa agcttgtcta acaatgcgac agttgataaa
60

aactttacag gttcactaga aaataacaca aatacggaaat tacaaaactt taatcatgaa
120

ggtatagagc cgtttggttag tgtatcaaca attcaaacgg gtattggtat tgctggtaaa
180

atccttggtta acctaggcgt tccttttgct gggcaagtag ctagcctcta tagttttatc
240

ctaggtgagc tttggcccaa agggaaaagc caatgggaaa tctttatgga acatgtagaa
300

gagcttatta atcaaaagat atcgacttat gcaagaaaca aagcacttgc agatttaaaa
360

ggattaggag atgctttggc tgtctacat gaatcgctgg aaagttggat taaaaatcgc
420

aataacacaa gaactagaag tgttgtcaag agccaatata ttaccttgga acttatgttc
480

gtacaatcat taccttcttt tgcagtgtct ggagaggaag taccactatt accaatatat
540

gctcaagctg caaatctaca cttgttgcta ttaagagatg cgtctatctt tggaagaa
600

tggggattat cagactcaga aatttcgaca ttctataatc gtcaagtgga aagaacatca
660

gattattccg atcattgcac gaaatgggtt gatacgggct tgaatagatt aaagggtca
720

aatgctgaaa tctgggtaaa gtataatcaa ttccgtagag acatgacttt aatgggtacta
780

gatttagtgg cactattcca aagctatgat acacatatgt acccaattaa aactacagcc
840

caacttacta gagaagtata tacaaacgca attgggacag tacatccgca cccaagtttt
900

acaagtacga cttggtataa taataatgca ccttcgttct ctgccataga ggctgccgtt
960

atccgaagcc cgcacctact cgattttcta gaacaagtta caatttacag cttattaagt
1020

cgatggagta acactcagta tatgaatatg tggggaggac ataaactaga attccgaaca
1080

ataggaggaa cgttaaatac ctcaacacaa ggatctacta atacttctat taatcctgta
1140

acattaccgt tcacgtctcg agacatctat aggactgaat cattggcagg gctgaatcta
1200

tttttaactc aacctgttaa tggagtacct agggttgatt ttcattggaa attcgtcaca
1260

catccgatcg catctgataa tttctattat ccagggtatg ctggaattgg gacgcaatta
1320

caggattcag aaaatgaatt accacctgaa acaacaggac agccaaatta tgaatcttat
1380

agtcatagat tatctcatat aggactcatt tcagcatcac atgtgaaagc attggtatat
1440

tcttggacgc atcgtagtgc agatcgtacg aatacaattc attcagatag tataacacaa
1500

ataccactgg taaaagcaca tacccttcag tcaggacta ctgttgtaaa agggccaggg
1560

tttacagggtg gagatatacct ccgacgaact agtggaggac catttgcttt tagtaatgtt
1620

aatttagact ggaacttgtc acaaagatat cgtgctagaa tacgctatgc ttctactact
1680

aatctaagaa tgtacgtaac gattgcaggg gaacgaattt ttgctggtca atttaataaa
1740

acaatgaata ctggtgatcc attaacattc caatctttta gttacgcaac tattgataca
1800

gcatttacat tcccaacgaa agcgagcagc ttgactgtag gtgctgatac ttttagctca
1860

ggtaatgaag tttatgtaga tagatttgaa ttgatcccag ttactgcaac acttgaggca
1920

gtaactgatt tagaaagagc gcagaaggcg gttcatgaac tgtttacatc tacgaatccg
1980

ggaggattaa aaacggatgt agccaaagat cattacacaa ataccattag taaaagcgtt
2040

caatctgtct tcagggtgccg ctgtagtgag aggaccagga tttacagggtg gggatatacct
2100

tcgaagaaag aatactggta catttgggga tatacgagta aatattaa
2148

<210> 5
<211> 3285
<212> ADN

<213> Bacillus thuringiensis

<400> 5

atgtttcttg aacaaattga acaactaatt gaccagagaa tagagacagt cgaaaggaat
60

agggcaattc aaacattaat tgggttatca aatagttatg atgtatatat tgaagcgtaa
120

aaagaatggg aaaataatcc tgataattca gcttcacaag aaagagtacg taatcgcttt
180

cggacaactg acgatgcttt gataactggc attcctcttt tagccattcc gaattttgaa
240

atagctactt tatcggtgta tgttcaagct gccaatctac atttatcctt attaagggat
300

gcagtgtttt tcggagaaag atggggatta acacaaataa atgtagatga cttgtacagg
360

agattaacga ataatatcag aaactattca gatcattgtg cacgatggta taatgaagga
420

ttagataata tttctggttt atctcgatct attaacttcc aaagagaagt aacaatctct
480

gtcttagata ttgttgcgct tttcccgaac tatgacatcc gaacatatcc aatttcaaca
540

acaagccaat taacaaggga gatattcaca tctccaattg ttgtccctaa tgattttagt
600

gtagcctacg agggggtaag gagagcgcca cacctatttg aattttttaga gaaacttggt
660

atttataccg gtgatcgaag tgggattcgc cattgggcgg gacatgaaat aacttctaga
720

cgtactgatt cataccacgg tataattcgt taccctcttt atggaacagc agcaaagca
780

gaaagtccat atactcttgc cettcaacct tctgaaagta tttatagaac gttatcagaa
840

cctatatattt cacaaactgg tgggctgtct cctcatagaa ggagagtagt agaggggagta
900

gagttctcta ttgtaaataa taacgtaa at cttcgtcat ttgtatatag aagaaaggg
960

tcgttagatt cttttactga gttaccacct gaagatgaaa gtgtaccacc ttatattggc
1020

tacagtcatc aattatgcca tgttggtatt ggtcgtacaa atgtaatctt tgaaccaagt
1080

aatttcgcta gggttccagt attctcctgg acacatcgta gtgcaacccc tacaaataca
1140

attgatccag atagaattac ccaaatacct tcagtgaagg cgagttctct tcgtaattct
1200

actgttgta gtggaccagg atttactgga ggggatattg ttcgaatggg agcagtgcac
1260

caaatatatg caacggattht aagtatgaat gttcgaccta gtgttgcatg gagcagatat
1320

cttataagac ttcgctatgc ttgtaggggg agttcaaaca tagttataca cggtccttct
1380

attagatttg tatcgctccc aagtacaatg agtaatgatg aacctttaac atatcaatca
1440

tttagatacg caagtatcac aactccaatt acccggtccaa tatataacat gtttaattta
1500

tctatatcca gaatttcagg tgtccaaaat ttgtttatag atcgaataga attcattcca
1560

gtagatgcaa acttcgaagc agaacgagat ttagagagag cgcagaaggc ggtgaatgct
1620

ctgtttactt ccacaaacca aagaggacta aaaatagatg tgactgacta tcatattgat
1680

caagtatcca atttagttga ttgcttatcg gatgaatttt gtctggatga aaagcgagaa
1740

ttgtccgaaa aatccaaaaca tgcgaagcga ctcatgatg aacgcaattt actccaagat
1800

ttaaatttca aagacattaa taggcaacca gaacgtgggt ggagcggaag tacagggatt
1860

accatccaag gaggagatga cgtattcaaa gagaattacg ttacactacc aggtaccttt
1920

gatgagtgct atccaacata tttgtatcaa aaaatcgatg aatcaaaatt aaaagcctat
1980

acccgctatc aattaagagg atacatcgaa gatagtcaag acttagaaat ctatttaatt
2040

cgctacaatg caaaacatga gacagtaaat gttcctggct ctggctcctt atggccactt
2100

tcagtcgaaa gctcagttgg aaaatgcgga gagccaaatc gatgtgcac acggatggaa
2160

tggaatcccc acctagattg ctctgttagg gatggggaga agtgtgcccc tcattcccat
2220

catttctctt tggacattga tgtaggatgt acagacttaa atgaggattt aggtgtatgg
2280

gtgatattca agattaagac acaagatggc catgcgaaaa taggaaatct agaatttctc
2340

aaagagaagc tgttattagg agaagcatta gcacgtgtga agaaagcggg gaaaaaatgg
2400

agagacaaac gcgacaaatt ggaatgggaa acaaagtgtg tttataaaga ggcaaaagaa
2460

tctgtagatg ctttattcgt agattctcaa tatagtagat tacaagcggg tacgaacatc
2520

gcaatgattc atgcggcaga taaacgcgtt catcgaattc gagaagcgta tctcccagaa
2580

ctaactgtga ttccaggtgt caatgcatct attttcgaag aattagaggg acgtattttc
2640

acagcgtatt ccctatatgg tgcaagaaat gtcataaaaa atggcgattt caataatggt
2700

ttatcttgct ggaacgtgaa agggcatgta gaggtacaac agattcatca tcgttcggtt
2760

cttgttgttc caagttggaa aacagaagta tcacaagagg tgtgctctg tccaggacgt
2820

ggctatatcc ttctgtttac agcgtacaaa gaaggatatg gagaaggcaa cgtaaccatc
2880

catgagatcg agaacaatac agatgaactg aagtttagaa actgtgaaga agaggaagtc
2940

tatccaaaca acacagtaac gtgtaatgat tatactgtaa atcaagaaga atataagggt
3000

acgtgcactt ctctgtaatcg aggatatgac gaatcctatg aaagcagctc ttccgaatca
3060

gcttattatg cttcagtcta tgaagaaaaa gggatatacag atggacgaag agagaatctt
3120

tgtgaattta atagggggta tggggattac acgtcactac caactgctta tgtgacaaaa
3180

gaattagagt acttcccaga aaccgataag gtatggattg agattggaga aacagaagga
3240

gcattcatcc tggacagcgt ggaattactc cttatggagg aataa
3285

<210> 6
<211> 3510
<212> ADN

<213> *Bacillus thuringiensis*

<400> 6

atggagatta ataatcagaa ccaatgtgtc ccttataatt gtttgaataa tcctgaaagc
60

gagatattaa acgttgcaat ttttagtagc gaacaggtag cagaaattca cttaaagatc
120

acgcgcttaa ttttagagaa ttttttacca ggtgggagtt ttgcattcgg cttatttgat
180

ttaatatggg ggatttttaa tgaagatcaa tggagcgcac ttcttcggca ggtagaagaa
240

ttaattaatc aaaggataac ggaattcgca agagggaag caattcagag actagtaggg
300

tttgaagga gttatgatga atatatttta gcactaaaag aatgggaaaa cgatcctgat
360

aaccagctt caaaggaaag agtgcgcact cgatttcgga caactgatga tgccttgcta
420

accggtgttc ctcttatggc aattccaggt tttgaattag ctactttatc tgtttatgct
480

caatcagcca atctacattt agccctatta agagatgctg tatttttttg ggagagatgg
540

ggattgacac aaacaaatat aaatgattta tatagtagat taaaaaactc cattcgtgat
600

tatacaaatc attgtgttcg tttttataat ataggtttag ggaatttaaa tgttataaga
660

ccagagtatt accgtttcca aagagaatta acaatatctg tcttagatct tgtagctctt
720

tttccaaatt acgatatccg aacatatcca ataccaacta aaagtcaatt aacaagagaa
780

atttatacag atccgattat ttcacctggc gcacaggcag gttatactct tcaagatggt
840

ttgagagAAC cacaccttat ggacttttta aaccgactta ttatttatac tggtagatg
900

cgcggaattc gtcactgggc aggacatgaa gtagaatcta gtagaacagg tatgatgact
960

aatataagat ttcctttgta tggaacagcc gcaacagcag aaccaacacg atttataact
1020

cctagtactt ttctgtgtct taatttattt tatagaacat tatcagctcc tatttttaga
1080

gatgaaccgg gagctaatat tattattaga tatagaacga gtttggtgga aggagtagga
1140

tttattcaac caaataacgg tgaacagctt tacagagtga gaggaacatt agattctctt
1200

gatcaattac cacttgaggg tgagagtagt ctaactgaat atagtcacg attatgccat
1260

gttagatttg cgcaatcatt gaggaatgca gaacctttag attatgcaag ggttccgatg
1320

ttttcttgga cacatcgtag tgcaaccctt acaaatacaa ttgatccaga tgatcatcacc
1380

caaataccgt tagtaaaggc tttcaatctt cattcaggtg ccacgattgt taaaggacca
1440

ggtttttacag gtggggatat ctttcgaaga acgaatgttg gtagcttttg agatatgcgt
1500

gtaaacatta ctgcaccact atcacaaaga tatcgcgtaa ggattcgta tgcttctacg
1560

acagattttac aattctatac gaatattaat ggaactacta ttaatattgg taatttctcg
1620

agcactatgg acagtgggga tgattttacag tacggaagat tcagggttgc aggttttact
1680

actccattta ctttttcaga tgcaatgagc acattcaca taggtgcttt tagcttctct
1740

tcaaacaacg aagtttatat agatcgaatt gaatttgtcc cggcagaagt aacatttgag
1800

gcagaatatg atttagagaa agctcagaaa gcggtgaatg cgctgtttac ttcttccaat
1860

caaatcgggt taaaaacaga tgtgacggac tatcatattg ataaagtatc caatctagtt
1920

gagtgtttat cagatgaatt ttgtctagat gaaaagcgag aattgtccga gaaagtcaaa
1980

catgcgaagc gactctgtga tgagcggaa ttacttcaag atccaaactt cagaggcatc
2040

aatagacaac cagaccgtgg ttggagagga agtacggata ttaccatcca aggaggagat
2100

gacgtattca aagagaatta cgttacgcta cgggtacct ttgatgagtg ctatccaacg
2160

tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct ataccgcta tgaattaaga
2220

gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacat
2280

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagccca aagtccaatc
2340

ggaaagtgtg gagaaccgaa tcgatgtgcg acacaccttg aatggaatcc tgacttagat
2400

tgttcgtgta gggatggaga aaagtgtgcc catcattegc atcatttctc cttagacatt
2460

gatgtaggat gtacagacct aaatgaggac ctaggtgtat gggatgatctt taagattaag
2520

acgcaagatg gtcatgcgag actaggaaat ctagaatttc tcgaagagaa accattagta
2580

ggagaagcgc tagctcgtgt gaagagagcg gagaaaaaat ggagagacaa acgcgaaaaa
2640

ttggaattgg aaacaaatat tgtttataaa gaggcaaaaa aatctgtaga tgctttatct
2700

gtgaactctc aatatgatag attacaagcg gatacgaata tcgcgataat tcatgcggca
2760

gataaacgcg ttcatagcat tcgagaagca tatcttccag agttgtctgt aattccgggt
2820

gtaaatgcag ctatTTTTga agaattagag ggacgtattt tcacagccta ctctctatat
2880

gatgcgagaa atgtcattaa aaatggcgat ttcaataatg gcttatcatg ctggaacgtg
2940

aaagggcatg tagatgtaga agaacagaac aaccatcggt cggtccttgt tgttccagaa
3000

tgggaagcag aagtgtcaca agagggtcgt gtctgtccag gtcgtggcta tatccttcgt
3060

gttacagcgt acaaagaggg atatggagag ggctgtgtaa cgattcatga gatcgaagac
3120

aatacagacg aactgaaatt cagcaactgt gtagaagagg aaatatatcc aaacaacacg
3180

gtaacgtgta atgattatac tgcgactcaa gaagaatatg agggtagta cacttctcgt
3240

aatcgaggat atgacggagc ctatgaaagc aattcttctg taccagctga ttatgcatca
3300

gcctatgaag aaaaagcgta tacagatgga agaagagaca atacttgtga atctaacaga
3360

ggatatggggg attacacacc actaccagct ggctatgtga caaaagaatt agagtacttc
3420

ccagaaaccg ataaggtatg gattgagatt ggagaaacgg aaggaacatt tatcgtggac
3480

agcgtggaat tactccttat ggaggaatag
3510

<210> 7

<211> 3414

<212> ADN

<213> *Bacillus thuringiensis*

<400> 7

atggaagtaa atcatcaaaa tgaatgtgta ccctataatt gtttaaagaa tcccaagatt
60

gagatgttag atatcgaggg aataagtagt agaagtagag aacaagtagc ggaaattagc
120

ttagggctca cgcgctttct tttggagtca ctgctccctg gtgcaagttt tggatttggt
180

ttatttgata tcatttgggg agttattggt cctgatcaat ggagcctatt tcttacacaa
240

attgaacaat taattgatca acggatagag gcacatgtaa gaaatcaagc gattagtaga
300

ttagaagggc ttggagatag ttatgaagtg tatattgagt cgcttagaga atgggagggc
360

tctcctaata atgaatcctt gcagcaggat gtgcgtaatc gatttagtaa taccgataat
420

gccttgatta ccgcgatacc tctcctgaga gaacagggct ttgaaattcc tcttttaacg
480

gtttatgttc aggctgccaa tttgcattta tccttattaa gggatgccgt atattttgga
540

cagagatggg gacttgatac agctacagta aacaatcatt ataatcgatt aattaatctt
600

attaacactt attccgatca ttgtgctcaa tggttcaata gaggtctgga taattttggt
660

gttgtaacag caagatattt agatttccaa agagaggtaa caatatctgt cttagatatc
720

gttgctcttt tcccgaatta tgatattcgg acatatccaa tccaaacact aagtcaatta
780

acaagagaaa tttatacatc tccagttgct gaacctggtg caagtttaaa tgtggatttg
840

cggaatatat tgagagaacc tcacctaattg gatttttttaa ctgccttgt tatttatacg
900

ggtgtgcagg gtggtattta tcattgggca ggacatgaga tatcttctag gactacaggt
960

aatttatcga gcaacatcca attcccacta tatggaacat ccgcaaatgc agataggcca
1020

tttaatttag ccattcacta ttctgaaaca atttatagaa cattatcagc ccctatttat
1080

tcggtttctg gtggtatttc tcctaataga acgagagcag tgggaaggggt aaggttctta
1140

accgcaagag ataataatct aaattccttg ccatttttat atagaaaaga aggttcttta
1200

gattctttta ccgagttacc acctgaagat gagaatgaac caccttatat tgggtatagt
1260

catcggttat gccatgctag atttgctagg tcatccgtag ttcttgagcc aagtaatttt
1320

gcgagaattc cagtattttc atggacacat cgtagcgctg gccctactaa tgaagtaagt
1380

tcatctagaa taacacaaat cccatgggta aaagcacata cccttgattc aggagcattt
1440

gttataaagg gtcttgatt tacaggaggg gatattctta ctaggcctaa tttgggtacc
1500

ttaggagctt taagagtgc ccttacagga caattaccac aaacatataa tataagaatc
1560

cgatatgctt cgatagcaaa tagagggtgt aactgattt tttcacagcc accttcatat
1620

ggcctcacat ttccaaaaac aatggatata gatgagccct taacatctcg ttcgtttgct
1680

cgtacaactc ttttcacacc aataactttt acacaagcac aggccgaatt aaatctaaca
1740

atacaacagg gtgtttatat agatagaatt gaattcatcc cagttaatgc aacctttgag
1800

gcagaatatg atttagaaag agcacaaaag gcggtgaatg ctctgtttac ttcttccaat
1860

caactaggat taaaaacaga ttgacggat tatcatattg atcaagtttc caatttagtg
1920

gattgtttat ccgatgaatt ctgtatagat gaaaagcgag aattgtccga gaaagtcaaa
1980

catgcgaagc gactcagtga tgaacggaat ttactccaag attcaaactt tagaggcatc
2040

aatagacaac cagaccgtgg ctggagagga agtacggata ttaccatcca aggaggaaat
2100

gacgtattca aagaaaatta cgtcacacta ccgggtacct tcgatgaatg ttatccaacg
2160

tatttgtatc aaaaaataga tgagtcaaaa ttaaaagcct atacccgta tcagttaaga
2220

gggtatatcg aagatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac
2280

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagtcga aagtccaatc
2340

ggaaagtgcg gagaaccgaa tcgatgtgtg ccacagcttg aatggaattc taatctagat
2400

tgttcctgca gagacggaga aaaatgtgca catcattccc atcattttctc cctagacatt
2460

gatgttggat gtacagactt acatgatgac ttaggtgtat gggatgatatt caagattaag
2520

acgcaagatg gccatgcaag actagggaaat ttagagtttc tcgaagagaa accattagta
2580

ggagaagcgc tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgcgaaaca
2640

ttgcaattgg aaacgaatat cgtttacaaa gaggcaaaag aatctgtgga tgctttattt
2700

gcaaactctc aatataatag attacaagcg gatacgaaca tcgcaatgat tcatgcggca
2760

gataaacgcg ttcatcgaat ccgagaggct tatcttccgg aattatccgt tatcccaggt
2820

gtaaatgcgg ggatttttga agaattagaa ggtcgtattt tcaactgcatt ctctttatat
2880

gatgcaagaa atgtcattaa aaacagtgat ttcaataatg gcttatcatg ctggaacgtg
2940

aaagggcatg tagatataga agaacagaac aaccaccgtt cggtccttgt tgttcctgaa
3000

tggaagcag aagtatcaca aaaagttcat gtttgtccag gtcgtggta tatccttcgt
3060

gtcacagcgt acaaggaggg atatggagaa ggttgcgtaa ccattcatga aattgaagat
3120

catacagatg aactaaagtt tagaaactgt gaagaagacg aagtctatcc aaacaacacg
3180

agaacgtgta atgcttatcc tgcagatcaa gaaggatatg agggggcgtg cacttctcgt
3240

aatcgtggat atgacgaagt ctatggaaac accccttccc taccagctga ttatgcgcca
3300

atztatgagg agaatgcata cacagatgga cgaagaggta atccttgtga atctagcaga
3360

gggtatggggg attacactcc actaccagct ggttatgaaa caaaggaatt agag
3414

<210> 8

<211> 2160

<212> ADN

<213> Bacillus thuringiensis

<400> 8

atgaaactaa agaatccaga taagcatcaa agcttgtcta gcaatgcgaa agtagataaa
60

atcgctacgg attcactaaa aaatgaaaca gatatagaat tgaaaaatat gaataatgaa
120

gattatttga gaatgtctga gcatgagagt attgatccgt ttgttagtgc atcaacaatt
180

caaacgggta ttggaattgc tggtgaagatt cttgggtactc taggtgttcc ttttgctgga
240

caaatagcta gcctctatag ttttatctta ggcgagcttt ggcctaaagg gaaaagtcaa
300

tgggaaatct ttatggaaca tgtagaagag attattaatc aaaaaatatt aacttatgca
360

agaaataaag cactttcaga cttgagagga ttaggggatg ctttagccgt ctaccatgaa
420

tcgcttgaaa gttgggttaa aaatcgtaat aacactcgag cgaggagtgt agtcaagaac
480

caatatatcg cattagaact gatgtttgtt caaaaactac cttcttttgc agtatctggt
540

gaggaagtac cattattacc gatatatgcc caagctgcca atttacattt gttgttatta
600

agagatgcat ctatTTTTGG aaaagaatgg ggattatcag cttcagaaat ttcaacattt
660

tataaccgtc aagtcgaacg aacaagagat tattccgacc attgtgtaaa atggtataat
720

acaggcctaa ataacttgag ggtacaaat gccaaaagtt gggttcgtta taatcaattt
780

cgtaaagata tgacattaat ggtattagat ttagttgcgc tattcccaag ctatgataca
840

cttgatatatc ctattaaaac cacttcacaa cttacaagag aagtatatac agacgcaatt
900

gggaccgtgc atccttcagg tgccgttgca agtacgactt ggtataataa taatgcacct
960

tcgtttctcta cgatagaggc tgctgttggt cgaaacccgc atctactoga ttttctagaa
1020

caagttacaa tttacagctt attaagtcga tggagtaaca ctcagtatat gaatatgtgg
1080

ggaggacata aactagaatt ccgaacaata ggaggacgt taaatacctc aacacaagga
1140

tctactaata cttctattaa tcctgtaaca ttaccgttca cgtctcgaga cgtctatagg
1200

actgaatcat tggcagggct aaatctattt ttaactcaac ctgttaatgg agtacctagg
1260

gttgattttc attggaaatt cgtcacacat ccgatcgcat ctgataatth ctattatcca
1320

gggtatgctg gaattgggac gcaattacag gattcagaaa atgaattacc acctgaaaca
1380

acaggacagc caaattatga atcttatagt catagactat ctcatatagg actcatttca
1440

gcaccccatg tgaaagcgtt ggtatattct tggacgcac gtagtgcaga tcgtacaaat
1500

accattaatt cagatagcat tacacaaata ccattagtaa aggcattcaa tcttccctca
1560

gggtgcttctg ttgttagagg accaggatth acagggtggg atatccttcg aagaacaaat
1620

actggtacat ttggggatat acgagtaaatt attaatccac catttgcaca acggtatcgc
1680

ttaagaattc gttatgcttc tactacaaat ttagaattcc atacgtcgat taacggaaaa
1740

gctattaatc aaggtaatth ttcagcaact atgaatagag gagaggactt agactataaa
1800

gccttttagaa ctgtaggctt tactactcca tttagcttht caaatgcaca aagtacattc
1860

acaataggtg cttggaactt ctcttttaggt aacgaagttt atatagatag aattgaattt
1920

gttccggtag aagtaacata tgaggcagag catgattttg aaaaagcgca agaggaggtt
1980

actgcactgt ttacatctac gaatccaggt gggttaaaaa caaatgtaac ggagtatcat
2040

attgaccagg tatcaaattt agtagagtct ttatcaaag aattctatct cgatgaaaag
2100

agagaattat tcgagatagt taaatacgcg aagcaacttc atattgggcg taacatgtag
2160

<210> 9

<211> 3495

<212> ADN

<213> Bacillus thuringiensis

<400> 9

atggagagaa ataatcagga tcaatgcatt ccttataatt gtttgaataa tcctgaggtt
60

gggatattag atattgaaaa tttcaatctc gaacttgtat cgcaagtcag cgtggggctc
120

acacgttttc ttttgaagc atctatccct ggggcaggtt ttgcattggg tctattcgac
180

atcatttggg gtgctctagg cgtcgatcaa tggagcttat tccttgcgca aattgagcaa
240

ttaattaatg aaagaataac aacagtagaa aggaatagag cgattcaagc attaagtggg
300

ctatcgagta gttatgaagt atatattgag gcattaagag aatgggagaa tgatctggat
360

aatccagctt cacgagatag agtggttgca cgttttcgtg caacagataa ttctctaata
420

acagatatat ctctattgga aattccaggt tttagatag ctactttatc agtctatact
480

caagcggcga atctacattt agctttgtta agagatgccg tttactttgg agaaagatgg
540

ggattaacac aaacaaatat tgaagatctg cacacaagac tcacgagata tattcaagaa
600

tactcagacc attgcgcaag atgggtataat caaggtttga ataattattgg agggataaat
660

acaagatatt tggacttcca aagagaatta acaatttcgg tcttagatat tgtcgctctt
720

ttcccaaatt acgacatccg aacatatcca attccgacac aaagccaatt aacaagggaa
780

atatatacat ctcccgtcgt tgcacctggt gtaaattgga ttttaagtat atcgaatgta
840

ttgagagccc ctcatctgat ggattttttt gatcgaataa ttatttatac tggtagagtt
900

agaagtacac cacattggga agggcatgaa gtcatatcta gaagaacagg gcaaggaaat
960

gagatacgct cgcctttata tggagtggct gcaaacgcag aaccaccagt tactataaga
1020

cctacaggat ttactgatga gcaaagacaa gtgtatagag tactatcacg tgttgcttct
1080

tttagaaatt caggaaccaa ctttagtctg gtagatgcag catcattcct aactatattt
1140

agcgctagct caatctatag aaatggtttt gggtttaatg ctgatactat tgatgaaatt
1200

ccaattgagg ggactgatcc atatattgga tatagccatc gattatgcca tgttggattt
1260

actgcgtcat ctccatttat ctcccagtat gcaagggcgc ccgtattttc ttggacgcat
1320

cgtagtgcca cttttacaaa tacaattgat ccagagagga ttacgcaaat accaatggta
1380

aaagcataca atcttcatgc aggtgccact gttgttagag gaccggggtt tacagggtgtt
1440

gatctcttac gaagaacgaa tactggtaca tttgcagata taagagtaaa tattactggg
1500

ccattatctc aaagatatcg tgtaagaatt cgctatgctt ctacgacaga tttacaattt
1560

ttcacgagaa tcaatggaac ttctgtaaata caaggtaatt tccaaagaac tatgaataga
1620

ggggataatt tagaatctgg aaactttagg actgcaggat ttagtacgcc ttttagtttt
1680

tcaaatgcgc aaagtacatt cacattgggt actcaggctt tttcaaata ggaagtttat
1740

atagatcgaa ttgaatttgt cccggcagaa gtaacattcg aggcagaatc tgatttagaa
1800

agagcgcaaa aggcggtgaa tgccctgttt acttctacaa gccaaactagg gctaaaaaca
1860

aatgtaacgg gttaccatat tgatcaagtg tccaatttag ttgctgtgtt atcggatgaa
1920

ttttgtctgg atgaaaagag agaattgtcc gagaaagtta aacatgcgaa gcgactcagt
1980

gataagcggg atttacttca agatccaaac ttcagaggga tcaataggca accagaccat
2040

ggctggagag gaagtacgga tattactatc caaggaggag atgacgtatt caaagagaat
2100

tacgttacgc taccgggtac ttttcatgag tgctatccaa cgtatttata tcaaaaaata
2160

gatgagtcga aattaaaagc ctatacccgt tatcaattaa gagggatatat cgaagatagt
2220

caagacttag aaatctatct aattcgttac aattcaaaac acgaaatagt aaatgtacca
2280

ggtacaggga gtttatggcc tctttctgta gaaaatcaaa ttggaccttg tggagaaccg
2340

aatcgaatgc cgccacacct tgaatggaat cctgatttac actgttcctg cagagacggg
2400

gaaaaaatgtg tgcatcattc tcatcatttc tctttggaca ttgatgtcgg atgtacagat
2460

ttaaatgagg acctaggtgt atggttgata ttcaagatta agacgcaaga tggccacgca
2520

agactaggga atctagagtt tctcgaagag gaaccgttat taggcgaagc gttagcacgt
2580

gtgaagagag cggagaagaa gtggagagac aaacgcgaga aactgcagtt ggaaacaaat
2640

attgtctata aagaggcaaa agaattctgta gatgctttat ttgtaaactc tcaatatgat
2700

agattacaag cggatacgaa catcgcgatg attcatgcgg cagataaacg cgttcataga
2760

atccgggaag cgtatctgcc agagttgtct gtgattccag gtgtcaatgc ggccattttc
2820

gaagaattag agggacgtat ttttacagcg tattccttat atgatgcgag aaatgttatt
2880

aaaaatggca atttcaataa tggcttatta tgctggaacg tgaaagggca tgtagatgta
2940

gaagagcaaa acaaccaccg ttcggtcctt gttatcccag aatgggaagc agaagtgtca
3000

caaaaagtgc gtgtctgtcc aggtcgtggt tatatccttc gtgttacagc gtacaaagag
3060

ggatatggag aaggttgctg aacgatccat gagatcgaag acaatacaga cgaattaaag
3120

tttagcaact gtgtagaaga aggatatcca aacaacacgg taacgtgtaa tgagtatact
3180

atgaatcaag gggtaggaga gtgtacggat gcatgtaatg tccgtaatcg tggatatgag
3240

gatgcatatg gacacaatcc ttcaacgcct gttcattaca caacgccgta cgaagaagaa
3300

acgtatacag atgaacgaag agagaatcct tgtgaagcta acaaagggtg tgtgaattac
3360

acgccactac cagttgggta tgtgacaaaa gaattagaat acttcccaga aaccgacaca
3420

gstatggattg aaattggaga aacggaagga acattcattg tggacagcgt ggaattactc
3480

cttatggagg aatag
3495

<210> 10
<211> 3501
<212> ADN
<213> Bacillus thuringiensis

<400> 10
atgaagataa ataatacagaa ccaatgcata ccatataatt gcttaagtaa tcctgaggaa
60

gtacttttgg atggggagag gatattacct gatatcgatc cactcgaagt ttctatgtcg
120

cttttgcaat ttcttctaaa taactttggt ccagggggggg ggggttatttc aggattaatt
180

aataaaatat ggggggcttt gagaccatct gaatgggaat tatttcttgc acagattgaa
240

cagttgattg atcgaagaat agaagcagca gtaagagcaa aagcaatcgc tgaattagaa
300

ggtttaggga gaagttatca actatatgga gaggcattta aagagtggga aaaaactcca
360

gataacacag cggctcggtc tagagtaact gagagatttc gtataattga tgctcaaatt
420

gaagcaaata tcccttcggt tcgggtttcc ggatttgaag tgccacttct atcggtttat
480

acccaagcag ctaatttgca tctcgctcta ttaagagatt ctgttgtttt tggagagaga
540

tggggattga cgactacaaa tgtcaatgat atctataata gacaagttaa gagaattgat
600

gaatatagcg atcattgtgt agatacgat aaaacagaat tagaacgtct agagtttagc
660

tctatagcgc aatggagaat atataatcag tttagaaggg aattgacact aacggtatta
720

gatattgtcg ctcttttccc gaactatgat ggtagactgt atccgattcg aacaatttct
780

caattgacaa gagatattta tacatcccca gtaagcgaat tttattatgg tcccatttat
840

aattataata tagttgggtcg cttactgaa cagcagctaa ggcgaccaca tcttatggac
900

ttctttaact ccatgatcat gtatacgtca gataatagac gagaacatta ttggtcagga
960

cttgaaatga aggctactga tacttcagga aaccaagtgt cattcccttt agctgggact
1020

agaggggaatt cagctccacc agtaactgtt agaaataatg gtgagggagt ttatagaata
1080

ttatcagagc cattttattc atcacctttt ctaggcacca gtgtgctagg aagtcgtggg
1140

gaagaatttg cttttgcac taatactact acaagtctgc catctacaat atatagaaat
1200

cgtggaacag tagattcatt agtcagcata ccgccacaag attatagcgt accaccgcac
1260

aggggggtata gtcatttatt aagtcacgtt acgatgcaca atagttctcc tatattccac
1320

tggacgcac gttagtgcaac ccctagaaat acaattgatc cagatagtat cactcaaact
1380

ccagtagtta aggcttcgca cctctctggg gggttcagtta taaaagggcc tggacatata
1440

ggtaggagatt taataagcct acctgtaaat aactttactc atttccgaat cccatttcag
1500

gcaaactc cacaaggta tcgtattaga attcgttatg cggcagactc agatgggact
1560

ttggatagt gagttttctt aagtcagca gcaggggatg gttttaatac aacttcttat
1620

agggccacaa tgagccctgg aggttcctta acatctcgtg attttcaatt tttagattta
1680

aacacatcgt ttacctccga tgtagcatct aacttatggt tacattttat acgttatata
1740

cgaccaggga atttgtatat agatagagcg gaatttatcc cagtggatgc aaccttcgag
1800

gcaggttata atttagaaag ggcgcaaaag gcggtgaatg ccctgtttac ttctacaaac
1860

caaaaaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatctagtt
1920

gattgtttat ctgatgagtt ttgcttagat gaaaagcgag aattgtccga gaaagtcaaa
1980

caggcgaagc gactcagtga tgagcggaat ttactccagg attcaaattt cagaggcatc
2040

aatagggaaac aagaccgtgg atggagagga agtacgcata ttactatcca aggaggaaat
2100

gatgtattca aagaaaattt tgttacacta ccagggtgcct ttgatgcgtg ttatccaacg
2160

tatttgtatc aaaaaataga tgaatctaaa ttaaaagcct atacacgtta tgaattaaga
2220

ggatatatag aagatagtca agatttggat atttacttga tccggttaca tacgaaacat
2280

gaaacattaa atgttccagg tactaagtct ccatggtcgc tttgtacgga gagcccactt
2340

ggaaagtgtg gggaaccaa tcatgctgca tcacaaatag aatggaatcc tgatctagac
2400

tgctcttgca gagacggaga aaaatgtgcg catcattcgc atcattttctc cttggacatt
2460

gatgttggat gtacagattt gaatgagaac ctaggtatat gggatgatatt caagattaag
2520

acacaggatg gtcacgcaag actaggaaat ctagaatttc tcgaagagaa accgttatta
2580

ggagaagcgt tagcccgtgt gaagagagcg gagaaaaaat ggagagacaa acgtgaaaaa
2640

ttgcaatcag aaacaaatat tgtttataaa gaagcaaaag aagctgtaga tggttttattc
2700

gtagattccc aatatgagag attacaagct gatacgaata tcgccatgat tcatgcggca
2760

gataaacacg ttcattcaaat ccgagagggt tattttcccag agctttctgt gattccaggt
2820

gtcaatgcag cgattttcga agaattagaa ggccgtatgt tcacagcgta ttccctatat
2880

gatgcgagaa atgtcattaa aaacgggtgat ttcaataatg gcttatcatg ctggaacgtg
2940

aaagggcatg tagatgtaga agtacaaaac aaccaccgtt cagtcctagt tatcgcgga
3000

tgggaagcag aagtttcaca agaagttcct gtctgtccag gtcgtggcta tatccttcgt
3060

gtcacagcgt acaaagaagg atatggagaa ggttgcgtaa ccattcatga gattgaagat
3120

catacagatg aactgaagtt tagaaactgt gaagaagagg aagtatatcc aaacaacaca
3180

gtaacgtgta atgggttatac tgcgactcaa gaagaatata aggatgcata cacttctcgt
3240

aatcgaggat atgacgaagc ctatggaaat aacccttccg taccagcaga ttatgcgtca
3300

gtctatgaag aaaaagcgtg tacagatgga cgaagagaga atccttgtga aatggaaaga
3360

ggttacacac ctttaccagt tgggttatata acaaaagaat tagaatatgt cccggaaacg
3420

gatacagtat ggattgagat tggagaaacg gaaggaacgt ttatcataga tagcgtggaa
3480

ttactcctca tggaggaata g
3501

<210> 11

<211> 3546

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 11

atggacaata atccgaacat caatgagtg atcccttaca actgcctgtc taaccctgag
60

gtggagggtgc tggggggcgga gcgcatcgag acaggctaca ctccaatcga catttcgctg
120

tctctcacgc agttcctcct gtcggagttc gtgcctggcg ctgggttcgt gctgggcctc
180

gtcgatatca tttgggggat cttcgggccc tcccagtggg acgctttcct ggtccagatc
240

gagcagctca ttaaccagcg catcgaggag ttcgctagga atcaggccat ttcgcggctg
300

gagggcctct ctaacctgta ccagatctac gcggagtcc tccgcgagtg ggaggctgat
360

ccgaccaacc ccgccctgag ggaggagatg cggattcagt tcaacgacat gaattcggcg
420

ctgaccacgg ctatcccgct cttcgccgtc cagaactacc aggttcccct cctgtctgtg
480

tacgtccagg ccgcgaatct gcacctctca gtgctgcgcg acgtttccgt gttcgggcag
540

cgctggggct tcgacgctgc tactattaac tcccgtaca atgatctcac taggctgatc
600

ggcaactaca ccgactacgc cgtgcgctgg tacaataccg ggctcgagag ggtttggggc
660

ccggacagca gggattgggt caggtacaac cagttccgca gggagctgac cctcacggtt
720

ctggacatcg tggccctctt ctcgaaactac gattctcggc gctaccaat taggaccgtg
780

agccagctga cgcgggagat ctacacaaat cctgtcctcg agaacttcga cgggtcgttc
840

aggggcatgg ctcagcgcat tgagcagaat atccgccagc cgcacctgat ggatatcctc
900

aattcgatca ccatttacac ggacgtgcac aggggggttc actactggtc tggccatcag
960

atcaccgctt ctccagtcgg gttctcaggc ccggagtctg ctttcccact gttcgggaac
1020

gctggcaatg ctgctccacc cgtcctcggt tcgctgacgg gcctcgggat cttccgcaca
1080

ctgtccagcc cactctacag gcggatcatt ctcggctctg ggcctaacaa tcaggagctg
1140

ttcgtgctcg acggcacgga gttctctttc gcgtcactga caactaacct cccatccact
1200

atttacaggc agaggggcac cgtggacagc ctcgatgtca tcccacctca ggacaactcg
1260

gtcccgccaa gggctggctt ctcacaccgc ctgtcccatg tcaccatgct cagccaggct
1320

gctggcgctg ttacactct gagggctccc accttctcct ggcagcaccg gagcgccgag
1380

ttcaacaata tcattgcgtc agattccatc acgcagattc cggcgggtgaa ggggaacttc
1440

ctcttcaatg gctccgtcat tagcggcccc gggttcacag gcggggacct ggtgcgcctc
1500

aactcgtctg ggaacaatat ccagaatagg ggctacattg aggtcccaat ccacttccct
1560

agcacatcga ctgcgtaccg cgtgagggtc cgctacgcta gcgtgactcc gattcatctg
1620

aacgtcaatt ggggcaattc atccatcttc tcgaacaccg tgcccgtac cgctacgtct
1680

ctggacaacc tccagtcctc cgatttcggc tacttcgagt ctgctaattgc cttcacgtct
1740

tcactcggga atatcgttgg cgtgcggaac ttctcagga cagccggcgt tatcattgac
1800

cgcttcgagt tcatcccggt gacagccact ctggaggcgg agtacaacct cgagcgcgct
1860

cagaaggccg tgaatgcgt gttcacctcc acgaaccagc tgggcctcaa gacaaatgtc
1920

actgactacc acatcgatca agtcagcaac ctgggttacct gcctctcgga cgagttctgc
1980

ctggatgaga agaggagct gagcgagaag gtgaagcatg ccaagaggct ctcagacgag
2040

cggaacctcc tgcaggactc caatttcaag gatatcaaca ggcagccaga gagggggtgg
2100

ggcgggtcaa ccggcatcac gattcagggc ggggacgatg tcttcaagga gaactacgtt
2160

acactgagcg gcactttcga tgagtgtac ccacttacc tctaccagaa gatcgacgag
2220

tcaaagctga aggcgttcac ccgctaccag ctccggggct acattgagga ctcccaggat
2280

ctggagggtg acctcatccg ctacaacgcc aagcacgaga cgctgaatgt ccagggcaca
2340

gggtctctgt ggccactcgc ggttaagtca ccaattgggc gctgcggcga gccaaacagg
2400

tgcgctccta ggatcgagtg gaagcctgac gtggattgct cctgcaggga cggcgagaag
2460

tcgcgtcacc atagccacca tttctcgctc gacatcgatg tcgggtgcac ggacctgaac
2520

gaggatctcg gcgtctgggt tatcttcaag attaagacac aggacgggca tgccaagatc
2580

ggcaacctgg agttcctcga ggagaagctc ctgctcggcg aggctctggc taggggaag
2640

aaggccgaga agaagtggcg cgacaagagg gagaagctcg agtgggagac caacatcgtc
2700

tacaaggagg ccaaggagtc cgttgacgcg ctgttcgtgg atagccagta caacaggctc
2760

cagacagata ctaatatcgc catgattcac gctgccgaca agcgggtgca tcgcatccgc
2820

gaggcgtacc tgccagagct gagcgtgatt cctggcgta acgcggctat ctccgaggag
2880

ctggagggcc tcattttcac ggctttctcc ctgtacgacg cccgcaacgt gatcaagaat
2940

ggggatttca actacggcct cagctgctgg aacgtcaagg gccacgtgga cgtcgaggag
3000

cagaacaatc ataggtccgt tctggtgatc ccggagtggg aggtgaagt cagccaggaa
3060

gtcagggttt gcccaggcag ggggtacatc ctccgggtta ccgcctacaa ggagggctac
3120

ggggagggct gcgtgacgat ccacgagatt gaggacaaca cagatgagct gaagttcagc
3180

aattgcgttg aggaggaggt gtacccgaac aataccgtga cgtgcaacga ctacacagcc
3240

actcaggagg agtacgaggg cacctacacg tcgaggaacc ggggggtacga tggcgcttac
3300

gagtctaatt ccagcgtccc agccgactac gcctcagcgt acgaggagaa ggcgtacacg
3360

gacggccgca gggataatcc ttgcgagtcc aaccggggct accgcgatta cactccactg
3420

cctgccggct acgtcaccaa ggagctggag tatttcccgg agacagacaa ggtttgatc
3480

gagattggcg agacggaggg gaccttcatt gtggattcgg ttgagctgct gctgatggag
3540

gagtag
3546

<210> 12
 <211> 3531
 <212> ADN
 <213> Trình tự Nhân tạo

 <220>
 <223> Trình tự mã hóa Cry tổng hợp

 <400> 12
 atggagatcg tgaacaatca gaatcagtgc gtgccttaca actgcctcaa taatcccagag
 60
 atcgagatcc tggagggggg gcgcatttcg gtgggcaaca cgcctatcga catttccttc
 120
 agcctgacac agttcctcct gtctgagttc gttccaggcg ctgggttcgt gctcggcctg
 180
 atcgacctca tttgggggtt cgtgggcccg tcccagtggg atgctttcct cgcccaggtc
 240
 gagcagctga tcaaccagag gattgcggag gctgttcgga ataccgcgat ccaggagctg
 300
 gagggcatgg ctgcgctgta caggacatac gccactgcgt tcgctgagtg ggagaaggct
 360
 ccagacgatc cagagctgag ggaggctctg aggacacagt tcacggcgac agagacttac
 420
 atctccggcc gcattagcgt cctcaagatc cagaccttcg aggttcagct cctgagcgtg
 480
 ttcgctcagg ctgctaacct ccacctgtcg ctccctgaggg acgtggtctt cttcgggcag
 540
 cgctggggct tctcgaccac gacagtgaac aattactaca acgacctcac tgagggcatc
 600
 tctacctaca cggattacgc cgtccgctgg tacaataccg ggctggagag ggtgtggggc
 660
 cccgacagca gggattgggt ccggtacaac cagttccgca gggagctgac actgactgtg
 720
 ctcgacatcg tcgcgctgtt ctcgaactac gattctcggc gctaccccat ccggacagtg
 780
 tcgcagctca ctgcgagat ttacaccaac ccagtcctgg agaatttcga cgggtccttc
 840
 aggggcatgg ctcagcggat cgagcagaac attaggcagc cccacctcat ggacatcctg
 900
 aattccatca ccatttacac ggatgtgcac cgcgggttca actactggag cggccatcag
 960

atcacccgctt ctccagtcgg gttctcaggc ccggagttcg ctttccccct cttcgggaac
 1020
 gctggcaatg ctgctccacc cgtgctcgtc tccctgactg gcctcgggat cttcaggacc
 1080
 ctctccagcc cactgtacag gcggatcatt ctgggcagcg ggccatacaa tcaggagctg
 1140
 ttctgtctcg acggcacgga gttctcattc gcgtccctca ctaccaacct gccctccacg
 1200
 atctacaggc agcggggcac agtcgactct ctcgatgtta ttccacctca ggacaactca
 1260
 gtcccccaa gggctggctt ctcacacagg ctctcccatg ttaccatgct gagccaggct
 1320
 gctggcgctg tgtacacgct gcgggctccg acattctcct ggcagcaccg cagcgcgagg
 1380
 ttcaacaata tcattccgtc gtctcagatc acgcagattc ccctcacaaa gtccactaac
 1440
 ctgggcagcg ggacatcggg tgtgaagggc ccgggggttca ctggcgggga catcctcagg
 1500
 aggacctgc caggccagat ctctactctc cgcgttaata ttaccgctcc actgagccag
 1560
 cgctacaggg tccgcatccg ctacgcctca acgacaaacc tgcagttcca cacgtccatc
 1620
 gacgggaggg ctattaacca gggcaatttc tctgccacta tgtcatccgg gtcaaacctc
 1680
 cagtcgggct ctttccgcac cgtgggggtc actaccccg tcaacttctc caatggcagc
 1740
 tcggtgttca ccctgtcagc gcatgtcttc aactccggca atgaggttta catcgaccgg
 1800
 attgagttcg ttccggccga agtgactttc gaggcggagt acgatctcga gcgcgcgcag
 1860
 aaggctgtga acgagctgtt cacgtcttca aatcagatcg gcctgaagac cgacgtcacg
 1920
 gattaccaca ttgaccaggt ctccaacctc gttgagtgcc tgagcgacga gttctgcctc
 1980
 gatgagaagc aggagctgtc cgagaagggt aagcatgcc aagcctcag cgacgagagg
 2040
 aacctcctgc aggaccgaa cttccgcggc atcaataggc agctggacag ggggtggagg
 2100

ggcagcacag atatcactat tcagggcggg gacgatgtct tcaaggagaa ctacgttacc
2160

ctcctgggca cgttcgacga gtgctaccca acgtacctgt accagaagat cgatgagtca
2220

aagctcaagg cgtacacacg ctaccagctg cgcggctaca ttgaggacag ccaggatctc
2280

gagatctacc tgattaggta caacgccaa gacgagacgg ttaatgtgcc tggcacaggc
2340

tccctctggc cactctctgc tcagtcaccc atcgggaagt gcggcgagcc aaacagggtgc
2400

gctcctcatc tcgagtggaa tccggacctg gattgcagct gcagggacgg cgagaagtgc
2460

gctcaccata gccaccattt ctcgctcgac atcgatgtgg gctgcacgga cctcaatgag
2520

gatctggggg tctgggttat cttcaagatt aagacacagg acgggcacgc taggctgggc
2580

aacctcgagt tcctggagga gaagccactc gtgggcgagg ctctggctag ggtcaagatc
2640

gccgagaaga agtggaggga caagcgggag aagctggagt gggagacgaa catcgtgtac
2700

aaggaggcta aggagtcggt ggatgccctc ttcgtcaact ctcagtacga ccagctgcag
2760

gctgatacca atatcgccat gattcacgcc gcggacaagc ggggccattc gatccgcgag
2820

gcctacctcc cagagctgtc tgtgatccct ggcgtcaacg ctgccatttt cgaggagctg
2880

gagggcagga tcttcaccgc gttcagcctg tacgacgctc ggaacgtgat taagaatggg
2940

gatttcaaca atggcctctc gtgctggaac gtcaagggcc acgtggacgt cgaggagcag
3000

aacaatcagc gctctgtcct ggtcgttcct gagtgggagg ccgaggtttc acaggaagtt
3060

agggtgtgcc caggcagggg gtacatcctc cgcgtgaccg cgtacaagga gggctacggg
3120

gagggctgcg tcacgatcca tgagattgag aacaatacag acgagctgaa gttcagcaac
3180

tgcgtggagg aggagatcta ctgaacaat accgttacgt gcaacgatta cactgtgaat
3240

caggaggagt acggcggggc ttacacctca aggaacaggg gctacaatga ggctccatcc
3300

gtccctgctg actacgctag cgtttacgag gagaagtcgt acacggatgg caggagggag
3360

aacccatgcg agttcaatcg cggctacagg gactacactc cactccctgt cggctacgtt
3420

accaaggagc tggagtatctt cccagagacc gataaggtct ggatcgagat tggcgagact
3480

gaggggacat tcattgtgga tagcgtggag ctgctgctga tggaggagtg a
3531

<210> 13

<211> 3714

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 13

atgaatagca acaggaagaa tgagaacgag atcattgacg cttcgttcat cccagccgtc
60

tcgaacgagt ctgttactat ttctaaggag tacgcgcaga ccaatcagct ccagaacaat
120

tcaattgagg acgggctgtg catcgctgag ggcgagtaca ttgatccttt cgtgtcagcg
180

tccaccgtcc agacggggat ctccattgct ggccgcatcc tgggggtcct gggcgttcca
240

ttcgctggcc agctcgctc attctactcc ttcacgtggt gggagctgtg gcctaagggc
300

agggaccagt gggagatctt catggagcac gttgagcagc tcgtgcgcca gcagattact
360

gctaatagcca ggaacacggc cctggctcgg ctccagggcc tggggaactc cttccgcgcc
420

taccagcaga gcctcgagga ctggctggag aataggaacg atgcccggac gcgcagcgtg
480

ctctacacac agtacatgcg gctggagctg gacttctca acgcgatgcc actgttcgtt
540

attcgcgagc aggaggttcc tctcctgatg gtgtacgctc aggcggccaa cctgcacctc
600

ctgctcctgc gcgatgcctc gctctacggg agggagttcg gctgacttc tcaggagatc
660

cagcgctact acgagaggca ggtcgagagg acccgggact actcagatca ttgcgtgcag
720

tggtaacaac cggggctcaa caatctgagg ggcacaaacg cggagagctg ggtccgctac
780

aatcagttcc gcagggacct gaccctgggc gtcctggatc tcgttgccct cttcccgtca
840

tacgacacca ggacgtaccc catcaacaca tccgcgcagc tgactcggga ggtctacacc
900

gatgccattg gggcgactgg cgttaacatg gcttcgatga attggtacaa caataacgcc
960

ccaagcttct cggcgatcga gaccgctgtg attaggtctc ctcacctcct ggacttcctc
1020

gagcagctga agatcttctc ggcctccagc aggtgggtcta acacccggca catgacgtac
1080

tggagggggc atacgatcca gagcaggccc attcgcggcg ccctcatcac aagcactcac
1140

ggcaatacca acacgtcgat caaccagtc acgttccagt tcccttctcg ggacgtttac
1200

cgcacagagt catacgccgg ggtgctcctg tggggcatct acctcgagcc aatccacggc
1260

gtgccgacgg tcaggttcaa tttccggaac cccagaata cgttcgagcg cggcacagcg
1320

aactactctc agccgtacga gtcacccggc ctgcagctca aggactccga gacggagctg
1380

ccgccggaga ccaccgagag gccaaactac gagtcttact cacaccggct gtcccatatc
1440

ggcatcattc tccagactag gctgaatgtg cctgtctact cctggaccca ccgcagcgct
1500

gacaggacaa aactatcgg cccaaatcgg atcacgcaga ttcttgccgt gaaggggaac
1560

ctcctgttca atggcagcgt catctcgggc ccgggggttca caggcgggga cctcgtccgc
1620

ctgaataaca gcgggaataa catccagaac aggggctacc tcgaggtgcc cattcagttc
1680

acctccacga gcacacgcta cagggttcgg gtgcgctacg cttccgtgac ccccatccac
1740

ctgtcagtc actggggcaa ttccaacatt ttctcgtcta cagtgccgcg gactgctgcc
1800

agcctcgaca atctgcagtc gcgggatttc ggctacttcg agtctaccaa cgccttcaca
 1860

tcagttactg ggaatgtggt cggcgtgcgc aacttctccg agaatgcgcg cgttatcatt
 1920

gacaggttcg agttcatccc agtgactgcc accttcgagg cggagtagca tctggagagg
 1980

gctcaggagg cgtcaacgc gctcttcacg aatacaaacc cgaggaggct gaagacgggc
 2040

gtgaccgact accacatcga tgagggtgctg aacctcgctg cctgcctgtc tgacgagttc
 2100

tgctcgtatg agaagcgcga gctgctggag aaggatgaagt acgcgaagcg gctgagcgac
 2160

gagcgcaacc tcctgcagga cccgaacttc acgtccatca ataagcagcc cgacttcatt
 2220

agcacgaacg agcagtcgaa tttcacatct atccacgagc agtcagagca tgggtggtgg
 2280

ggctccgaga acatcaccat tcaggagggc aacgatgttt tcaaggagaa ttacgtgacc
 2340

ctcccaggca cgtacaacga gtgctaccct acgtacctgt accagaagat cggcgagtcg
 2400

gagctgaagg cctacacacg ctaccagctg cgcggctaca tcgaggactc tcaggatctc
 2460

gagatctacc tgattaggta caacgcgaag cacgagactc tggacgtgcc cggcaccgag
 2520

tcggtctggc cgctgtctgt tgagtccccg atccgcaggt gcggcgagcc caacaggctg
 2580

gctccccatt tcgagtggaa tccggacctc gattgctcct gcagggacgg cgagaagtgc
 2640

gccaccatt cacaccattt ctccctggac attgatgtcg ggtgcatcga tctccacgag
 2700

aacctgggcg tgtgggttgt gttcaagatc aagacgcagg aggggcatgc taggctgggc
 2760

aacctggagt tcattgagga gaagcccctc ctgggcgagg ctctctccag ggtcaagcgc
 2820

gcggagaaga agtggaggga caagcgggag aagctgcagc tcgagactaa gcgcgtgtac
 2880

accgaggcta aggaggccgt ggatgcgctc ttcgtcgaca gccagtagca taggctgcag
 2940

gccgacacga acatcggcat gattcacgcg gctgataagc tgggtgcatcg catccgcgag
3000

gcgtacctct ccgagctgag cgttatcccc ggctgaacg ctgagatttt cgaggagctg
3060

gaggggcgga tcattaccgc gatctccctg tacgacgctc gcaacgctcg taagaatggg
3120

gatttcaata acggcctcgc ctgctggaac gtcaagggcc acgtcgacgt tcagcagagc
3180

caccatcgct cgggtgctggg catccccgag tgggaggctg aggtgagcca ggcgggttagg
3240

gtgtgccccg gcagggggta catcctccgc gtcaccgctg acaaggaggg ctacggggag
3300

ggctgcgtga ctatccacga gattgagaat aacaccgacg agctgaagtt caagaactgc
3360

gaggaggagg aggtctaccc aactgacacc gggacgtgca atgattacac ggcccatcag
3420

ggcacagccg cgtgcaactc caggaatgct gggtagcagg acgcctacga ggtcgataca
3480

actgcgagcg ttaactacaa gccacctac gaggaggaga catacactga cgtgcgcagg
3540

gataaccact gcgagtagca ccgcggctac gtcaattacc cacctctccc ggccggctac
3600

gtgacgaagg agctggagta cttccccgag acagacaagg tctggatcga gattggggag
3660

acagagggca ctttcatcgt ggatagcatt gagctgctgc tcatggagga gtaa
3714

<210> 14

<211> 2148

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 14

atgaagtcta agaatcagaa catgcatcag agcctcagca acaacgctac tgtggataag
60

aatttcacgg gcagcctgga gaataacaca aatactgagc tgcagaattt caaccacgag
120

ggcatcgagc cattcgctctc agtttccaca attcagactg gcatcgggat tgcgggcaag
180

atcctgggca acctcggcgt gccgttcgct ggccagggtg cttccctcta cagcttcato
240

ctgggcgagc tgtggcccaa ggggaagagc cagtgggaga ttttcatgga gcatgtcgag
300

gagctgatca atcagaagat ttcgacctac gcccggaaca aggctctggc tgacctcaag
360

ggcctggggg atgctctcgc tgtgtaccac gagtcgctgg agtcttgat caagaaccgc
420

aacaatacca ggacgcggtc tgtggtcaag tcacagtaca ttaccctcga gctgatgttc
480

gtccagtcac tgccgtcctt cgccgtttcc ggcgaggagg tgccgctcct gccaatctac
540

gctcaggctg cgaatctcca tctcctgctc ctgcgcgacg cctcgatctt cggcaaggag
600

tgggggctga gcgattcga gatttctact ttctacaacc ggcaggtcga gcgcaccagc
660

gactactcgg atcactgcac aaagtgggtc gacactggcc tcaataggct gaaggggtcc
720

aacgcggaga tctgggtgaa gtacaaccag ttccgcaggg acatgactct catggtcctc
780

gatctgggtg cctgtttcca gagctacgat acacacatgt accccatcaa gaccacggct
840

cagctcacca gggagggtta cacgaacgcc attggcacag tgcaccaca tccttctttc
900

acctcaacaa cttggtacaa caataacgct ccgtctttct cagccatcga ggctgccgtc
960

attcggagcc cccatctcct ggacttctg gagcagggtta ctatctactc cctcctgagc
1020

agggtggtcga ataccagta catgaacatg tggggcgggc acaagctcga gttccggaca
1080

attggcggga ctctgaacac atccactcag ggcagcacca atacgtcgat caaccagtc
1140

acctccctt tcacgtcaag ggacatctac cggacggagt ccctggccgg cctcaatctg
1200

ttcctcacac agccagttaa cggggtgccc agggtcgact tccactggaa gttcgtgacg
1260

catccaatcg cgtccgataa cttctactac cctggctacg ctggcattgg caccagctc
1320

caggacagcg agaagtagct gccaccagag accacgggcc agccaaacta cgagtcctac
1380

agccacagge tctcccatat cggcctgatt tcggcgtctc acgtgaaggc tctcgtctac
1440

tcttggaacc atcgctcagc ggacaggaca aacactatcc actcagattc catcacgcag
1500

attcctctcg tcaaggccca cacactgcag agcggcacia ctgttgtgaa gggcccaggg
1560

ttcacgggcg gggacatcct gaggcgcaca tccggcgggc ctttcgcgtt cagcaatgtt
1620

aacctcgatt ggaatctgag ccagcgctac cgcgcccga tccgctacgc gtcgaccacg
1680

aacctccgga tgtatgtgac gatcgcgggc gagcgcattt tcgctgggca gttcaataag
1740

accatgaaca cgggcgaccc actgaccttc cagagcttct cgtacgctac aatcgatact
1800

gccttcacct tccctacgaa ggctccagc ctactgtgg gcgctgacac cttctcgtct
1860

gggaacgagg tctacgtgga tcgcttcgag ctgatcccag tgaccgctac gctcgagggt
1920

gtcacggacc tggagagggc tcagaaggcc gtgcatgagc tgttcacctc cactaatcca
1980

ggcgggctga agaccgacgt cgctaaggat cactacacca acacgatctc taagtcagtt
2040

cagtcggtgt tcaggtgccg gtgctctgag cgcacgagga tctaccgctg gggctacccg
2100

tctaagaagg agtactggta catttggggc tacacctcaa agtactga
2148

<210> 15

<211> 3285

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 15

atgttcctcg agcagattga gcagctgatc gaccagcgca ttgagaccgt cgagaggaat
60

cgggcgatcc agacgctcat tggcctgtcg aactcttacg atgtgtacat cgaggcgctc
120

aaggagtggg agaacaatcc ggacaattca gcttcccagg agcgcgtcag gaaccgggtc
180

cgcaccacgg acgatgcgct catcacaggc attccgctcc tggctattcc caatttcgag
240

atcgccactc tgtcagtcta cgttcaggcg gccaacctcc acctgtccct cctgagggac
300

gctgttttct tcggcgagcg ctgggggctg acccagatca atgtggacga tctctaccgc
360

aggctgacga acaatatccg gaactactcc gatcattgcg cgcgctggta caatgagggc
420

ctcgacaaca ttagcgggct gagccggctg atcaatttcc agcgcgaggt cacaatttcg
480

gttctcgata tcgtggccct gttcccgaac tacgacattc gcacttacct catcagcaca
540

acttcgcagc tcaccaggga gattttcacg tcgccgatcg tggccccaa cgacttctcc
600

gttgccctacg agggcgctcg gcgcgctccc cacctgttcg agttcctcga gaagctcgtg
660

atctacacgg gcgatcgag cgggatcagg cactggggccg gccatgagat tacgtctagg
720

cggacagact cataccacgg catcatccgc taccgctct acgggacagc tgccaacgcc
780

gagtccccat acactctcgc gctgcagcct tctgagtcaa tctaccgcac cctctccgag
840

cctatitttca gccagacggg cgggctgtct ccgcacagga ggagggttgt ggagggcgct
900

gagttctcaa tcgtcaacaa taacgttaac ccgtccagct tcgtgtaccg caggaagggc
960

tctctcgatt cattcacaga gctgccgccg gaggacgagt ctgtcccacc ttacatcggc
1020

tactcacacc agctctgcca tgtgggcttc gggaggacga atgtcatctt cgagccgtcg
1080

aacttcgcca gggtgcccg cttctcgtgg acacaccggg ctgcgacacc aactaatacc
1140

atcgaccctg atcggatcac tcagattcct tccgtgaagg cgctgtctct gcgcaactcg
1200

acagtcgttt ctggcccggg gttcactggc ggggatatcg tccgcatggg cgctgttcac
1260

cagatctacg ccaccgacct cagcatgaac gtgaggccgt cagtcgctct gtcccgtac
1320

ctcatcaggc tgcggtagcg ctgcaggggc tcatccaata tcgttattca tgggccatcc
1380

atccggttcg tgtccctccc tagcaccatg tcgaacgacg agccgctgac gtaccagtca
1440

ttccgctacg cctccatcac cagccaatt accaggccta tctacaatat gttcaacctc
1500

tccatcagcc ggattagcgg cgtccagaat ctgttcatcg accgcattga gttcatcccc
1560

gtggatgcga acttcgaggc tgagagggac ctggagcgcg ctcagaaggc tgtcaatgcg
1620

ctcttcacgt ccacaaacca gaggggcctg aagattgacg tgacggatta ccacatcgat
1680

caggttagca acctcgtgga ctgcctgtcg gatgagttct gcctcgacga gaagcgcgag
1740

ctgtctgaga agtcaaagca tgccaagagg ctgagcgacg agcggaacct cctgcaggac
1800

ctgaatttca aggatatcaa ccgccagccg gagagggggg ggtccggctc tactgggac
1860

accattcagg gcggggacga tgttttcaag gagaactacg tgactctccc cggcaccttc
1920

gatgagtgt acccaacct cctgtaccag aagatcgacg agtccaagct caaggcttac
1980

acgcgctacc agctgagggg ctacatcgag gacagccagg atctcgagat ctacctcatc
2040

aggtacaatg ccaagcacga gaccgttaac gtgccaggct caggctccct ctggcctctg
2100

tccgtcgagt cctccgttgg caagtgcggg gagccaaata ggtgcgccag ccgcatggag
2160

tggaaccag acctcgattg ctctgcagg gacggcgaga agtgcgctca ccatagccac
2220

catttctcgc tggacatcga tgtcggctgc accgacctca acgaggatct gggggctctg
2280

gttattttca agatcaagac gcaggacggc cagccaaga tcgggaacct cgagttcctg
2340

aaggagaagc tgctgctggg cgaggctctg gccaggggtga agaaggcgga gaagaagtgg
2400

agggacaagc gggataagct ggagtgggag accaacgtgg tctacaagga ggcgaaggag
2460

agcgtggacg ctctcttcgt cgattcccag tacagccgcc tgcaggccga tacgaacatt
2520

gcgatgatcc acgcggctga caagaggggtg catcgcatca gggaggctta cctcccggag
2580

ctgaccgtga ttcccggcgt caacgcctcg atcttcgagg agctggaggg caggatcttc
2640

acggcctact ctctgtacgg cgccaggaat gtcataaga acggcgactt caataacggg
2700

ctctcctgct ggaacgtgaa gggccacgtg gaggtccagc agatccacca taggtctgtc
2760

ctggttgtgc cctcatggaa gacagaggtt tcccaggagg tttgcgtgtg cccaggcagg
2820

gggtacatcc tcagggtgac tgcctacaag gagggtacg gggagggcaa tgtcacaatt
2880

cacgagatcg agaataacac tgacgagctg aagttccgca attgagagga ggaggaggtc
2940

taccogaata aactgttac ctgcaatgac tacaccgtca accaggagga gtacaagggc
3000

acgtgcacat cccggaaccg cgggtacgac gagagctacg agtcttcac cagcgagagc
3060

gcttactacg cctcgggtga cgaggagaag ggctacacgg atgggcggcg cgagaatctc
3120

tgcgagttca acagggggta cggcgactac accagcctcc ctacagccta cgtgactaag
3180

gagctggagt acttcccgga gacagataaa gtgtggatcg agattggcga gactgagggg
3240

gccttcatcc tggactccgt ggagctgctc ctgatggagg agtaa
3285

<210> 16

<211> 3510

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 16

atggagatca acaatcagaa ccagtgcgtc ccgtacaatt gcctcaacaa tcccaggtcc
60

gagatcctga acgtggccat tttctccagc gagcaggtcg cggagattca cctcaagatc
120

accgcgctga ttctcgagaa cttctgccc ggcggtctt tcgctttcgg cctgttcgac
180

ctcatttggg ggatcttcaa tgaggatcag tggtcagcgt tcctcaggca ggtggaggag
240

ctgatcaacc agaggattac ggagttcgct aggggccagg ctatccagag gctgggtgggc
300

ttcgggaggt cctacgacga gtacattctg gctctcaagg agtgggagaa cgacccggat
360

aatcccgcca gcaaggagag ggttcggaca cgcttcagga ccacggacga tgccctcctg
420

actggcgtgc cactcatggc tatccctggg ttcgagctgg ctaccctctc ggtctacgt
480

cagtcggcca acctgcacct cgccctcctg agggacgctg tgttcttcgg ggagaggtgg
540

ggcctcacc cagacgaacat caatgacctc tactcgcggc tgaagaactc tattcgcgat
600

tacaccaatc attgcgttag gttctacaac atcggcctgg ggaacctcaa tgtgattcgg
660

ccagagtact accggttcca gcgcgagctg acgatcagcg ttctggacct cgtggcgctg
720

ttccctaact acgatatccg gacttaccgg attcccacca agtcccagct cacgcgcgag
780

atctacacag acccaatcat ttcgccgggg gcccaggccg gctacacact ccaggacgtc
840

ctgagggagc cgcacctgat ggatttcctg aaccggctca tcatctacac tggggagtag
900

cggggcatcc ggcactgggc tggccatgag gtggagtcgt ctgcactgg catgatgacc
960

aacatccgct tccccctcta cggcacggcg gccacagccg agccaacacg cttcattact
1020

ccgtcgacct tccccggcct gaatctcttc taccggactc tgtcgcccc catcttcgc
1080

gacgagccgg gcgctaacat cattatccgc taccgcacct ccctggtgga gggcggtggg
1140

ttcatccagc cgaacaatgg cgagcagctg tacaggggtg ggggcaccct cgactccctg
1200

gatcagctgc ccctcgaggg cgagtcaccc ctcacggagt actcacacag gctgtgccat
1260

gtccggttcg ctcagtcctt ccggaacgcg gagccactgg actacgcccg ggtgcctatg
1320

ttctcatgga cacaccgctc cgccacacca actaatacca tcgaccctga tgttattact
1380

cagatcccac tcgtgaaggc gttcaacctg cattcgggcg ctactatcgt caagggccct
1440

gggttcaccg gcggggacat tctgcgagg acgaatgttg ggtctttcgg cgatatgcgc
1500

gtgaacatca cggcgcccct cagccagcgc tacaggggtgc ggattcgcta cgcttcgaca
1560

actgacctgc agttctacac caacatcaat gggaccacga ttaacatcgg caatttcagc
1620

tcgacgatgg attccgggga cgatctccag tacggcaggt tccgggtggc ggggttcaca
1680

actccgttca cgttcagcga cgctatgtcg acgttcacaa tcggcgcctt cagcttctct
1740

tcaaacaatg aggtgtacat tgaccgcacg gagttcgtcc ccgctgaggt taccttcgag
1800

gccgagtacg atctggagaa ggctcagaag gccgtcaatg cgctcttcac gtccagcaac
1860

cagatcggcc tgaagacgga cgtgacagat taccacattg acaaggtgag caacctcgtc
1920

gagtgcctgt cggacgagtt ctgcctcgat gagaagaggg agctgtccga gaaggtcaag
1980

catgcgaagc gcctctgcga cgagaggaac ctctccagg acccgaactt ccgcggcatc
2040

aatcggcagc ctgaccgcgg gtggaggggc tccactgata ttaccatcca gggcggggac
2100

gatgtgttca aggagaacta cgtcacgctc ccgggcacat tcgacgagtg ctacccaca
2160

tacctgtacc agaagatcga tgagtcaaag ctcaaggcct aactcggta cgagctgcgc
2220

ggctacatcg aggactccca ggatctcgag atctacctca tccgctacaa tgcgaagcac
2280

gagactgtta acgtgccagg caccgggtcg ctgtggcccc tctcggcgca gtccccaatc
2340

gggaagtgcg gcgagccaaa tcgctgcgcg acccatctcg agtggaaacc tgacctggat
2400

tgcagctgca gggacggcga gaagtgcgct caccattctc accatttctc actggacatc
2460

gatgtgggct gcaccgacct caacgaggat ctgggcgtct gggttatatt caagatcaag
2520

accaggacg ggcacgctag gctgggcaac ctcgagttcc tcgaggagaa gccgctgggtg
2580

ggcgaggctc tggctagggt taagaggcg gagaagaagt ggagggacaa gcgggagaag
2640

ctggagctgg agacaaacat cgtctacaag gaggcgaaga agtctgtcga tgctctcttc
2700

gttaattcac agtacgacag gctgcaggcc gatacgaaca tcgcgattat ccacgctgcc
2760

gacaagcggg tgcattccat tcgcgaggcc tacctcccag agctgagcgt tatccctggc
2820

gtgaacgcgg ctattttcga ggagctggag ggccgcatct tcaccgccta cagcctgtac
2880

gacgcgagga acgttattaa gaatggggat ttcaacaatg gcctctcgtg ctggaacgtg
2940

aagggccacg tggacgtcga ggagcagaac aatcatcgct ctgtcctcgt ggtcccggag
3000

tgggaggctg aggtttcaca ggaggtagg gtgtgcccg gcagggggta catcctccgc
3060

gtcacggcgt acaaggaggg ctacggggag ggctgcgtta caattcacga gatcgaggac
3120

aatactgatg agctgaagtt ctccaactgc gtggaggagg agatctaccc aaacaatact
3180

gtcacctgca acgactacac ggccacacag gaggagtacg agggcactta caccagccgc
3240

aatagggggt acgacggcgc ctacgagtcg aactcgtctg tgctgcgga ttacgcgtct
3300

gcttacgagg agaaggctta caccgacggc aggcgcgata acacatgcga gagcaatagg
3360

gggtacggcg actacacccc actccctgcg ggctacgtga cgaaggagct ggagtacttc
3420

ccggagaccg acaaggctctg gattgagatc ggggagacgg agggcacatt catcgtggat
3480

tccgtcgagc tgctgctcat ggaggagtaa
3510

<210> 17
<211> 3510
<212> ADN
<213> Trình tự Nhân tạo

<220>
<223> Trình tự mã hóa Cry tổng hợp

<400> 17
atggaggtca accaccagaa tgagtgcgtt ccatacaact gcctgaagaa tcctaagatc
60

gagatgctcg acatcgaggg catttcgtcc aggtcgaggg agcaggtggc cgagatctcc
120

ctgggcctca caaggttcct cctggagtct ctctgccgg gcgcttcatt cggcttcggg
180

ctgttcgata tcatttgggg cgtgatcggc ccggaccagt ggtcgctgtt cctcaccag
240

attgagcagc tcatcgacca gcgcattgag gcgcacgtga ggaaccaggc tatttcgagg
300

ctggaggggc tgggggactc ctacgaggtc tacatcgaga gcctgagggg gtgggaggct
360

tccccgaaca atgagtctct ccagcaggat gtgaggaacc gggtcagcaa caccgacaat
420

gctctgatta cggccatccc gattctccgc gagcagggct tcgagatccc gtcctgacc
480

gtctacgtgc aggtcgcaa cctgcacctc agcctcctgc gggatgcggt gtacttcggg
540

cagaggtggg gcctcgacac ggccacggtc aacaatcact acaaccggct gatcaacctc
600

attaatacct actccgatca ttgcgctcag tggttcaacc gcgggctgga caatttcggc
660

gtggtcactg cccgctacct cgatttcag agggaggtga ccattctagt cctggatatt
720

gttgcgctct tcccaaacta cgacatcagg acatacccta ttcagactct gtcccagctc
780

accggggaga tctacacgtc tccagtcgcg gagcctggcg cttcactcaa cgttgacctg
840

aggaatatcc tccgggagcc acacctgatg gatttcctga cagcctcgt gatctacact
900

ggcgtccagg gcgggatcta ccaactgggct gggcatgaga tctcgtctag gaccacgggc
 960
 aacctgtcat ccaatatcca gttcccactg tacggcacgt ccgccaacgc tgaccggcct
 1020
 ttcaatctgg cgatccatta ctccgagacc atctaccgca cgctcagcgc tccaatctac
 1080
 tccgtcagcg gcgggatttc gccgaacagg accagggccg ttgagggcgt gcgcttcctc
 1140
 acagctaggg acaacaatct gaactcgctc ccgttcctgt accggaagga gggctcactg
 1200
 gattccttca ccgagctgcc gccggaggac gagaacgagc caccttacat cggctactcc
 1260
 cacaggctgt gccatgccag gttcgctcgc agctcggttg tgctcgagcc gtctaaacttc
 1320
 gctcgcatcc cgggtgttctc atggaccac cgctccgccg gcccgacaaa tgaggtctct
 1380
 tcattccagga tcaactcagat tccttgggtt aaggctcata ccctggatag cggcgccttc
 1440
 gtgattaagg gcccggggtt caccggcggg gacatcctca ccaggccaaa cctcgggaca
 1500
 ctgggcgcgc tcagggtcac actgactggc cagctcccgc agacttacia catccggatt
 1560
 cgctacgcct ccattgcgaa tcgcggcggc accctcatct tctctcagcc gccatcatac
 1620
 ggcttcacat tcccgaagac tatggacatc gatgagcccc tgacgtcgcg ctctttcgcg
 1680
 aggacaactc tcttcacacc aattaccttc acgcaggccc aggcggagct gaacctcact
 1740
 atccagcagg gcgtgtacat cgataggatt gagttcatcc ctgtcaatgc cacgttcgag
 1800
 gcggagtacg acctggagcg ggctcagaag gccgtgaacg cgctcttcac cagctcgaat
 1860
 cagctgggcc tcaagacaga cctgactgat taccacattg atcaggtgtc gaacctggtc
 1920
 gactgcctct ctgatgagtt ctgcatcgac gagaagcgcg agctgagcga gaaggtaag
 1980
 catgccaaag ggctctcaga cgagcgcaac ctctgcagg attccaactt caggggcac
 2040

aataggcagc cggacagggg gtggaggggc tccaccgata tcacgattca gggcggaac
2100

gacgttttca aggagaatta cgtgaccctg ccaggcacgt tcgacgagtg ctaccctacg
2160

tacctctacc agaagatcga tgagtccaag ctgaaggcgt acacacgcta ccagctccgg
2220

ggctacatcg aggacagcca ggatctggag atctaccta tcaggtacaa cgccaagcac
2280

gagactgtta atgtgccggg caccgggagc ctgtggccac tcagcgtgga gtcgcctatc
2340

gggaagtgcg gcgagcccaa ccgctgcgtc ccacagctgg agtggaacag caatctcgac
2400

tgctcctgca gggatggcga gaagtgcgcc caccattccc accatttcag cctcgacatc
2460

gatgtggggg gcacggacct gcacgacgat ctcggcgtct gggttatttt caagatcaag
2520

acgcaggacg ggcattgctag gctgggcaac ctcgagttcc tcgaggagaa gccgctggg
2580

ggcgaggctc tggctagggg caagagggcg gagaagaagt ggcgggacaa gcgcgagaca
2640

ctgcagctcg agactaacat cgtttacaag gaggccaagg agtcggtgga cgctctgttc
2700

gccaactctc agtacaatcg cctccaggct gatacgaaca ttgccatgat ccacgctgcc
2760

gacaagaggg tgcattcgcat ccgcgaggct tacctccccg agctgtctgt tatccccggg
2820

gtgaacgcgg gcattttcga ggagctggag ggccggattt tcaccgcctt cagcctctac
2880

gatgcgcgca acgtcatcaa gaattcagac ttcaacaatg ggctgtcctg ctggaacgta
2940

aagggccacg ttgacatcga ggagcagaac aatcatagga gcgttctcgt cgttccggag
3000

tgggagggcg aggtgtcgca gaaggtgcac gtctgcccg gagggggta catcctgcgg
3060

gtcaccgcct acaaggaggg ctacggggag ggctgcgtta ctattcacga gatcaggagc
3120

cataccgatg agctgaagtt ccgcaactgc gaggaggatg aggtgtacct gaacaatata
3180

aggacttgca atgcgtaccc ggctgaccag gaggggtacg agggcgctg cacttctcgc
3240

aacaggggggt acgacgaggt ctacggcaat acccgcgcac tccccgccga ttacgcgcc
3300

atctacgagg agaacgccta cacggacggg aggaggggca atccttgcca gtcttcacgc
3360

gggtacggcg actacacgcc actgcctgcc ggctacgaga caaaggagct ggagtacttc
3420

cgggagaccg atacggtgtg gatcaagatt ggggagacag agggcacttt catcgctcac
3480

tccgttgagc tgctgctcat ggaggagtaa
3510

<210> 18

<211> 2160

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 18

atgaagctca agaatccaga caagcaccag agcctgtcca gcaacgctaa ggtggacaag
60

atcgccacag attcgctcaa gaatgagact gacattgagc tgaagaacat gaacaatgag
120

gattacctcc gcatgtcgga gcatgagtct atcgaccctt tcgtttcggc ctctactatt
180

cagaccggca tcgggattgc cggcaagatc ctcggcaccc tcggcgtgcc gttecgctggc
240

cagatcgctt ctctctactc attcattctg ggcgagctgt ggcctaaggg gaagagccag
300

tgggagatct tcatggagca cgtggaggag atcattaatc agaagattct gacctacgcg
360

aggaacaagg ctctgtcgga cctcaggggc ctgggggatg ccctcgctgt ctaccatgag
420

tccctggaga gctgggttaa gaatcgcaac aatacgcggg ctcgctccgt ggtcaagaac
480

cagtacatcg ccctcgagct gatgttcgtc cagaagctgc catcgttcgc cgtgtccggc
540

gaggaggtcc cgctcctgcc aatctacgcc caggctgcca acctccacct cctgctcctg
600

cgcgacgctt ctatcttcgg caaggagtgg gggctgtcag cgtccgagat tagcacattc
660

tacaacaggc aggttgagag gactcgggac tactcggatc attgcgtgaa gtggtacaac
720

accggcctca acaatctgag ggggacgaat gccaaagagct ggggccggta caaccagttc
780

cgcaaggaca tgacactcat ggtgctcgat ctggtcgcgc tcttcccatc atacgacact
840

ctggtctacc ctatcaagac cacgtcccag ctgactaggg aggtttacac cgatgccatt
900

ggcacggttc acccctcagg ggcggtggcc tccacaactt ggtacaacaa taacgcgccg
960

agcttctcga ccatcgaggc tgccgttgtg cggaaccccc acctcctgga cttcctcgag
1020

caggtgacca tctactcgct cctgtctagg tggtaaata cgcagtacat gaacatgtgg
1080

ggcgggcata agctcgagtt ccggactatc ggcggcaccc tcaacacaag cactcagggc
1140

tcgaccaata cgtctattaa ccagtgacg ctccctttca catctagggg cgtctaccgg
1200

acagagtcac tggccggcct caatctgttc ctcacgcagc ccgtgaacgg cgtccccagg
1260

gttgacttcc actggaagtt cgtcacccat ccaatcgct ccgataactt ctactaccct
1320

ggctacgctg gcattgggac gcagctccag gactccgaga atgagctgcc gcccgagacc
1380

accggccagc ccaactacga gtcttactca cacaggctct cccatatcgg cctgatttcc
1440

gccagccacg ttaaggcgct cgtgtacagc tggacacatc gctcggccga caggacaaat
1500

actatcaact cggattctat cacgcagatt ccgctcgtga aggccttcaa cctgcgctcg
1560

ggcgctcgg tcgttcgggg ccccggttc accggcgggg acatcctgcg caggaccaac
1620

acgggcacat tcggggatat ccgcgtcaat attaaccac ctttcgcccc gcgctacagg
1680

ctccgcatcc gctacgcgtc cacaactaat ctggagttcc acaccagcat caacggcaag
1740

gctattaatc aggggaactt ctcggtacc atgaaccgcg gcgaggacct cgattacaag
1800

gccttcagga cgggtgggctt caccacgccc ttctcattct ccaacgcgca gtcgactttc
1860

accatcggcg cttggaattt ctctctgggg aacgaggtct acatcgacag gattgagttc
1920

gtgccggtcg aggttaccta cgaggcggag cacgatttcg agaaggcgca ggaggaggtg
1980

acggctctct tcacgagcac aaatccaggc ggcctgaaga cgaacgtcac agagtaccat
2040

atcgaccagg tctcgaatct cgttgagagc ctgtcgaacg agttctacct cgatgagaag
2100

cgggagctgt tcgagatcgt gaagtacgcc aagcagctgc acattggccg caacatgtaa
2160

<210> 19

<211> 3495

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 19

atggagcgca acaatcagga ccagtgcac ccgtacaact gcctcaacaa tcccagaggt
60

ggcattctgg atatcgagaa cttcaatctg gagctgggtg cccaggtgtc agtcggcctc
120

accaggttcc tcctggaggc ctccatccca ggcgcgggt tcgctctggg cctcttcgac
180

atcatttggg gcgccctcgg cgtggatcag tgggccctgt tcctcgctca gattgagcag
240

ctgatcaatg agcgcattac caccgtggag aggaaccgcg ctatccaggc cctgtcgggc
300

ctctccagct cgtacgaggt ctacattgag gctctcaggg agtgggagaa cgacctggat
360

aatccggctt ctagggacag ggtggtggct aggttccggg ctaccgacaa ctcaactcatt
420

acggatatcc cgctcctgga gatccccggc ttogagattg ccacactgtc cgtgtacacg
480

caggcgcca acctgcacct cgccctcctg agggacgctg tctacttcgg ggagcgctgg
540

ggctgacac agactaacat cgaggacctc cacacacggc tgactcgcta cattcaggag
 600
 tactccgatc attgcgcgcg ctggtacaac cagggcctca acaatattgg cgggatcaat
 660
 acccgctacc tcgacttcca gcgggagctg acgatcagcg tcctcgacat tgttgccctg
 720
 ttcccgaaact acgatatccg gacctacca attcctacgc agagccagct cacacgcgag
 780
 atctacactt cgccggttgt gggccccggc gtgaattgga ttctgtcgat ctctaacgtc
 840
 ctgagggcgc cacacctgat ggactttctc gatcggatca ttatctacac gggcacggtg
 900
 cgctcaaccc ctactggga ggggcatgag gtcattctcca ggaggacggg ccaggggaac
 960
 gagattagga gcccgctcta cggcgtggct gctaattgcc agccgccagt tacgatcagg
 1020
 ccaacgggct tcacggacga gcagcggcag gtctaccgcg ttctgtccag ggtcgccagc
 1080
 ttccggaact cgggcacaaa tttctcactc gttgacgcgg ctctcttctt gactatcttc
 1140
 agcgcgtctt caatctaccg caacggcttc ggggtcaatg ctgacaccat cgatgagatt
 1200
 ccaatcgagg ggacggaccc ttacattggc tactcccacc gcctctgcc a tgtgggcttc
 1260
 accgcctcca gcccgttcat cagccagtac gcgagggctc ccgtcttctc ctggacacac
 1320
 cggagcgcta cattactaa caccatcgac ccagagcgca ttacgcagat ccctatggtg
 1380
 aaggcctaca acctgcatgc gggcgccaca gtcgttaggg gcccggggtt cactggcggg
 1440
 gacctcctgc ggcgacgaa cacaggcact ttcgcggata tcagggtcaa tattaccggc
 1500
 cccctctccc agcgtaccg cgtgcgcac cgtacgctt ctacaactga cctgcagttc
 1560
 ttcaccagga tcaacgggac gagcgtgaac cagggcaatt tccagcggac gatgaaccgc
 1620
 ggggacaatc tcgagtctgg caacttccgg acagccgggt tctcaactcc attctcattc
 1680

tccaatgctc agtccacctt cacgctgggc acccaggcct tcagcaacca ggaggtctac
1740

attgacagga tcgagttcgt gcctgctgag gtcacgttcg aggccgagtc cgatctcgag
1800

cgggctcaga aggccgtgaa cgcgctgttc acaagcactt cgcagctggg cctcaagaca
1860

aacgtcactg gctaccacat cgaccagggt tcgaatctcg tggcgtgcct gtctgacgag
1920

ttctgcctcg atgagaagcg cgagctgtcg gagaagggtga agcacgcaa ggggctctct
1980

gacaagcgca acctcctgca ggaccogaac ttcaggggca tcaataggca gcctgaccac
2040

ggctggaggg gctcgaccga tattacgatc cagggcgggg acgatgtttt caaggagaac
2100

tacgtgacct tcccaggcac gttcgacgag tgctacccta catacctgta ccagaagatc
2160

gatgagtcca agctcaaggc gtacactcgc taccagctga ggggctacat cgaggactca
2220

caggatctcg agatctacct catccgctac aattccaagc acgagatcgt taacgtgccc
2280

ggcaccgggt cgctgtggcc actctctgtg gagaatcaga tcgggccatg cggcgagcct
2340

aaccggtgcg ccccgcacct cgagtggaaat ccagacctgc attgctcctg cggggatggc
2400

gagaagtgcg tccaccattc tcaccatttc tcaactcgaca tcgatgttgg ctgcaccgac
2460

ctcaacgagg atctgggggt gtggctcatt ttcaagatca agaccagga cgggcacgcc
2520

aggctgggca acctcgagtt cctcgaggag gagccccctc tgggcgaggc tctggctagg
2580

gtgaagaggg ccgagaagaa gtggcgcgac aagagggaga agctgcagct cgagaccaac
2640

atcgtctaca aggaggcgaa ggagtcagtg gatgctctct tcgtcaactc ccagtacgac
2700

cgcctgcagg ccgatacgaa tattgcgatg atccacgccg cggacaagag ggtgcatcgc
2760

atccgcgagg cctacctccc agagctgtcg gtcacccctg gcgttaacgc tgccattttc
2820

gaggagctgg agggccgcat cttcaccgcc tactctctgt acgacgcgag gaacgtgatt
2880

aagaacggga atttcaacaa tggcctcctg tgctggaacg tcaagggcca cgtcgatgtt
2940

gaggagcaga acaatcatcg cagcgtgctc gtcacccccg agtgggaggc cgaggtctcg
3000

cagaaggtta ggggtgtgcc gggcaggggg tacatcctgc gcgtgaccgc ttacaaggag
3060

ggctacgggg agggctgcgt caccattcac gagatcgagg acaatacggg tgagctgaag
3120

ttctccaact gcgttgagga gggctacccg aacaataccg tgacgtgcaa cgagtacacc
3180

atgaaccagg ggggtgggca gtgcactgac gcttgcaacg ttcgcaatag gggctacgag
3240

gatgcgtacg ggcacaacc aagcaccct gtgcattaca ccacgccata cgaggaggag
3300

acatacactg acgagaggcg ggagaaccct tgcgaggcca ataagggcta cgtgaactac
3360

acaccgctcc ccgtgggcta cgtgactaag gagctggagt acttccccga gaccgacacg
3420

gtctggattg agatcgggga gacagagggc actttcatcg ttgatagcgt ggagctgctg
3480

ctcatggagg agtaa
3495

<210> 20

<211> 3501

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa Cry tổng hợp

<400> 20

atgaagatca acaatcagaa ccagtgcatt ccgtacaact gcctgagcaa tccggaggag
60

gtgctcctgg acggcgagcg cattctgccg gacatcgatc ccctcgaggt cagcatgtcg
120

ctcctgcagt tcctcctgaa caatttcgtg ccgggcgggg gcgtcatctc gggcctgatc
180

aacaagattt ggggcgccct ccggccgtcc gaggggagc tgttcctcgc ccagattgag
240

cagctgatcg acaggaggat tgaggcggcg gtcagggcta aggccatcgc tgagctggag
300

gggctggggc ggtcttacca gctctacggc gaggcgttca aggagtggga gaagacaccc
360

gataacaccg cggcgcgctc cagggtgacc gagcgccttca ggatcattga cgcgcagatc
420

gaggctaaca ttccatcttt ccgcgtttca ggcttcgagg tgccctctct gtctgtctac
480

accagggcgg ctaacctgca tctcgccctc ctgcgggatt cagtgggtgt cggcgagagg
540

tgggggctga ccaccacaaa cgtgaatgac atctacaacc gccagggtcaa gaggattgac
600

gagtacagcg atcactgcgt cgacacctac aagacggagc tggagcgcct cgagttctcc
660

agcatcgcgc agtggaggat ctacaaccag ttccggaggg agctgacctt caccgtcttc
720

gatatcggtg ccctgttccc gaattacgac ggccgcctgt accccatcag gaccatttcc
780

cagctcacgc gggacatcta cacatctcca gtttcagagt tctactacgg ccctatctac
840

aactacaata ttgtggggag gctcactgag cagcagctga ggcggccaca tctcatggat
900

ttcttcaact ccatgatcat gtacaccagc gacaatcgca gggagcacta ctgggtccggc
960

ctggagatga aggccacaga cactagcggg aaccagggtg cgttccccct ggctggcacg
1020

agggggaatt ccgccccgcc agtcacagtt aggaacaatg gggagggcgt ctaccgcac
1080

ctctctgagc cattctactc gtctcctttc ctgggcacct ccgtgctggg ctcccgcggc
1140

gaggagttcg ccttcgcgag caacactacc acgtccctgc cgagcactat ctacaggaat
1200

cgcggcaccg tggattcgct cgtctctatt ccacctcagg actactcggg cccgccccac
1260

aggggctact cccacctcct gtctcatggt acgatgcaca actcatcccc gatcttccac
1320

tggacacatc gctccgccac cccacgcaat acgattgacc ctgatagcat cacgcagatt
1380

cccgttgtga aggcctcaca tctgtccggg ggcagcgta tcaaggggcc gggccacaca
1440

gggggcgacc tgatttcctt ccccgtaac aatttctctc acttccgcat cccggtccag
1500

gccaacaccc cgcagcgcta ccgcatccgc atccgctacg ctgccgactc ggatggcacc
1560

ctggattcgg ggggtgttct ctcagctgcg gccggggacg gcttcaacac aacttcgtac
1620

agggctacca tgtctccggg cggctccctc acatctaggg atttccagtt cctggacctc
1680

aacacatcat tcacttccga cgtcgcgtcc aatctgtggc tccatttcat caggtacatt
1740

cggccaggca acctgtacat cgatagggcc gagttcttcc ctgtggacgc caccttcgag
1800

gcgggctaca atctcgagcg ggcgcagaag gctgtgaacg ccctgttcac ctccacgaat
1860

cagaagggcc tccagaccga cgtcacggat taccacatcg atcagggttc aaacctcgtg
1920

gattgcctgt ccgacgagtt ctgcctggac gagaagcgcg agctgagcga gaagggtcaag
1980

caggcgaaga ggctgtctga tgagcggaac ctctgcagg actcaaactt ccggggcacc
2040

aatagggagc aggatagggg ctggaggggg tcgacacata tcaactattca gggcggcaac
2100

gacgtcttca aggagaattt cgttactctg ccgggcgcgt tcgatgcttg ctaccccacc
2160

tacctctacc agaagatcga cgagtccaag ctgaaggcct acacccgcta cgagctgcgg
2220

ggctacatcg aggacagcca ggacctcgat atctacctga ttcgctacaa cacaagcac
2280

gagactctca atgtgccggg cactaagtcc ccctgggtccc tgtgcacgga gtccccactc
2340

ggcaagtgcg gggagcctaa cagggtgcgc tccagatcg agtggaatcc ggacctggat
2400

tgctcctgca gggacggcga gaagtgcgct caccattcac accatttctc cctcgacatc
2460

gatgtgggct gcacggacct caacgagaat ctgggggatct gggtcatttt caagatcaag
2520

acgcaggacg gccacgccag gctggggaac ctcgagttcc tcgaggagaa gccgctcctg
 2580
 ggcgaggctc tggctagggc caagcgcgcc gagaagaagt ggagggacaa gcgggagaag
 2640
 ctccagagcg agacgaacat cgtgtacaag gaggccaagg aggcgggtga tggcctgttc
 2700
 gtggactcgc agtacgagcg cctccaggcg gataccaaca ttgctatgat ccacgctgcc
 2760
 gacaagcacg tgcacagat caggagagtc tacttcccag agctgtccgt catccctggc
 2820
 gttaacgcgg ctattttcga ggagctggag ggccggatct tcacggccta ctcaactctac
 2880
 gatgcgcgca acgtcattaa gaatggcgac ttcaacaatg ggctgtcctg ctggaacgtt
 2940
 aagggccatg tggacgtcga ggttcagaac aatcaccgct cagtgtcctg catcgtcgag
 3000
 tgggaggcgg aggtgtccca ggaggttccg gtgtgcccg ggaggggcta catcctgagg
 3060
 gttaccgcct acaaggaggg gtacggcgag ggtgctgta cgattcatga gatcgaggac
 3120
 cacacagatg agctgaagtt caggaactgc gaggaggagg aggtttaccc gaacaataca
 3180
 gtgacttgca atggctacac cgcgacgcag gaggagtaca aggatgctta cacctcgagg
 3240
 aacaggggct acgacgaggc ctacgggaac aatccaagcg tccctgctga ttacgcctcg
 3300
 gtttacgagg agaaggctta cacggacggc aggaggaga acccatgcga gatggagagg
 3360
 ggctacactc cactgcctgt ggggtacatc accaaggagc tggagtactt ccctgagaca
 3420
 gatactgtct ggatcgagat tggcgagacc gaggggacgt tcatcattga cagcgtggag
 3480
 ctgctgctca tggaggagta a
 3501

<210> 21
 <211> 3546
 <212> ADN
 <213> Trình tự Nhân tạo
 <220>

<223> Trình tự mã hóa BT-0001 đột biến tổng hợp

<400> 21

atggacaata atccgaacat caacgagtgc atcccttaca actgcctgtc taaccctgag
60

gtggaggtgc tggggggcga gcgcatcgag acaggctaca ctccaatcga catttcgctg
120

tctctcacgc agttcctcct gtcggagttc gtgcctggcg ctgggttcgt gctgggcctc
180

gtcgatatca tttgggggat cttcggtccc tccagtgagg acgctttcct ggtccagatc
240

gagcagctca ttaaccagcg catcgaggag ttcgctagga atcaggccat ttcgcggctg
300

gagggcctct ctaacctgta ccagatctac gcggagtcct tccgcgagtg ggaggctgat
360

ccgaccaacc ccgccctgag ggaggagatg cggattcagt tcaacgacat gaattcggcg
420

ctgaccacgg ctatcccgtc cttcgccgtc cagaactacc aggttcccct cctgtctgtg
480

tacgtccagg ccgcgaatct gcacctctca gtgctgcgcg acgtttccgt gttcgggcag
540

cgctggggct tcgacgtgc tactattaac tcccgtaca atgatctcac taggctgac
600

ggcaactaca ccgactacgc cgtgcgctgg tacaataccg ggctcgagag ggtttggggc
660

ccggacagca gggattgggt caggtacaac cagttccgca gggagctgac cctcacggtt
720

ctggacatcg tggccctctt ctggaactac gattctcggc gctaccaat taggaccgtg
780

agccagctga cgcgggagat ctacacaaat cctgtcctcg agaacttcga cgggtcgttc
840

aggggcatgg ctacgcgcgt tgagcagaat atccgccagc cgcacctgat ggatatactc
900

aattcgatca ccatttacac ggacgtgcac aggggggttca actactggtc tggccatcag
960

atcacgcgtt ctccagtcgg gttctcaggc ccggagttcg ctttcccact gttcgggaac
1020

gctggcaatg ctgctccacc cgtcctagtt tcgctgacgg gcctcgggat cttccgcaca
1080

ctgtccagcc cactctacag gcggatcatt ctcggtcttg ggcctaacaa tcaggagctg
1140

ttcgtgctcg acggcacgga gttctctttc gcgtcactga caactaacct cccatccact
1200

atttacaggc agagggggcac cgtggacagc ctcgatgtaa tcccacctca ggacaactcg
1260

gtcccgccaa gggctggctt ctcacaccgc ctgtcccatg taaccatgct cagccaggct
1320

gctggcgctg tttaactctt gagggctccc accttctcct ggcagcaccg gagcgccgag
1380

ttcaacaata tcattgcgtc agattccatc acgcagattc cggcgggtgaa ggggaacttc
1440

ctcttcaatg gctccgtaat tagcggcccc gggttcacag gcggggacct ggtgcgcctc
1500

aactcgtctg ggaacaatat ccagaatagg ggatacatcg aggtcccaat ccacttccct
1560

agcacatcga ctcgctaccg cgtgagggtc cgctacgcta gcgtgactcc gattcatctg
1620

aacgtcaatt ggggcaattc atccatcttc tcgaacaccg tgcccgctac cgctacgtct
1680

ctggacaacc tccagtcctc cgatttcggc tacttcgagt ctgctaatac cttcacgtct
1740

tcactagggg atatcgttgg cgtgcggaac ttctcagggg cagccggcgt tatcattgac
1800

cgcttcgagt tcattcccgtt gacagccact ctggaggcgg agtacaacct cgagcgcgct
1860

cagaaggccg tgaacgcgct gttcacctcc acgaaccagc tgggcctcaa gacaaatgta
1920

actgactacc acatcgatca agtcagcaac ctggttacct gcctctcgga cgagttctgc
1980

ctggatgaga agagggagct gagcgagaag gtgaagcacg ccaagaggct ctgagacgag
2040

cggaacctcc tgcaggactc caatttcaag gatataca ggcagccaga ggggggtg
2100

ggcgggtcaa ccggcatcac gattcagggc ggggacgatg tcttcaagga gaactacgta
2160

acactgagcg gcactttcga cgagtgtctc ccacttacc tctaccagaa gatcgacgag
2220

tcaaagctga aggcgttcac ccgctaccag ctccggggat acatcgagga ctcccaggat
2280

ctggaggtgt acctcatccg ctacaacgcc aagcacgaga cgctgaacgt cccaggcaca
2340

gggtctctgt ggccactcgc ggttaagtca ccaattgggc gctgcggcga gccaaacagg
2400

tgcgctccta ggatcgagtg gaagcctgac gtggattgct cctgcaggga cggcgagaag
2460

tgcgctcacc atagccacca tttctcgctc gacatcgatg tcgggtgcac ggacctgaac
2520

gaggatctcg gcgtctgggt tatcttcaag attaagacac aggacgggca tgccaagata
2580

ggcaacctgg agttcctcga ggagaagctc ctgctcggcg aggctctggc tagggtgaag
2640

aaggccgaga agaagtggcg cgacaagagg gagaagctcg agtgggagac caacatcgtc
2700

tacaaggagg ccaaggagtc cgttgacgcg ctgttcgtgg atagccaata caacaggctc
2760

cagacagata ctaatatcgc catgattcac gctgccgaca agcgggtgca tcgcatccgc
2820

gaggcgtacc tgccagagct gagcgtgatt cctggcgctca acgcggctat cttcgaggag
2880

ctggagggcc tcattttcac ggctttctcc ctgtacgacg cccgcaacgt gatcaagaat
2940

ggggatttca actacggcct aagctgctgg aacgtcaagg gccacgtgga cgtcgaggag
3000

cagaacaatc ataggtccgt tctggtgatc ccggagtggg aggctgaagt cagccaggaa
3060

gtcagggttt gcccaggcag gggatacatc ctccgggtta ccgcctacaa ggagggtac
3120

ggggagggct gcgtgacgat ccacgagatt gaggacaaca cagatgagct gaagttcagc
3180

aattgcgttg aggaggaggt gtacccgaac aataccgtga cgtgcaacga ctacacagcc
3240

actcaggagg agtacgaggg cacctacacg tcgaggaacc gggggtacga cggcgcttac
3300

gagtctaatt ccagcgtccc agccgactac gcctcagcgt acgaggagaa ggcgtacacg
3360

gacggccgca gggataatcc ttgcgagtcc aaccggggct accgcgatta cactccactg
3420

cctgccggct acgtcaccaa ggagctggag tatttcccgg agacagacaa ggtttggatc
3480

gagcttggcg agacggaggg gaccttcctt gtggattcgg ttgagctgct gctgatggag
3540

gagtag
3546

<210> 22

<211> 3531

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0003 đột biến tổng hợp

<400> 22

atggagatcg tgaacaatca gaatcagtcg gtgccttaca actgcctcaa taatcccagag
60

atcgagatcc tggagggggg gcgcatttcg gtgggcaaca cgcctatcga catttcctc
120

agcctgacac agttcctcct gtctgagttc gttccaggcg ctgggttcgt gctcggcctg
180

atcgacctca tttggggggt cgtggggccc tcccagtgagg atgctttcct cgcccaggtg
240

gagcagctga tcaaccagag gattgcggag gctgttcgga ataccgcgat ccaggagctg
300

gagggcatgg ctgcgctgta caggacatac gccactgcgt tcgctgagtg ggagaaggct
360

ccagacgac cagagctgag ggaggctctg aggacacagt tcacggcgac agagacttac
420

atctccggcc gcattagcgt cctcaagatc cagaccttcg aggttcagct cctgagcgtg
480

ttcgctcagg ctgctaacct ccacctgtcg ctctgaggg acgtggtctt ctccgggcag
540

cgctggggct tctcgaccac gacagtgaac aattactaca acgacctcac tgagggcatc
600

tctacctaca cggattacgc cgtccgctgg tacaataccg ggctggagag ggtgtggggc
660

cccagacgca gggattgggt ccggtacaac cagttccgca gggagctgac actgactgtg
720

ctcgacatcg tcgcgctggt ctcgaaactac gattctcggc gctaccccat ccggacagtg
 780
 tcgcagctca ctgcgagat ttacaccaac ccagtcctgg agaatttcga cgggtccttc
 840
 aggggcatgg ctcagcggat cgagcagaac attaggcagc cccacctcat ggacatcctg
 900
 aattccatca ccatttacac ggacgtgcac cgcgggttca actactggag cggccatcag
 960
 atcacgcgtt ctccagtcgg gttctcaggc ccggagttcg ctttccccct cttcgggaac
 1020
 gctggcaatg ctgctccacc cgtgctagtc tcctgactg gcctcgggat cttcaggacc
 1080
 ctctccagcc cactgtacag gcggatcatt ctgggcagcg ggcctaacaa tcaggagctg
 1140
 ttctgtctcg acggcacgga gttctcattc gcgtccctca ctaccaacct gccctccacg
 1200
 atctacaggc agcggggcac agtcgactct ctcgatgtaa ttccacctca ggacaactca
 1260
 gtcccgccaa gggctggctt ctcacacagg ctctcccatg ttaccatgct gagccaggct
 1320
 gctggcgctg tgtacacgct gcgggctccg acattctcct ggagcaccg cagcgcggag
 1380
 ttcaacaata tcattccgtc gtctcagatc acgcagattc ccctcacaaa gtccactaac
 1440
 ctgggcagcg ggacatcggg tgtgaagggc ccgggggttca ctggcgggga catcctcagg
 1500
 aggacctcgc caggccagat ctctactctc cgcgttaata ttaccgctcc actgagccag
 1560
 cgctacaggg tccgcacccg ctacgcctca acgacaaacc tgcagttcca cacgtccatc
 1620
 gacgggaggg ctattaacca gggcaatttc tctgccacta tgtcatccgg gtcaaacctc
 1680
 cagtcgggct ctttccgcac cgtgggggtt actaccccg tcaacttctc caacggcagc
 1740
 tcggtgttca ccctgtcagc gcacgtcttc aactccggca atgaggttta catagaccgg
 1800
 attgagttcg ttccggccga agtgactttc gaggcggagt acgatctcga gcgcgcgcag
 1860

aaggctgtga acgagctgtt cacgtcttca aatcagatcg gcctgaagac cgacgtcacg
1920

gattaccaca ttgaccaggt ctccaacctc gttgagtgcc tgagcgacga gttctgcctc
1980

gacgagaagc aggagctgtc cgagaagggtg aagcatgcca agcgcctgag cgacgagagg
2040

aacctcctgc aggacccgaa cttccgcgga atcaacaggc agctggacag ggggtggagg
2100

ggcagcacag atatcactat tcagggcggg gacgacgtct tcaaggagaa ctacgttacc
2160

ctcctgggca cgttcgacga gtgctacca acgtacctgt accagaagat cgacgagtca
2220

aagctcaagg cgtacacacg ctaccagctg cgcgataca tcgaggacag ccaggatctc
2280

gagatctacc tgattaggtg caacgccaag cacgagacgg ttaacgtgcc tggcacaggc
2340

tccctctggc cactctctgc tcagtcaccc atcgggaagt gcggcgagcc aaacagggtgc
2400

gtcctcatc tcgagtggaa tccggacctg gattgcagct gcagggacgg cgagaagtgc
2460

gtcaccata gccaccattt ctgctcgac atcgatgtag gctgcacgga cctcaatgag
2520

gatctagggg tctgggttat cttcaagatt aagacacagg acgggcacgc taggctgggc
2580

aacctcgagt tcctggagga gaagccactc gtgggcgagg ctctggctag ggtcaagatc
2640

gccgagaaga agtggagggg caagcgggag aagctggagt gggagacgaa catcgtgtac
2700

aaggaggcta aggagtcggt ggatgccctc ttcgtaaact ctcagtacga ccagctgcag
2760

gctgatacca atatcgccat gattcacgcc gcggacaagc gggtcatttc gatccgcgag
2820

gcctacctcc cagagctgtc tgtgatccct ggcgtcaacg ctgccatttt cgaggagctg
2880

gagggcagga tcttcaccgc gttcagcctg tacgacgctc ggaacgtgat taagaacggg
2940

gatttcaaca atggcctctc gtgctggaac gtaaagggcc acgtggacgt cgaggagcag
3000

aacaatcagc gctctgtcct ggtcgttcct gagtgggagg ccgaggtttc acaggaagtt
3060

aggggtgtgcc caggcagggg atacatcctc cgcgtgaccg cgtacaagga gggctacggg
3120

gagggctgcg tcacgatcca tgagatagag aacaatacag acgagctgaa gttcagcaac
3180

tgcgtggagg aggagatcta ctgaacaat accgttacgt gcaacgatta cactgtgaat
3240

caggaggagt acggcggggc ttacacctca aggaacaggg gctacaacga ggctccatcc
3300

gtccctgctg actacgctag cgtttacgag gagaagtcgt acacggacgg caggagggag
3360

aacctatgtg agttcaatcg cggctacagg gactacactc cactccctgt cggctacgtt
3420

accaaggagc tggagtatatt cccagagacc gataaggtct ggctcgagat tggcgagact
3480

gaggggacat tcattgtgga tagcctggag ctgctgctga tggaggagtg a
3531

<210> 23

<211> 3714

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0020 đột biến tổng hợp

<400> 23

atgaatagca acaggaagaa tgagaacgag atcattgacg cttcgttcat cccagccgtc
60

tcgaacgagt ctgttactat ttctaaggag tacgcgcaga ccaatcagct ccagaacaat
120

tcaattgagg acgggctgtg catcgctgag ggcgagtaca ttgatccttt cgtgtcagcg
180

tccaccgtcc agacggggat ctccattgct ggccgcatcc tgggggtcct gggcgttcca
240

ttcgttggcc agctcgctc attctactcc ttcacgtgg gggagctgtg gcctaagggc
300

agggaccagt gggagatctt catggagcac gttgagcagc tcgtgcgcca gcagattact
360

gctaatacca ggaacacggc cctggctcgg ctccagggcc tggggaactc cttccgcgcc
420

taccagcaga gcctcgagga ctggctggag aataggaacg atgcccgga cgcagcgtg
480

ctctacacac agtacatcgc gctggagctg gacttcctca acgcgatgcc actgttcgct
540

attcgcgagc aggaggttcc tctcctgatg gtgtacgctc aggcggccaa cctgcacctc
600

ctgctcctgc gcgatgcctc gctctacggg agggagttcg gcctgacttc tcaggagatc
660

cagcgctact acgagaggca ggtcgagagg acccgggact actcagatca ttgcgtgcag
720

tggtacaaca cggggctcaa caatctgagg ggaaccaacg cggagagctg ggtccgctac
780

aatcagttcc gcagggacct gaccctgggc gtcttgatc tcgttgccct cttcccgta
840

tacgacacca ggacgtaccc aatcaacaca tccgcgcagc tgactcggga ggtctacacc
900

gatgccattg gggcgactgg cgttaacatg gcttcgatga attggtacaa caataacgcc
960

ccaagcttct cggcgatcga gaccgctgtg attaggtctc ctcacctctt ggacttcctc
1020

gagcagctga agatcttctc ggctccagc aggtgggtcta acaccggca catgacgtac
1080

tggaggggcc atacgatcca gagcaggccc attcggggcg cctcatcac aagcactcac
1140

ggcaatacca acacgtcgat caaccagtc acgttccagt tcccttctcg ggacgtttac
1200

cgcacagagt catacgccgg ggtgctcctg tggggcatct acctcgagcc aatccacggc
1260

gtgccgacgg tcaggttcaa tttccggaac cccagaata cgttcgagcg cggcacagcg
1320

aactactctc agccgtacga gtcacccggc ctgcagctca aggactccga gacggagctg
1380

ccgccggaga ccaccgagag gccaaactac gagtcttact cacaccggct gtcccatatc
1440

ggcatcattc tccagactag gctgaatgtg cctgtctact cctggacca ccgcagcgct
1500

gacaggacaa acactatcgg cccaaatcgg atcacgcaga ttctgcccgt gaaggggaac
1560

ctcctgttca atggcagcgt catctcgggc ccgggggttca caggcgggga cctcgtccgc
1620

ctgaataaca gcggaataa catccagaac aggggctacc tcgaggtgcc cattcagttc
1680

acctccacga gcacacgcta cagggttcgg gtgcgctacg cttccgtgac ccccatccac
1740

ctgtcagtca actggggcaa ttccaacatt ttctcgtcta cagtgcccg cactgctgcc
1800

agcctcgaca atctgcagtc gcgggatttc ggctacttcg agtctaccaa cgccttcaca
1860

tcagttactg ggaatgtggt cggcgtgcgc aacttctccg agaatgcgcg cgttatcatt
1920

gacaggttcg agttcatccc agtgactgcc accttcgagg cggagtacga tctggagagg
1980

gctcaggagg ccgtcaacgc gctcttcacg aatacaaacc cgaggaggct gaagacgggc
2040

gtgaccgact accacatcga tgaggtgtcg aacctcgtcg cctgcctgtc tgacgagttc
2100

tgacctgatg agaagcgcg gctgctggag aaggtgaagt acgcgaagcg gctgagcgac
2160

gagcgcaacc tcctgcagga cccgaacttc acgtccatca ataagcagcc cgacttcatt
2220

agcacgaacg agcagtcgaa tttcacatct atccacgagc agtcagagca tgggtggtgg
2280

ggctccgaga acatcaccat tcaggagggc aacgatgttt tcaaggagaa ttacgtgacc
2340

ctcccaggca cgtacaacga gtgctaccct acgtacctgt accagaagat cggcgagtcg
2400

gagctgaagg cctacacacg ctaccagctg cgcggataca tcgaggactc tcaggatctc
2460

gagatctacc tgattaggtg caacgcgaag cacgagactc tggacgtgcc cggcaccgag
2520

tcgggtctggc cgctgtctgt tgagtccccg atccgcaggt gcggcgagcc caacagggtgc
2580

gctccccatt tcgagtggaa tccggacctc gattgctcct gcagggacgg cgagaagtgc
2640

gccaccatt cacaccattt ctccctggac attgatgtcg ggtgcatcga tctccacgag
2700

aacctgggcg tgtgggttgt gttcaagatc aagacgcagg aggggcatgc taggctgggc
2760

aacctggagt tcattgagga gaagcccctc ctgggcgagg ctctctccag ggtcaagcgc
2820

gcggagaaga agtggaggga caagcgggag aagctgcagc tcgagactaa gcgcgtgtac
2880

accgaggcta aggaggccgt ggatgcgctc ttcgtcgaca gccagtacga taggctgcag
2940

gccgacacga acatcggcat gattcacgcg gctgataagc tgggtgcatcg catccgcgag
3000

gcgtacctct ccgagctgag cgttatcccg ggcgtgaacg ctgagatttt cgaggagctg
3060

gaggggcgga tcattaccgc gatctccctg tacgacgctc gcaacgctcg taagaatggg
3120

gatttcaata acggcctcgc ctgctggaac gtcaagggcc acgtcgacgt tcagcagagc
3180

caccatcgct cggtgctggt catccccgag tgggaggctg aggtgagcca ggcggtagg
3240

gtgtgcccgg gcaggggata catcctccgc gtcaccgcgt acaaggaggg ctacggggag
3300

ggctgcgtga ctatccacga gattgagaat aacaccgacg agctgaagtt caagaactgc
3360

gaggaggagg aggtctaccc aactgacacc gggacgtgca atgattacac ggcccatcag
3420

ggcacagccg cgtgcaactc caggaatgct gggtagcagg acgcctacga ggtcgataca
3480

actgcgagcg ttaactacaa gcccacctac gaggaggaga catacactga cgtgcgcagg
3540

gataaccact gcgagtacga ccgcggctac gtcaattacc cacctctccc ggccggctac
3600

gtgacgaagg agctggagta cttccccgag acagacaagg tctggatcga gcttggggag
3660

acagagggca ctttcctcgt ggatagcatt gagctgctgc tcattggagga gtaa
3714

<210> 24

<211> 2148

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0022 đột biến tổng hợp

<400> 24

atgaagtcta agaatcagaa catgcatcag agcctcagca acaacgctac tgtggataag
60

aatttcacgg gcagcctgga gaataacaca aatactgagc tgcagaattt caaccacgag
120

ggcatcgagc cattcgtctc agtttccaca attcagactg gcatcgggat tgcgggcaag
180

atcctgggca acctcggcgt gccgttcgct ggccagggtg cttccctcta cagcttcac
240

ctgggcgagc tgtggcccaa ggggaagagc cagtgggaga ttttcatgga gcatgtcgag
300

gagctgatca atcagaagat ttcgacctac gcccggaaca aggctctggc tgacctcaag
360

ggcctggggg atgctctcgc tgtgtaccac gagtcgctgg agtcttggat caagaaccgc
420

aacaatacca ggacgcggtc tgtggtcaag tcacaatata ttaccctcga gctgatgttc
480

gtccagtcac tgccgtcctt cgcggtttcc ggcgaggagg tgccgctcct gccaatctac
540

gctcaggctg cgaatctcca tctcctgctc ctgcgcgacg cctcgatctt cggcaaggag
600

tgggggctga gcgattcgga gatctctact ttctacaacc ggcaggtcga gcgcaccagc
660

gactactcgg atcactgcac aaagtgggtc gacactggcc tcaataggct gaaggggtcc
720

aacgcggaga tctgggtgaa atacaaccag ttccgcaggg acatgactct catggtcctc
780

gatctgggtg ccctgttcca gagctacgat acacacatgt acccaatcaa gaccacggct
840

cagctcacca gggagggtta cacgaacgcc attggcacag tgcaccacaca tccttctttc
900

acctcaacaa cttggtacaa caataacgct ccgtctttct cagccatcga ggctgccgct
960

attcggagcc cccatctcct ggacttcctg gagcagggtta ctatctactc cctcctgagc
1020

aggtagtcga atacccaata catgaacatg tggggcgggc acaagctcga gttccggaca
1080

attggcggga ctctgaacac atccactcag ggcagcacca atacgtcgat caaccagtc
1140

accctccctt tcacgtcaag ggacatctac cggacggagt ccctggccgg cctcaatctg
1200

ttcctcacac agccagttaa cggggtgccc agggtcgact tccactggaa gttcgtgacg
1260

catccaatcg cgtccgataa cttctactac cctggctacg ctggcattgg caccagctc
1320

caggacagcg agaatgagct gccaccagag accacggggc agccaaacta cgagtcctac
1380

agccacagcg tctcccatat cggcctgatt tcggcgtctc acgtgaaggc tctcgtctac
1440

tcttggaacc atcgctcagc ggacaggaca aacactatcc actcagattc catcacgcag
1500

attcctctcg tcaaggccca cacactgcag agcggcacia ctgttgtgaa gggcccaggg
1560

ttcacgggcg gggacatcct gaggcgcaca tccggcgggc ctttcgcgtt cagcaatggt
1620

aacctcgatt ggaatctgag ccagcgctac cgcgcccga tccgctacgc gtcgaccacg
1680

aacctccgga tgtatgtgac gatcgcgggc gagegcattt tcgctgggca gttcaataag
1740

accatgaaca cgggcgaccc actgaccttc cagagcttct cgtacgctac aatcgatact
1800

gccttcacct tccctacgaa ggcctccagc ctcaactgtgg gcgctgacac cttctcgtct
1860

gggaacgagg tctacgtgga tcgcttcgag ctgatcccag tgaccgctac gctcgaggct
1920

gtcacggacc tggagagggc tcagaaggcc gtgcatgagc tgttcacctc cactaatcca
1980

ggcgggctga agaccgacgt cgctaaggat cactacacca acacgctctc taagtcagtt
2040

cagtcggtgt tcaggtgccg gtgctctgag cgcacgaggc tctaccgctg gggctacccg
2100

tctaagaagg agtactggtg catttggggc tacacctcaa agtactga
2148

<210> 25
<211> 3285
<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0027 đột biến tổng hợp

<400> 25

atgttcctcg agcagattga gcagctgac gaccagcgca ttgagaccgt cgagaggaat
60

cggcgcatcc agacgctcat tggcctgtcg aactcttacg atgtgtacat cgagggcgctc
120

aaggagtggg agaacaatcc ggacaattca gcttcccagg agcgcgtcag gaaccgggttc
180

cgcaccacgg acgatgcgct catcacaggc attccgctcc tggctattcc caatttcgag
240

atcgccactc tgtcagtcta cgttcaggcg gccaacctcc acctgtccct cctgagggac
300

gctgttttct tcggcgagcg ctgggggctg acccagatca atgtggacga tctctaccgc
360

aggctgacga acaatatccg gaactactcc gatcattgcy cgcgctggta caatgagggc
420

ctcgacaaca ttagcgggct gagccggtcg atcaatttcc agcgcgaggt cacaatttcg
480

gttctcgata tcgtggccct gttcccgaac tacgacattc gcacttacc catcagcaca
540

acttcgcagc tcaccaggga gatcttcacg tcgccgatcg tgggtcccaa cgacttctcc
600

gttgccctacg agggcggtcg gcgcgctccc cacctgttcg agttcctcga gaagctcgtg
660

atctacacgg gcgatcgag cgggatcagg cactgggccc gccatgagat tacgtctagg
720

cggacagact cataccacgg catcatccgc taccgctct acgggacagc tgccaacgcc
780

gagtccccat acactctcgc gctgcagcct tctgagtcaa tctaccgcac cctctccgag
840

cctatatttca gccagacggg cgggctgtct ccgcacagga ggaggggttg ggagggcgctc
900

gagttctcaa tcgtcaacaa taacgttaac ccgtccagct tcgtgtaccg caggaagggc
960

tctctcgatt cattcacaga gctgccgccc gaggacgagt ctgtcccacc atacatcggc
1020

tactcacacc agctctgcca tgtgggcttc gggaggacga atgtcatctt cgagccgtcg
 1080
 aacttcgcca gggtgcccggt cttctcgtgg acacaccggg ctgcgacacc aactaatacc
 1140
 atcgaccctg atcggatcac tcagattcct tccgtgaagg cgtcgtctct gcgcaactcg
 1200
 acagtcgttt ctggcccggg gttcactggc ggggatatcg tccgcatggg cgctgttcac
 1260
 cagatctacg ccaccgacct cagcatgaac gtgaggccgt cagtcgctct gtcccgtac
 1320
 ctcatcaggc tgcggtacgc ctgcaggggc tcatccaata tcgttattca tgggccatcc
 1380
 atccggttcg tgtccctccc tagcaccatg tcgaacgacg agccgctgac gtaccagtca
 1440
 ttccgctacg cctccatcac cagcceaatt accaggccta tctacaatat gttcaacctc
 1500
 tccatcagcc ggattagcgg cgtccagaat ctgttcatcg accgcattga gttcatcccc
 1560
 gtggatgcga acttcgaggc tgagaggac ctggagcgcg ctcagaaggc tgtcaatgcg
 1620
 ctcttcacgt ccacaaacca gaggggcctg aagattgacg tgacggatta ccacatcgat
 1680
 caggttagca acctcgtgga ctgcctgtcg gatgagttct gcctcgacga gaagcgcgag
 1740
 ctgtctgaga agtcaaagca tgccaagagg ctcagcgacg agcggaaacct cctgcaggac
 1800
 ctgaatttca aggatatcaa ccgccagccg gagagggggg ggtccggctc tactgggac
 1860
 accattcagg gcggggacga tgttttcaag gagaactacg tgactctccc cggcaccttc
 1920
 gatgagtgct acccaacctc cctgtaccag aagatcgacg agtccaagct caaggcttac
 1980
 acgcgctacc agctgagggg atacatcgag gacagccagg atctcgagat ctacctcatc
 2040
 aggtacaatg ccaagcacga gaccgttaac gtgccaggct caggctccct ctggcctctg
 2100
 tccgtcgagt cctccgttgg caagtgcggg gagccaaata ggtgcgccag ccgcatggag
 2160

tggaaccag acctcgattg ctctgcagg gacggcgaga agtgcgctca ccatagccac
2220

catttctcgc tggacatcga tgtcggctgc accgacctca acgaggatct gggggctctgg
2280

gttattttca agatcaagac gcaggacggc cacgccaaga tcgggaacct cgagttcctg
2340

aaggagaagc tgctgctggg cgaggctctg gccaggggtga agaaggcggg gaagaagtgg
2400

agggacaagc gggataagct ggagtgggag accaacgtgg tctacaagga ggcgaaggag
2460

agcgtggacg ctctcttcgt cgattcccag tacagccgcc tgcaggccga tacgaacatt
2520

gcatgatcc acgcggctga caagaggggtg catcgcatca gggaggctta cctcccggag
2580

ctgaccgtga ttcccggcgt caacgcctcg atcttcgagg agctggaggg caggatcttc
2640

acggcctact ctctgtacgg cgccaggaat gtcacaaaga acggcgactt caataacggg
2700

ctctcctgct ggaacgtgaa gggccacgtg gaggtccagc agatccacca taggtctgtc
2760

ctggttgtgc cctcatggaa gacagaggtt tcccaggagg tttgcgtgtg cccaggcagg
2820

ggatacatcc tcagggtgac tgcctacaag gagggctacg gggagggcaa tgtcacaatt
2880

cacgagatcg agaataacac tgacgagctg aagttccgca attgcgagga ggaggaggtc
2940

taccgaata aactgttac ctgcaatgac tacaccgtca accaggagga gtacaagggc
3000

acgtgcacat cccggaaccg cgggtacgac gagagctacg agtcttcac cagcgagagc
3060

gcatactacg cctcgggtga cgaggagaag ggctacacgg atgggcggcg cgagaatctc
3120

tgcgagttca acagggggta cggcgactac accagcatcc ctacagccta cgtgactaag
3180

gagctggagt acttcccgga gacagataaa gtgtggatcg agcttggcga gactgagggg
3240

gccttcatcc tggactccgt ggagctgctc ctgatggagg agtaa
3285

<210> 26
 <211> 3509
 <212> ADN
 <213> Trình tự Nhân tạo

 <220>
 <223> Trình tự mã hóa BT-0029 đột biến tổng hợp

 <400> 26
 atggagatca acaatcagaa ccagtgcgtc ccgtacaatt gcctcaacaa tcccgagtc
 60
 gagatcctga acgtggccat tttctccagc gagcaggtcg cggagattca cctcaagatc
 120
 acccgctga ttctcgagaa cttcctgccg ggcgggtctt tcgctttcgg cctgttcgac
 180
 ctcatattggg ggatcttcaa tgaggatcag tggtcagcgt tcctcaggca ggtggaggag
 240
 ctgatcaacc agaggattac ggagttcgct aggggccagg ctatccagag gctggtgggc
 300
 ttcgggaggt cctacgacga gtacattctg gctctcaagg agtgggagaa cgacccggat
 360
 aatcccgcca gcaaggagag ggttcggaca cgcttcagga ccacggacga tgccctcctg
 420
 actggcgtgc cactcatggc tatccctggg ttcgagctgg ctaccctctc ggtctacgct
 480
 cagtcggcca acctgcacct cgccctcctg agggacgctg tgttcttcgg ggagaggtgg
 540
 ggcctcacc agacgaacat caatgacctc tactcgcggc tgaagaactc tattcgcat
 600
 tacaccaatc attgcgttag gttctacaac atcggcctgg ggaacctcaa tgtgattcgg
 660
 ccagagtact accggttcca gcgcgagctg acgatcagcg ttctggacct cgtggcgtg
 720
 ttccctaact acgatatccg gacttaccg attcccacca agtcccagct cacgcgcgag
 780
 atctacacag acccaatcat ttcgccgggg gcccaggccg gctacacact ccaggacgtc
 840
 ctgagggagc cgcacctgat ggatttcctg aaccggtca tcatctacac tggggagtac
 900
 cggggcatcc ggcactgggc tggccatgag gtggagtcgt ctgcactgg catgatgacc
 960

aacatccgct tccccctcta cggcacggcg gccacagccg agccaacacg cttcattact
1020

ccgtcgacct tccccggcct gaatctcttc taccggactc tgtcggcccc catcttccgc
1080

gacgagccgg gcgctaacat cattatccgc taccgcacct ccctggtgga gggcggtggg
1140

ttcatccagc cgaacaatgg cgagcagctg tacaggggtgc ggggcaccct cgactccctg
1200

gatcagctgc ccctcgaggg cgagtcatcc ctcacggagt actcacacag gctgtgccat
1260

gtccggttcg ctcagtccct ccggaacgcg gagccactgg actacgcccg ggtgcctatg
1320

ttctcatgga cacaccgctc cgccacacca actaatacca tcgaccctga tgttattact
1380

cagatcccac tcgtgaaggc gttcaacctg cattcggggc ctactatcgt caagggccct
1440

gggttcaccg gcgggggacat tctgcgcagg acgaatgttg ggtctttcgg cgatatgcgc
1500

gtgaacatca cggcgcccct cagccagcgc tacaggggtgc ggattcgcta cgcttcgaca
1560

actgacctgc agttctacac caacatcaat gggaccacga ttaacatcgg caatttcagc
1620

tcgacgatgg attccgggga cgatctccag tacggcaggt tccgggtggc ggggttcaca
1680

actccgttca cgttcagcga cgctatgtcg acgttcacaa tcggcgcctt cagcttctct
1740

tcaaacaatg aggtgtacat tgaccgcacg gagttcgtcc ccgctgaggt taccttcgag
1800

gccgagtacg atctggagaa ggctcagaag gccgtcaatg cgctcttcac gtccagcaac
1860

cagatcggcc tgaagacgga cgtgacagat taccacattg acaaggtagag caacctcgtc
1920

gagtgcctgt cggacgagtt ctgcctcgat gagaagaggg agctgtccga gaaggtaag
1980

catgcgaagc gcctctgcga cgagaggaac ctccctccagg acccgaactt ccgcggaatc
2040

aaccggcagc ctgaccgcgg gtggaggggc tccactgata ttaccatoca gggcggggac
2100

gatgtgttca aggagaacta cgtcacgctc ccgggcacat tcgacgagtg ctaccccaca
2160

tacctgtacc agaagatcga tgagtcaaag ctcaaggcct aactcggta cgagctgcgc
2220

ggatacatcg aggactccca ggatctcgag atctacctca tccgctacaa tgcgaagcac
2280

gagactgtta acgtgccagg caccgggtcg ctgtggcccc tctcggcgca gtccccaatc
2340

gggaagtgcg gcgagccaaa tcgctgcgcg acccatctcg agtgggaacc tgacctggat
2400

tgcagctgca gggacggcga gaagtgcgct caccattctc accatttctc actggacatc
2460

gatgtgggct gcaccgacct caacgaggat ctgggcgtct gggttatttt caaaatcaag
2520

accaggacg ggcacgctag gctgggcaac ctcgagttcc tcgaggagaa gccgctggtg
2580

ggcgaggctc tggctagggt taagaggcg gagaagaagt ggagggacaa gcgggagaag
2640

ctggagctgg agacaaacat cgtctacaag gaggcgaaga agtctgtcga tgctctcttc
2700

gttaattcac agtacgacag gctgcaggcc gatacgaaca tcgcgattat ccacgctgcc
2760

gacaagcggg tgcattccat tcgcgaggcc tacctcccag agctgagcgt tatccctggc
2820

gtgaacgcgg ctattttcga ggagctggag ggccgcatct tcaccgcta cagcctgtac
2880

gacgcgagga acgttattaa gaatggggat ttcaacaatg gcctctcgtg ctggaacgtg
2940

aagggccacg tggacgtcga ggagcagaac aatcatcgct ctgtcctcgt ggtcccgga
3000

tgggaggctg aggtttcaca ggaggtagg gtgtgcccgg gcaggggata catcctccgc
3060

gtcacggcgt acaaggaggg ctacggggag ggctgcgtta caattcacga gatcgaggac
3120

aatactgatg agctgaagtt ctccaactgc gtggaggagg agatctaccc aaacaatact
3180

gtcacctgca acgactacac ggccacacag gaggagtacg agggcactta caccagccgc
3240

aatagggggt acgacggcgc ctacgagtcg aactcgtctg tgcctgcgga ttacgcgtct
3300

gcttacgagg agaaggctta caccgacggc aggcgcgata acacatgcga gagcaatagg
3360

gggtacggcg actacacccc actccctgcg ggctacgtga cgaaggagct ggagtacttc
3420

ccggagaccg acaaggtctg gcttgagatc ggggagacgg agggcacatt catcgtggat
3480

tccgtcgagc tgctgatcat ggaggagta
3509

<210> 27

<211> 3510

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0030 đột biến tổng hợp

<400> 27

atggaggtca accaccagaa tgagtgcgtt ccatacaact gcctgaagaa tcctaagatc
60

gagatgctcg acatcgaggg catttcgtcc aggtcgaggg agcaggtggc cgagatctcc
120

ctgggcctca caaggttcct cctggagtct ctctgcgcg gcgcttcatt cggcttcggg
180

ctgttcgata tcatttgggg cgtgatcggc ccggaccagt ggtcgctggt cctcaccag
240

attgagcagc tcatcgacca gcgcattgag gcgcacgtga ggaaccaggc tatttcgagg
300

ctggagggcc tgggggactc ctacgaggtc tacatcgaga gcctgagggg gtgggagggt
360

tccccgaaca atgagtctct ccagcaggat gtgaggaacc ggttcagcaa caccgacaat
420

gctctgatta cggccatccc gattctccgc gagcagggtc tcgagatccc gtcctgacc
480

gtctacgtgc aggtgcgcaa cctgcacctc agcctcctgc gggatgcggt gtacttcggg
540

cagaggtggg gcctcgacac ggccacggtc aacaatcact acaaccggct gatcaacctc
600

attaatacct actccgatca ttgcgctcag tggttcaacc gcgggctgga caatttcggc
660

gtggtcactg cccgctacct cgatttccag agggaggtga ccattctcagt cctggatatt
 720
 gttgcgtctt tcccaaacta cgacatcagg acatacccta ttcagactct gtcccagctc
 780
 acccgggaga tctacacgtc tccagtcgcg gagcctggcg cttcactcaa cgttgacctg
 840
 aggaatatcc tccgggagcc acacctgatg gatttcttga cagcctcgt gatctacact
 900
 ggcgtccagg gcgggatcta ccaactgggct gggcatgaga tctcgtctag gaccacgggc
 960
 aacctgtcat ccaatatcca gttcccactg tacggcacgt ccgccaacgc tgaccggcct
 1020
 ttcaatctgg cgatccatta ctccgagacc atctaccgca cgctcagcgc tccaatctac
 1080
 tccgtcagcg gcgggatttc gccgaacagg accagggccg ttgagggcgt gcgcttcctc
 1140
 acagctaggg acaacaatct gaactcgctc ccgttctctgt accggaagga gggctcactg
 1200
 gattccttca ccgagctgcc gccggaggac gagaacgagc caccttacat cggctactcc
 1260
 cacaggctgt gccatgccag gttcgtcgc agctcggttg tgctcgagcc gtctaacttc
 1320
 gctcgcatcc cgggtgttctc atggacccac cgctccgccg gcccgacaaa tgaggtctct
 1380
 tcatccagga tcaactcagat tccttgggtt aaggctcata ccctggatag cggcgccttc
 1440
 gtgattaagg gcccgggggt caccggcggg gacatcctca ccaggccaaa cctcgggaca
 1500
 ctgggcgcgc tcagggtcac actgactggc cagctcccgc agacttaca catccggatt
 1560
 cgctacgcct ccattgcgaa tcgcggcggc accctcatct tctctcagcc gccatcatac
 1620
 ggccctcacat tcccgaagac tatggacatc gatgagcccc tgacgtcgcg ctctttcgcc
 1680
 aggacaactc tcttcacacc aattaccttc acgcaggccc aggcggagct gaacctcact
 1740
 atccagcagg gcgtgtacat cgataggatt gagttcatcc ctgtcaatgc caggttcgag
 1800

gcggagtacg acctggagcg ggctcagaag gccgtgaacg cgctcttcac cagctcgaat
 1860
 cagctggggc tcaagacaga cctgactgat taccacattg atcaggtgtc gaacctggtc
 1920
 gactgcctct ctgatgagtt ctgcatcgac gagaagcgcg agctgagcga gaaggtcaag
 1980
 catgccaagc ggctctcaga cgagcgcaac ctctgcagg attccaactt caggggcatc
 2040
 aataggcagc cggacagggg gtggaggggc tccaccgata tcacgattca gggcgggaac
 2100
 gacgttttca aggagaatta cgtgacctg ccaggcacgt tcgacgagtg ctaccctacg
 2160
 tacctctacc agaagatcga tgagtccaag ctgaaggcgt acacacgcta ccagctccgg
 2220
 ggatacatcg aggacagcca ggatctggag atctacctca tcaggtacaa cgccaagcac
 2280
 gagactgtta atgtgccggg caccgggagc ctgtggccac tcagcgtgga gtcgcctatc
 2340
 gggaagtgcg gcgagcccaa ccgctgcgtc ccacagctgg agtggaacag caatctcgac
 2400
 tgctcctgca gggatggcga gaagtgcgcc caccattccc accatttcag cctcgacatc
 2460
 gatgtggggt gcacggacct gcacgacgat ctggcgctct gggttatttt caagatcaag
 2520
 acgcaggacg ggcattgctag gctgggcaac ctcgagttcc tcgaggagaa gccgctgggt
 2580
 ggcgaggctc tggctagggg caagagggcg gagaagaagt ggcgggacaa gcgcgagaca
 2640
 ctgcagctcg agactaacat cgtttacaag gaggccaagg agtcggtgga cgctctgttc
 2700
 gccaaactctc agtacaatcg cctccaggct gatacgaaca ttgccatgat ccacgctgcc
 2760
 gacaagaggg tgcattcgcat ccgcgaggct tacctccccg agctgtctgt tatccccggg
 2820
 gtgaacgcgg gcattttcga ggagctggag ggccggattt tcaccgcctt cagcctctac
 2880
 gatgcgcgca acgtcatcaa gaattcagac ttcaacaatg ggctgtcctg ctggaacgtc
 2940

aagggccacg ttgacatcga ggagcagaac aatcatagga gcgttctcgt cgttccggag
3000

tgggagggcg aggtgtcgca gaaggtgcac gtctgcccgg gcaggggata catcctgcgg
3060

gtcaccgcct acaaggaggg ctacggggag ggctgcgtta ctattcacga gatcgaggac
3120

cataccgatg agctgaagtt ccgcaactgc gaggaggatg aggtgtaccc gaacaatata
3180

aggacttgca atgcgtaccc ggctgaccag gagggggtacg agggcgccctg cacttctcgc
3240

aacaggggggt acgacgaggt ctacggcaat accccgtcac tccccgccga ttacgcgccca
3300

atctacgagg agaacgccta cacggacggg aggaggggca atccttgcca gtcttcacgc
3360

gggtacggcg actacacgcc actgcctgcc ggctacgaga caaaggagct ggagtacttc
3420

ccggagaccg atacggtgtg gatcaagctt ggggagacag agggcacttt catcgtcgac
3480

tccgttgagc tgctgatcat ggaggagtaa
3510

<210> 28

<211> 2160

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0031 đột biến tổng hợp

<400> 28

atgaagctca agaatccaga caagcaccag agcctgtcca gcaacgctaa ggtggacaag
60

atgccacag attcgctcaa gaatgagact gacattgagc tgaagaacat gaacaatgag
120

gattacctcc gcatgtcgga gcatgagtct atcgaccctt tcgtttcggc ctctactatt
180

cagaccggca tcgggattgc cggcaagatc ctcggcaccc tcggcgtgcc gttcgctggc
240

cagatcgctt ctctctactc attcattctg ggcgagctgt ggcctaaggg gaagagccag
300

tgggagatct tcatggagca cgtggaggag atcattaatc agaagattct gacctacggc
360

aggaacaagg ctctgtcgga cctcaggggc ctgggggatg ccctcgctgt ctaccatgag
 420
 tccttgagga gctgggttaa gaatcgcaac aatacgcggg ctcgctccgt ggtcaagaac
 480
 cagtacatcg ccctcgagct gatgttcgtc cagaagctgc catcgttcgc cgtgtccggc
 540
 gaggaggtcc cgctcctgcc aatctacgcc caggetgcca acctccacct cctgctcctg
 600
 cgcgacgctt ctatcttcgg caaggagtgg gggctgtcag cgtccgagat tagcacattc
 660
 tacaacaggc aggttgagag gactcgggac tactcggatc attgcgtgaa gtggtacaac
 720
 accggcctca acaatctgag gggaaccaac gccaaagact gggtcgggta caaccagtcc
 780
 cgcaaggaca tgacactcat ggtgctcgat ctggctcggc tcttcccatc atacgacact
 840
 ctggtctacc caatcaagac cacgtcccag ctgactaggg aggtttacac cgatgccatt
 900
 ggcacggttc acccctcagg ggcggtggcc tccacaactt ggtacaacaa taacgcgcgc
 960
 agcttctcga ccatcgaggc tgccgttggt cggaaccccc acctcctgga cttcctcgag
 1020
 caggtgacca tctactcgct cctgtctagg tggtaaata cgcagtacat gaacatgtgg
 1080
 ggcgggcata agctcgagtt ccggactatc ggcggcacc tcaacacaag cactcagggc
 1140
 tcgaccaata cgtctattaa ccagtgacg ctccctttca catctaggga cgtctaccgg
 1200
 acagagtcac tggccggcct caatctgttc ctacgcagc ccgtgaacgg cgtccccagg
 1260
 gttgacttcc actggaagtt cgtcacccat ccaatcgct cggataactt ctactaccct
 1320
 ggctacgctg gcattgggac gcagctccag gactccgaga atgagctgcc gcccgagacc
 1380
 accggccagc ccaactacga gtcttactca cacaggctct cccatatcgg cctgatttcc
 1440
 gccagccacg ttaaggcgct cgtgtacagc tggacacatc gtcggccga caggacaaat
 1500

actatcaact cggattctat cacgcagatt ccgctcgtga aggccttcaa cctgccgtcg
1560

ggcgccctcg tcgttcgggg ccccggttc accggcgggg acatcctgcg caggaccaac
1620

acgggcacat tcggggatat ccgcgtcaat attaaccac ctttcgcca gcgctacagg
1680

ctccgcatcc gctacgcgtc cacaactaat ctggagttcc acaccagcat caacggcaag
1740

gctattaatc aggggaactt ctgggtacc atgaaccgcg gcgaggacct cgattacaag
1800

gccttcagga cgggtgggctt caccacgccc ttctcattct ccaacgcgca gtcgactttc
1860

accatcggcg cttggaattt ctctctgggg aacgaggtct acatcgacag gattgagttc
1920

gtgccggtcg aggttaccta cgaggcggag cacgatttcg agaaggcgca ggaggagggtg
1980

acggctctct tcacgagcac aaatccaggc ggctgaaga cgaacgtcac agagtaccat
2040

atcgaccagg tctcgaatct cgttgagagc ctgtcgaacg agttctacct cgatgagaag
2100

cgggagctgt tcgagctcgt gaagtacgcc aagcagctgc accttggccg caacatgtaa
2160

<210> 29

<211> 3495

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Trình tự mã hóa BT-0201 đột biến tổng hợp

<400> 29

atggagcgca acaatcagga ccagtgcac ccgtacaact gcctcaacaa tcccagaggtt
60

ggcattctgg atatcgagaa cttcaatctg gagctgggtg cccaggtgtc agtcggcctc
120

accaggttcc tcctggaggc ctccatccca ggcgcgggt tcgctctggg cctcttcgac
180

atcatttggg gcgccctcgg cgtggatcag tggtcctgt tcctcgctca gattgagcag
240

ctgatcaatg agcgcattac caccgtggag aggaaccgcg ctatccaggc cctgtcgggc
300

ctctccagct cgtacgaggt ctacattgag gctctcaggg agtgggagaa cgacctggat
360
aatccggctt ctagggacag ggtggtggct aggttccggg ctaccgacaa ctactcatt
420
acggatatcc cgctcctgga gatccccggc ttcgagattg ccacactgtc cgtgtacacg
480
cagggggcca acctgcacct cgccctcctg agggacgctg tctacttcgg ggagcgcctgg
540
ggcctgacac agactaacat cgaggacctc cacacacggc tgactcgcta cattcaggag
600
tactccgatac attgcgcgcg ctggtacaac cagggcctca acaatattgg cggaatcaac
660
acccgctacc tcgacttcca gcgggagctg acgatcagcg tcctcgacat tgttgccctg
720
ttcccgaact acgatatccg gacctacca attcctacgc agagccagct cacacgcgag
780
atctacactt cgccggttgt ggccccggc gtgaattgga ttctgtcgat ctctaacgtc
840
ctcagggcgc cacacctgat ggacttcttc gatcggatca ttatctacac gggcacggtg
900
cgctcaaccc ctactggga ggggcatgag gtcacttcca ggaggacggg ccaggggaac
960
gagattagga gcccgtctta cggcgtggct gctaatgccg agccgccagt tacgatcagg
1020
ccaacgggct tcacggacga gcagcggcag gtctaccgcg ttctgtccag ggtcgccagc
1080
ttocggaact cgggcacaaa tttctcactc gttgacgcgg cttccttcct gactatcttc
1140
agcgcgtctt caatctaccg caacggcttc gggttcaatg ctgacaccat cgatgagatt
1200
ccaatcgagg ggacggaccc atacatcggc tactcccacc gcctctgcca tgtgggcttc
1260
accgctcca gcccgttcac cagccagtac gcgagggctc ccgtcttctc ctggacacac
1320
cggagcgtc cattcactaa caccatcgac ccagagcgca ttacgcagat ccctatggtg
1380
aaggcctaca acctgcatgc gggcgccaca gtcgttaggg gcccggggtt cactggcggg
1440

gacctcctgc ggcgcacgaa cacaggcact ttgcgggata tcaggggtcaa tattaccggc
1500

ccccctctccc agcgcctaccg cgtgcgcacg cgctacgctt ctacaactga cctgcagttc
1560

ttcaccagga tcaacgggac gagcgtgaac cagggcaatt tccagcggac gatgaaccgc
1620

ggggacaatc tcgagtctgg caacttccgg acagccgggt tctcaactcc attctcattc
1680

tccaatgctc agtccacctt cacgctgggc acccaggcct tcagcaacca ggaggtctac
1740

attgacagga tcgagttcgt gcctgctgag gtcacgttcg aggccgagtc cgatctcgag
1800

cgggctcaga aggccgtgaa cgcgctgttc acaagcactt cgcagctggg cctcaagaca
1860

aacgtcactg gctaccacat cgaccagggt tcgaatctcg tggcgtgcct gtctgacgag
1920

ttctgcctcg atgagaagcg cgagctgtcg gagaagggtga agcacgcaa gcggctctct
1980

gacaagcgca acctcctgca ggaccggaac ttcaggggaa tcaacaggca gcctgaccac
2040

ggctggaggg gctcgaccga tattacgacg cagggcgggg acgatgtttt caaggagaa
2100

tacgtgaccc tcccaggcac gtctgacgag tgctacccta catacctgta ccagaagatc
2160

gatgagtcca agctcaaggc gtacactcgc taccagctga ggggatacat cgaggactca
2220

caggatctcg agatctacct catccgctac aattccaagc acgagatcgt taacgtgccc
2280

ggcaccgggt cgctgtggcc actctctgtg gagaatcaga tcgggccatg cggcgagcct
2340

aaccggtgcg ccccgcacct cgagtggaaat ccagacctgc attgctcctg ccgggatggc
2400

gagaagtgcg tccaccattc tcaccatttc tcaactcgaca tcgatgttgg ctgcaccgac
2460

ctcaacgagg atctgggggt gtggctcatt ttcaagatca agaccagga cgggcacgcc
2520

aggctgggca acctcgagtt cctcgaggag gagccccctc tgggcgaggc tctggctagg
2580

gtgaagaggg ccgagaagaa gtggcgcgac aagagggaga agctgcagct cgagaccaac
2640

atcgtctaca aggagggcga ggagtcagtg gatgctctct tcgtcaactc ccagtacgac
2700

cgctgcagg ccgatacgaa tattgcgatg atccacgccg cggacaagag ggtgcatcgc
2760

atccgcgagg cctacctccc agagctgtcg gtcacccctg gcgttaacgc tgccattttc
2820

gaggagctgg agggccgcat cttcaccgcc tactctctgt acgacgcgag gaacgtgatt
2880

aagaacggga atttcaacaa tggcctcctg tgctggaacg tcaagggcca cgtcgatggt
2940

gaggagcaga acaatcatcg cagcgtgctc gtcacccccg agtgggaggg cgaggtctcg
3000

cagaaggtta ggggtgtgccc gggcagggga tacatcctgc gcgtgaccgc ttacaaggag
3060

ggctacgggg agggctgcgt caccattcac gagatcgagg acaatacgga tgagctgaag
3120

ttctccaact gcgttgagga gggctacccg aacaataccg tgacgtgcaa cgagtacacc
3180

atgaaccagg ggggtgggca gtgcactgac gcttgcaacg ttcgcaatag gggctacgag
3240

gatgcgtacg ggcacaaccc aagcaccctt gtgcattaca ccacgccata cgaggaggag
3300

acatacactg acgagagggc ggagaaccct tgcgaggcca ataagggcta cgtgaactac
3360

acaccgctcc ccgtgggcta cgtgactaag gagctggagt acttccccga gaccgacacg
3420

gtctggattg agatcgggga gacagagggc actttcctcg ttgatagcgt ggagctgctg
3480

atcatggagg agtaa
3495

<210> 30
<211> 3501
<212> ADN
<213> Trình tự Nhân tạo

<220>
<223> Trình tự mã hóa BT-0202 đột biến tổng hợp
<400> 30

atgaagatca acaatcagaa ccagtgcatt ccgtacaact gcctgagcaa tccggaggag
 60
 gtgtcctcgtg acggcgagcg cattctgccg gacatcgatc ccctcgaggt cagcatgtcg
 120
 ctctgcagt tcctcctgaa caatttcgtg ccgggcgggg gcgtcatctc gggcctgatc
 180
 aacaagattt ggggcgcctt ccggccgtcc gagtgggagc tgttcctcgc ccagattgag
 240
 cagctgatcg acaggaggat tgaggcgggc gtcagggcta aggccatcgc tgagctggag
 300
 gggctgggcc ggtcttacca gctctacggc gaggcgttca aggagtggga gaagacacc
 360
 gataacaccg cggcgcgctc cagggtgacc gagcgcttca ggatcattga cgcgcagatc
 420
 gaggctaaca ttccatcttt ccgcgtttca ggcttcgagg tgcctctcct gtctgtctac
 480
 acccaggcgg ctaacctgca tctcgccctc ctgcgggatt cagtgggtgtt cggcgagagg
 540
 tgggggctga ccaccacaaa cgtgaatgac atctacaacc gccagggtcaa gaggattgac
 600
 gagtacagcg atcactgcgt cgacacctac aagacggagc tggagcgcct cgagttctcc
 660
 agcatcgcg agtgaggat ctacaaccag ttccggaggg agctgaccct caccgtcctc
 720
 gatatcgttg ccctgttccc gaattacgac ggccgcctgt acccatcag gaccatttcc
 780
 cagctcacgc gggacatcta cacatctcca gtttcagagt tctactacgg ccctatctac
 840
 aactacaata ttgtggggag gctcactgag cagcagctga ggcggccaca tctcatggat
 900
 ttcttcaact ccatgatcat gtacaccagc gacaatcgca gggagcacta ctggtcggc
 960
 ctggagatga aggccacaga cactagcggg aaccaggtgt cgttccccct ggctggcacg
 1020
 agggggaatt ccgccccgcc agtcacagtt aggaacaatg gggagggcgt ctaccgcac
 1080
 ctctctgagc cattctactc gtctcctttc ctgggcacct ccgtgctggg ctcccgcggc
 1140

gaggagtctcg ccttcgcgag caacactacc acgtccctgc cgagcactat ctacaggaat
1200

cgcgggcaccg tggattcgct cgtctctatt ccacctcagg actactcggg cccgccccac
1260

aggggctact cccacctcct gtctcatggt acgatgcaca actcatcccc gatcttccac
1320

tggacacatc gctccgccac cccacgcaat acgattgacc ctgatagcat cacgcagatt
1380

cccgttgtga aggcctcaca tctgtccggg ggcagcgta tcaaggggcc gggccacaca
1440

ggggggcgacc tgatttccct ccccgtaac aatttactc acttccgcat cccgttccag
1500

gccaacaccc cgcagcgcta ccgcatccgc atccgctacg ctgccgactc ggatggcacc
1560

ctggattcgg ggggtgttct ctcagctgcg gccggggacg gcttcaacac aacttcgtac
1620

agggctacca tgtctccggg cggctccctc acatctaggg atttccagtt cctggacctc
1680

aacacatcat tcacttccga cgtcgcgtcc aatctgtggc tccatttcat caggtacatt
1740

cgggcaggca acctgtacat cgatagggcc gagttcattc ctgtggacgc caccttcgag
1800

gcgggctaca atctcgagcg ggcgcagaag gctgtgaacg ccctgttcac ctccacgaat
1860

cagaagggcc tccagaccga cgtcacggat taccacatcg atcaggtttc aaacctcgtg
1920

gattgcctgt ccgacgagtt ctgcctggac gagaagcgcg agctgagcga gaaggtaag
1980

caggcgaaga ggctgtctga tgagcggaac ctctgcagg actcaaactt ccggggaatc
2040

aacagggagc aggatagggg ctggaggggg tcgacacata tcactattca gggcggcaac
2100

gacgtcttca aggagaatth cgttactctg ccggggcgct tcgatgcttg ctacccacc
2160

tacctctacc agaagatcga cgagtccaag ctgaaggcct acacccgcta cgagctgcgg
2220

ggatacatcg aggacagcca ggacctgat atctacctga ttcgctacaa cacaagcac
2280

gagactctca atgtgccggg cactaagtcc ccctgggtccc tgtgcacgga gtccccactc
2340

ggcaagtgcg gggagcctaa caggtgcgcc tcccagatcg agtggaatcc ggacctggat
2400

tgctcctgca gggacggcga gaagtgcgct caccattcac accatttctc cctcgacatc
2460

gatgtgggct gcacggacct caacgagaat ctggggatct gggtcatttt caagatcaag
2520

acgcaggacg gccacgccag gctggggaac ctcgagttcc tcgaggagaa gccgctcctg
2580

ggcgaggctc tggctagggc caagcgcgcc gagaagaagt ggagggacaa gcggggagaag
2640

ctccagagcg agacgaacat cgtgtacaag gaggccaagg aggcggttga tggcctgttc
2700

gtggactcgc agtacgagcg cctccaggcg gataccaaca ttgctatgat ccacgctgcc
2760

gacaagcacg tgcattcagat cagggaggtc tacttcccag agctgtccgt catccctggc
2820

gttaacgcgg ctatttttoga ggagctggag ggccggatct tcacggccta ctactctac
2880

gatgcgcgca acgtcattaa gaatggcgac ttcaacaatg ggctgtcctg ctggaacgtt
2940

aagggccatg tggacgtoga ggttcagaac aatcacgcgt cagtgtcctg catcgtgag
3000

tgggaggcgg aggtgtccca ggaggttccg gtgtgcccg ggaggggata catcctgagg
3060

gttaccgcct acaaggaggg gtacggcgag ggggtgcgtga cgattcatga gatcgaggac
3120

cacacagatg agctgaagtt caggaactgc gaggaggagg aggtttaccc gaacaataca
3180

gtgacttgca atggctacac cgcgacgcag gaggagtaca aggatgctta cacctcgagg
3240

aacaggggct acgacgaggg ctacgggaac aatccaagcg tccttgctga ttacgcctcg
3300

gtttacgagg agaaggctta cacggacggc aggagggaga acccatgcga gatggagagg
3360

ggctacactc cactgcctgt gggatacatc accaaggagc tggagtactt ccctgagaca
3420

gataactgtct ggatcgagct tggcgagacc gaggggacgt tcctcattga cagcgtggag
3480

ctgctgctca tggaggagta a
3501

<210> 31
<211> 1181
<212> PRT
<213> Bacillus thuringiensis

<400> 31

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Val Arg Leu Asn Ser Ser Gly Asn Asn Ile Gln Asn Arg Gly Tyr
 500 505 510

Ile Glu Val Pro Ile His Phe Pro Ser Thr Ser Thr Arg Tyr Arg Val
 515 520 525

Arg Val Arg Tyr Ala Ser Val Thr Pro Ile His Leu Asn Val Asn Trp
 530 535 540

Gly Asn Ser Ser Ile Phe Ser Asn Thr Val Pro Ala Thr Ala Thr Ser
 545 550 555 560

Leu Asp Asn Leu Gln Ser Ser Asp Phe Gly Tyr Phe Glu Ser Ala Asn
 565 570 575

Ala Phe Thr Ser Ser Leu Gly Asn Ile Val Gly Val Arg Asn Phe Ser
 580 585 590

Gly Thr Ala Gly Val Ile Ile Asp Arg Phe Glu Phe Ile Pro Val Thr
 595 600 605

Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
 625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Cys Leu Ser
 645 650 655
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
 675 680 685
 Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
 690 695 700
 Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720
 Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
 740 745 750
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Val Tyr Leu Ile Arg Tyr
 755 760 765
 Asn Ala Lys His Glu Thr Leu Asn Val Pro Gly Thr Gly Ser Leu Trp
 770 775 780
 Pro Leu Ala Val Lys Ser Pro Ile Gly Arg Cys Gly Glu Pro Asn Arg
 785 790 795 800
 Cys Ala Pro Arg Ile Glu Trp Lys Pro Asp Val Asp Cys Ser Cys Arg
 805 810 815
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
 820 825 830
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
 835 840 845
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Lys Ile Gly Asn Leu Glu
 850 855 860
 Phe Leu Glu Glu Lys Leu Leu Leu Gly Glu Ala Leu Ala Arg Val Lys
 865 870 875 880

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Val Glu Leu Leu Leu Met Glu Glu
 1175 1180

<210> 32

<211> 1176

<212> PRT

<213> Bacillus thuringiensis

<400> 32

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

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

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

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

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

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

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

Ala Phe Ala Glu Trp Glu Lys Ala Pro Asp Asp Pro Glu Leu Arg Glu
 115 120 125

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu

340	345	350
Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg		
355	360	365
Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp		
370	375	380
Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr		
385	390	400
Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro		
	405	410
		415
Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser		
	420	425
		430
His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg		
	435	440
		445
Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Glu Phe Asn Asn Ile		
450	455	460
Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr Asn		
465	470	475
		480
Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly Gly		
	485	490
		495
Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg Val		
	500	505
		510
Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr		
	515	520
		525
Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg Pro		
530	535	540
Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn Leu		
545	550	555
		560
Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn Phe		
	565	570
		575

Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn Ser
 580 585 590

Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val
 595 600 605

Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn
 610 615 620

Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr
 625 630 635 640

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp
 645 650 655

Glu Phe Cys Leu Asp Glu Lys Gln Glu Leu Ser Glu Lys Val Lys His
 660 665 670

Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe
 675 680 685

Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp
 690 695 700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
 705 710 715 720

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

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

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

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

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

Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp

				805						810							815
Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	Asp		
			820						825				830				
Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly	Val	Trp	Val	Ile	Phe		
		835					840					845					
Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu	Gly	Asn	Leu	Glu	Phe		
	850					855						860					
Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	Ile		
865					870					875					880		
Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp	Glu	Thr		
				885					890						895		
Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe	Val		
			900					905						910			
Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr	Asn	Ile	Ala	Met	Ile		
		915					920					925					
His	Ala	Ala	Asp	Lys	Arg	Val	His	Ser	Ile	Arg	Glu	Ala	Tyr	Leu	Pro		
	930					935					940						
Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala	Ile	Phe	Glu	Glu	Leu		
945					950					955					960		
Glu	Gly	Arg	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr	Asp	Ala	Arg	Asn	Val		
				965					970						975		
Ile	Lys	Asn	Gly	Asp	Phe	Asn	Asn	Gly	Leu	Ser	Cys	Trp	Asn	Val	Lys		
			980					985					990				
Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	Gln	Arg	Ser	Val	Leu	Val		
		995					1000						1005				
Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val	Cys			
	1010					1015						1020					
Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu	Gly			
	1025					1030						1035					

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

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

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

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

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

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

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

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

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

Met Glu Glu
1175

<210> 33

<211> 1176

<212> PRT

<213> *Bacillus thuringiensis*

<400> 33

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

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

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

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

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

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

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

Ala Phe Ala Glu Trp Glu Lys Ala Pro Asp Asp Pro Glu Leu Arg Glu
 115 120 125

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr Asn
 465 470 475 480

Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly Gly
 485 490 495

Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg Val
 500 505 510

Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr
 515 520 525

Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg Pro
 530 535 540

Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn Leu
 545 550 555 560

Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn Phe
 565 570 575

Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn Ser
 580 585 590

Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val
 595 600 605

Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn
 610 615 620

Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr
 625 630 635 640

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp
 645 650 655

Glu Phe Cys Leu Asp Glu Lys Gln Glu Leu Ser Glu Lys Val Lys His
 660 665 670

Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe
 675 680 685

Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp
 690 695 700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
 705 710 715 720

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

Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly
 740 745 750
 Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn
 755 760 765
 Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro
 770 775 780
 Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys
 785 790 795 800
 Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp
 805 810 815
 Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
 820 825 830
 Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
 835 840 845
 Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
 850 855 860
 Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Ile
 865 870 875 880
 Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
 885 890 895
 Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
 900 905 910
 Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
 915 920 925
 His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
 930 935 940
 Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
 945 950 955 960
 Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
 965 970 975

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

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

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

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

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

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

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

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

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

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

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

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

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

Met Glu Glu
 1175

<210> 34
 <211> 715
 <212> PRT
 <213> *Bacillus thuringiensis*

<400> 34

Met Lys Ser Lys Asn Gln Asn Met His Gln Ser Leu Ser Asn Asn Ala
 1 5 10 15

Thr Val Asp Lys Asn Phe Thr Gly Ser Leu Glu Asn Asn Thr Asn Thr
 20 25 30

Glu Leu Gln Asn Phe Asn His Glu Gly Ile Glu Pro Phe Val Ser Val
 35 40 45

Ser Thr Ile Gln Thr Gly Ile Gly Ile Ala Gly Lys Ile Leu Gly Asn
 50 55 60

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

Leu Gly Glu Leu Trp Pro Lys Gly Lys Ser Gln Trp Glu Ile Phe Met
 85 90 95

Glu His Val Glu Glu Leu Ile Asn Gln Lys Ile Ser Thr Tyr Ala Arg
 100 105 110

Asn Lys Ala Leu Ala Asp Leu Lys Gly Leu Gly Asp Ala Leu Ala Val
 115 120 125

Tyr His Glu Ser Leu Glu Ser Trp Ile Lys Asn Arg Asn Asn Thr Arg
 130 135 140

Thr Arg Ser Val Val Lys Ser Gln Tyr Ile Thr Leu Glu Leu Met Phe
 145 150 155 160

Val Gln Ser Leu Pro Ser Phe Ala Val Ser Gly Glu Glu Val Pro Leu
 165 170 175

Leu Pro Ile Tyr Ala Gln Ala Ala Asn Leu His Leu Leu Leu Leu Arg
 180 185 190

Asp Ala Ser Ile Phe Gly Lys Glu Trp Gly Leu Ser Asp Ser Glu Ile
 195 200 205

Ser Thr Phe Tyr Asn Arg Gln Val Glu Arg Thr Ser Asp Tyr Ser Asp
 210 215 220

His Cys Thr Lys Trp Phe Asp Thr Gly Leu Asn Arg Leu Lys Gly Ser
 225 230 235 240

Asn Ala Glu Ile Trp Val Lys Tyr Asn Gln Phe Arg Arg Asp Met Thr
 245 250 255

Leu Met Val Leu Asp Leu Val Ala Leu Phe Gln Ser Tyr Asp Thr His
 260 265 270

Met Tyr Pro Ile Lys Thr Thr Ala Gln Leu Thr Arg Glu Val Tyr Thr
 275 280 285

Asn Ala Ile Gly Thr Val His Pro His Pro Ser Phe Thr Ser Thr Thr
 290 295 300

Trp Tyr Asn Asn Asn Ala Pro Ser Phe Ser Ala Ile Glu Ala Ala Val
 305 310 315 320

Ile Arg Ser Pro His Leu Leu Asp Phe Leu Glu Gln Val Thr Ile Tyr
 325 330 335

Ser Leu Leu Ser Arg Trp Ser Asn Thr Gln Tyr Met Asn Met Trp Gly
 340 345 350

Gly His Lys Leu Glu Phe Arg Thr Ile Gly Gly Thr Leu Asn Thr Ser
 355 360 365

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

Thr Ser Arg Asp Ile Tyr Arg Thr Glu Ser Leu Ala Gly Leu Asn Leu
 385 390 395 400

Phe Leu Thr Gln Pro Val Asn Gly Val Pro Arg Val Asp Phe His Trp
 405 410 415

Lys Phe Val Thr His Pro Ile Ala Ser Asp Asn Phe Tyr Tyr Pro Gly
 420 425 430

Tyr Ala Gly Ile Gly Thr Gln Leu Gln Asp Ser Glu Asn Glu Leu Pro

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

Thr Asn Thr Ile Ser Lys Ser Val Gln Ser Val Phe Arg Cys Arg Cys
 675 680 685

Ser Glu Arg Thr Arg Ile Tyr Arg Trp Gly Tyr Pro Ser Lys Lys Glu
 690 695 700

Tyr Trp Tyr Ile Trp Gly Tyr Thr Ser Lys Tyr
 705 710 715

<210> 35
 <211> 1094
 <212> PRT
 <213> *Bacillus thuringiensis*

<400> 35

Met Phe Leu Glu Gln Ile Glu Gln Leu Ile Asp Gln Arg Ile Glu Thr
 1 5 10 15

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

Tyr Asp Val Tyr Ile Glu Ala Leu Lys Glu Trp Glu Asn Asn Pro Asp
 35 40 45

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

Asp Ala Leu Ile Thr Gly Ile Pro Leu Leu Ala Ile Pro Asn Phe Glu
 65 70 75 80

Ile Ala Thr Leu Ser Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser
 85 90 95

Leu Leu Arg Asp Ala Val Phe Phe Gly Glu Arg Trp Gly Leu Thr Gln
 100 105 110

Ile Asn Val Asp Asp Leu Tyr Arg Arg Leu Thr Asn Asn Ile Arg Asn
 115 120 125

Tyr Ser Asp His Cys Ala Arg Trp Tyr Asn Glu Gly Leu Asp Asn Ile
 130 135 140

Ser Gly Leu Ser Arg Ser Ile Asn Phe Gln Arg Glu Val Thr Ile Ser
 145 150 155 160

Val Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ile Arg Thr Tyr
 165 170 175

Pro Ile Ser Thr Thr Ser Gln Leu Thr Arg Glu Ile Phe Thr Ser Pro
 180 185 190

Ile Val Val Pro Asn Asp Phe Ser Val Ala Tyr Glu Gly Val Arg Arg
 195 200 205

Ala Pro His Leu Phe Glu Phe Leu Glu Lys Leu Val Ile Tyr Thr Gly
 210 215 220

Asp Arg Ser Gly Ile Arg His Trp Ala Gly His Glu Ile Thr Ser Arg
 225 230 235 240

Arg Thr Asp Ser Tyr His Gly Ile Ile Arg Tyr Pro Leu Tyr Gly Thr
 245 250 255

Ala Ala Asn Ala Glu Ser Pro Tyr Thr Leu Ala Leu Gln Pro Ser Glu
 260 265 270

Ser Ile Tyr Arg Thr Leu Ser Glu Pro Ile Phe Ser Gln Thr Gly Gly
 275 280 285

Leu Ser Pro His Arg Arg Arg Val Val Glu Gly Val Glu Phe Ser Ile
 290 295 300

Val Asn Asn Asn Val Asn Pro Ser Ser Phe Val Tyr Arg Arg Lys Gly
 305 310 315 320

Ser Leu Asp Ser Phe Thr Glu Leu Pro Pro Glu Asp Glu Ser Val Pro
 325 330 335

Pro Tyr Ile Gly Tyr Ser His Gln Leu Cys His Val Gly Phe Gly Arg
 340 345 350

Thr Asn Val Ile Phe Glu Pro Ser Asn Phe Ala Arg Val Pro Val Phe
 355 360 365

Ser Trp Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Asp
 370 375 380

Arg Ile Thr Gln Ile Pro Ser Val Lys Ala Ser Ser Leu Arg Asn Ser
385 390 395 400

Thr Val Val Ser Gly Pro Gly Phe Thr Gly Gly Asp Ile Val Arg Met
405 410 415

Gly Ala Val His Gln Ile Tyr Ala Thr Asp Leu Ser Met Asn Val Arg
420 425 430

Pro Ser Val Ala Leu Ser Arg Tyr Leu Ile Arg Leu Arg Tyr Ala Cys
435 440 445

Arg Gly Ser Ser Asn Ile Val Ile His Gly Pro Ser Ile Arg Phe Val
450 455 460

Ser Leu Pro Ser Thr Met Ser Asn Asp Glu Pro Leu Thr Tyr Gln Ser
465 470 475 480

Phe Arg Tyr Ala Ser Ile Thr Thr Pro Ile Thr Arg Pro Ile Tyr Asn
485 490 495

Met Phe Asn Leu Ser Ile Ser Arg Ile Ser Gly Val Gln Asn Leu Phe
500 505 510

Ile Asp Arg Ile Glu Phe Ile Pro Val Asp Ala Asn Phe Glu Ala Glu
515 520 525

Arg Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser
530 535 540

Thr Asn Gln Arg Gly Leu Lys Ile Asp Val Thr Asp Tyr His Ile Asp
545 550 555 560

Gln Val Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp
565 570 575

Glu Lys Arg Glu Leu Ser Glu Lys Ser Lys His Ala Lys Arg Leu Ser
580 585 590

Asp Glu Arg Asn Leu Leu Gln Asp Leu Asn Phe Lys Asp Ile Asn Arg
595 600 605

Gln Pro Glu Arg Gly Trp Ser Gly Ser Thr Gly Ile Thr Ile Gln Gly
610 615 620

Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr Phe
625 630 635 640

Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys
645 650 655

Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser
660 665 670

Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr
675 680 685

Val Asn Val Pro Gly Ser Gly Ser Leu Trp Pro Leu Ser Val Glu Ser
690 695 700

Ser Val Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Ser Arg Met Glu
705 710 715 720

Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala
725 730 735

His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp
740 745 750

Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln
755 760 765

Asp Gly His Ala Lys Ile Gly Asn Leu Glu Phe Leu Lys Glu Lys Leu
770 775 780

Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Lys Ala Glu Lys Lys Trp
785 790 795 800

Arg Asp Lys Arg Asp Lys Leu Glu Trp Glu Thr Asn Val Val Tyr Lys
805 810 815

Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asp Ser Gln Tyr Ser
820 825 830

Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys
835 840 845

Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu Leu Thr Val Ile
 850 855 860

Pro Gly Val Asn Ala Ser Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe
 865 870 875 880

Thr Ala Tyr Ser Leu Tyr Gly Ala Arg Asn Val Ile Lys Asn Gly Asp
 885 890 895

Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Glu Val
 900 905 910

Gln Gln Ile His His Arg Ser Val Leu Val Val Pro Ser Trp Lys Thr
 915 920 925

Glu Val Ser Gln Glu Val Cys Val Cys Pro Gly Arg Gly Tyr Ile Leu
 930 935 940

Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Asn Val Thr Ile
 945 950 955 960

His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu
 965 970 975

Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr
 980 985 990

Val Asn Gln Glu Glu Tyr Lys Gly Thr Cys Thr Ser Arg Asn Arg Gly
 995 1000 1005

Tyr Asp Glu Ser Tyr Glu Ser Ser Ser Ser Glu Ser Ala Tyr Tyr
 1010 1015 1020

Ala Ser Val Tyr Glu Glu Lys Gly Tyr Thr Asp Gly Arg Arg Glu
 1025 1030 1035

Asn Leu Cys Glu Phe Asn Arg Gly Tyr Gly Asp Tyr Thr Ser Leu
 1040 1045 1050

Pro Thr Ala Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr
 1055 1060 1065

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Ala Phe Ile
 1070 1075 1080

Leu Asp Ser Val Glu Leu Leu Leu Met Glu Glu
 1085 1090

<210> 36
 <211> 1169
 <212> PRT
 <213> Bacillus thuringiensis

<400> 36

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

Asn Pro Glu Ser Glu Ile Leu Asn Val Ala Ile Phe Ser Ser Glu Gln
 20 25 30

Val Ala Glu Ile His Leu Lys Ile Thr Arg Leu Ile Leu Glu Asn Phe
 35 40 45

Leu Pro Gly Gly Ser Phe Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
 50 55 60

Ile Phe Asn Glu Asp Gln Trp Ser Ala Phe Leu Arg Gln Val Glu Glu
 65 70 75 80

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

Arg Leu Val Gly Phe Gly Arg Ser Tyr Asp Glu Tyr Ile Leu Ala Leu
 100 105 110

Lys Glu Trp Glu Asn Asp Pro Asp Asn Pro Ala Ser Lys Glu Arg Val
 115 120 125

Arg Thr Arg Phe Arg Thr Thr Asp Asp Ala Leu Leu Thr Gly Val Pro
 130 135 140

Leu Met Ala Ile Pro Gly Phe Glu Leu Ala Thr Leu Ser Val Tyr Ala
 145 150 155 160

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

Gly Glu Arg Trp Gly Leu Thr Gln Thr Asn Ile Asn Asp Leu Tyr Ser

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

Arg Leu Cys His Val Arg Phe Ala Gln Ser Leu Arg Asn Ala Glu Pro
 420 425 430

Leu Asp Tyr Ala Arg Val Pro Met Phe Ser Trp Thr His Arg Ser Ala
 435 440 445

Thr Pro Thr Asn Thr Ile Asp Pro Asp Val Ile Thr Gln Ile Pro Leu
 450 455 460

Val Lys Ala Phe Asn Leu His Ser Gly Ala Thr Ile Val Lys Gly Pro
 465 470 475 480

Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Val Gly Ser Phe
 485 490 495

Gly Asp Met Arg Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg
 500 505 510

Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu Gln Phe Tyr Thr Asn
 515 520 525

Ile Asn Gly Thr Thr Ile Asn Ile Gly Asn Phe Ser Ser Thr Met Asp
 530 535 540

Ser Gly Asp Asp Leu Gln Tyr Gly Arg Phe Arg Val Ala Gly Phe Thr
 545 550 555 560

Thr Pro Phe Thr Phe Ser Asp Ala Met Ser Thr Phe Thr Ile Gly Ala
 565 570 575

Phe Ser Phe Ser Ser Asn Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe
 580 585 590

Val Pro Ala Glu Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Lys Ala
 595 600 605

Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu
 610 615 620

Lys Thr Asp Val Thr Asp Tyr His Ile Asp Lys Val Ser Asn Leu Val
 625 630 635 640

Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser

645						650						655				
Glu	Lys	Val	Lys 660	His	Ala	Lys	Arg	Leu 665	Cys	Asp	Glu	Arg	Asn 670	Leu	Leu	
Gln	Asp	Pro 675	Asn	Phe	Arg	Gly	Ile 680	Asn	Arg	Gln	Pro	Asp 685	Arg	Gly	Trp	
Arg	Gly 690	Ser	Thr	Asp	Ile	Thr 695	Ile	Gln	Gly	Gly	Asp 700	Asp	Val	Phe	Lys	
Glu 705	Asn	Tyr	Val	Thr	Leu 710	Pro	Gly	Thr	Phe	Asp 715	Glu	Cys	Tyr	Pro	Thr 720	
Tyr	Leu	Tyr	Gln	Lys 725	Ile	Asp	Glu	Ser	Lys 730	Leu	Lys	Ala	Tyr	Thr 735	Arg	
Tyr	Glu	Leu	Arg 740	Gly	Tyr	Ile	Glu	Asp 745	Ser	Gln	Asp	Leu	Glu 750	Ile	Tyr	
Leu	Ile 755	Arg	Tyr	Asn	Ala	Lys	His 760	Glu	Thr	Val	Asn 765	Val	Pro	Gly	Thr	
Gly 770	Ser	Leu	Trp	Pro	Leu	Ser 775	Ala	Gln	Ser	Pro	Ile 780	Gly	Lys	Cys	Gly	
Glu 785	Pro	Asn	Arg	Cys	Ala 790	Thr	His	Leu	Glu	Trp 795	Asn	Pro	Asp	Leu	Asp 800	
Cys	Ser	Cys	Arg	Asp 805	Gly	Glu	Lys	Cys	Ala 810	His	His	Ser	His	His 815	Phe	
Ser	Leu	Asp	Ile 820	Asp	Val	Gly	Cys	Thr 825	Asp	Leu	Asn	Glu	Asp 830	Leu	Gly	
Val	Trp	Val	Ile 835	Phe	Lys	Ile	Lys 840	Thr	Gln	Asp	Gly	His 845	Ala	Arg	Leu	
Gly 850	Asn	Leu	Glu	Phe	Leu	Glu 855	Glu	Lys	Pro	Leu	Val 860	Gly	Glu	Ala	Leu	
Ala 865	Arg	Val	Lys	Arg	Ala 870	Glu	Lys	Lys	Trp	Arg 875	Asp	Lys	Arg	Glu	Lys 880	

Leu Glu Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Lys Ser Val
 885 890 895

Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr
 900 905 910

Asn Ile Ala Ile Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg
 915 920 925

Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala
 930 935 940

Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr
 945 950 955 960

Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser
 965 970 975

Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His
 980 985 990

Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu
 995 1000 1005

Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala
 1010 1015 1020

Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
 1025 1030 1035

Glu Asp Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu
 1040 1045 1050

Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala
 1055 1060 1065

Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly
 1070 1075 1080

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr
 1085 1090 1095

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp

```

1100                      1105                      1110

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

Pro Ala  Gly Tyr Val Thr Lys  Glu Leu Glu Tyr Phe  Pro Glu Thr
1130                      1135                      1140

Asp Lys  Val Trp Ile Glu Ile  Gly Glu Thr Glu Gly  Thr Phe Ile
1145                      1150                      1155

Val Asp  Ser Val Glu Leu Leu  Leu Met Glu Glu
1160                      1165

<210>  37
<211>  1138
<212>  PRT
<213>  Bacillus thuringiensis

<400>  37

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

Asn Pro Lys Ile Glu Met Leu Asp Ile Glu Gly Ile Ser Ser Arg Ser
20                      25                      30

Arg Glu Gln Val Ala Glu Ile Ser Leu Gly Leu Thr Arg Phe Leu Leu
35                      40                      45

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

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

Ile Glu Gln Leu Ile Asp Gln Arg Ile Glu Ala His Val Arg Asn Gln
85                      90                      95

Ala Ile Ser Arg Leu Glu Gly Leu Gly Asp Ser Tyr Glu Val Tyr Ile
100                      105                      110

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

```

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

Ala Ile Pro Ile Leu Arg Glu Gln Gly Phe Glu Ile Pro Leu Leu Thr
 145 150 155 160

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

Val Tyr Phe Gly Gln Arg Trp Gly Leu Asp Thr Ala Thr Val Asn Asn
 180 185 190

His Tyr Asn Arg Leu Ile Asn Leu Ile Asn Thr Tyr Ser Asp His Cys
 195 200 205

Ala Gln Trp Phe Asn Arg Gly Leu Asp Asn Phe Gly Val Val Thr Ala
 210 215 220

Arg Tyr Leu Asp Phe Gln Arg Glu Val Thr Ile Ser Val Leu Asp Ile
 225 230 235 240

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

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

Gly Ala Ser Leu Asn Val Asp Leu Arg Asn Ile Leu Arg Glu Pro His
 275 280 285

Leu Met Asp Phe Leu Thr Arg Leu Val Ile Tyr Thr Gly Val Gln Gly
 290 295 300

Gly Ile Tyr His Trp Ala Gly His Glu Ile Ser Ser Arg Thr Thr Gly
 305 310 315 320

Asn Leu Ser Ser Asn Ile Gln Phe Pro Leu Tyr Gly Thr Ser Ala Asn
 325 330 335

Ala Asp Arg Pro Phe Asn Leu Ala Ile His Tyr Ser Glu Thr Ile Tyr
 340 345 350

Arg Thr Leu Ser Ala Pro Ile Tyr Ser Val Ser Gly Gly Ile Ser Pro
 355 360 365

Asn Arg Thr Arg Ala Val Glu Gly Val Arg Phe Leu Thr Ala Arg Asp
 370 375 380

Asn Asn Leu Asn Ser Leu Pro Phe Leu Tyr Arg Lys Glu Gly Ser Leu
 385 390 395 400

Asp Ser Phe Thr Glu Leu Pro Pro Glu Asp Glu Asn Glu Pro Pro Tyr
 405 410 415

Ile Gly Tyr Ser His Arg Leu Cys His Ala Arg Phe Ala Arg Ser Ser
 420 425 430

Val Val Leu Glu Pro Ser Asn Phe Ala Arg Ile Pro Val Phe Ser Trp
 435 440 445

Thr His Arg Ser Ala Gly Pro Thr Asn Glu Val Ser Ser Ser Arg Ile
 450 455 460

Thr Gln Ile Pro Trp Val Lys Ala His Thr Leu Asp Ser Gly Ala Phe
 465 470 475 480

Val Ile Lys Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Pro
 485 490 495

Asn Leu Gly Thr Leu Gly Ala Leu Arg Val Thr Leu Thr Gly Gln Leu
 500 505 510

Pro Gln Thr Tyr Asn Ile Arg Ile Arg Tyr Ala Ser Ile Ala Asn Arg
 515 520 525

Gly Gly Thr Leu Ile Phe Ser Gln Pro Pro Ser Tyr Gly Leu Thr Phe
 530 535 540

Pro Lys Thr Met Asp Ile Asp Glu Pro Leu Thr Ser Arg Ser Phe Ala
 545 550 555 560

Arg Thr Thr Leu Phe Thr Pro Ile Thr Phe Thr Gln Ala Gln Ala Glu
 565 570 575

Leu Asn Leu Thr Ile Gln Gln Gly Val Tyr Ile Asp Arg Ile Glu Phe
 580 585 590

Ile Pro Val Asn Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala
 595 600 605

Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn Gln Leu Gly Leu
 610 615 620

Lys Thr Asp Leu Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val
 625 630 635 640

Asp Cys Leu Ser Asp Glu Phe Cys Ile Asp Glu Lys Arg Glu Leu Ser
 645 650 655

Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu
 660 665 670

Gln Asp Ser Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp
 675 680 685

Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asn Asp Val Phe Lys
 690 695 700

Glu Asn Tyr Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr
 705 710 715 720

Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg
 725 730 735

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

Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr
 755 760 765

Gly Ser Leu Trp Pro Leu Ser Val Glu Ser Pro Ile Gly Lys Cys Gly
 770 775 780

Glu Pro Asn Arg Cys Val Pro Gln Leu Glu Trp Asn Ser Asn Leu Asp
 785 790 795 800

Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe
 805 810 815

Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu His Asp Asp Leu Gly
 820 825 830

Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu
 835 840 845

Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu
 850 855 860

Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr
 865 870 875 880

Leu Gln Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val
 885 890 895

Asp Ala Leu Phe Ala Asn Ser Gln Tyr Asn Arg Leu Gln Ala Asp Thr
 900 905 910

Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg
 915 920 925

Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Gly
 930 935 940

Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr
 945 950 955 960

Asp Ala Arg Asn Val Ile Lys Asn Ser Asp Phe Asn Asn Gly Leu Ser
 965 970 975

Cys Trp Asn Val Lys Gly His Val Asp Ile Glu Glu Gln Asn Asn His
 980 985 990

Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Lys
 995 1000 1005

Val His Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala
 1010 1015 1020

Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
 1025 1030 1035

Glu Asp His Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu Glu Asp
 1040 1045 1050

Glu Val Tyr Pro Asn Asn Thr Arg Thr Cys Asn Ala Tyr Pro Ala
 1055 1060 1065

Asp Gln Glu Gly Tyr Glu Gly Ala Cys Thr Ser Arg Asn Arg Gly
 1070 1075 1080

Tyr Asp Glu Val Tyr Gly Asn Thr Pro Ser Leu Pro Ala Asp Tyr
 1085 1090 1095

Ala Pro Ile Tyr Glu Glu Asn Ala Tyr Thr Asp Gly Arg Arg Gly
 1100 1105 1110

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

Pro Ala Gly Tyr Glu Thr Lys Glu Leu Glu
 1130 1135

<210> 38
 <211> 719
 <212> PRT
 <213> Bacillus thuringiensis

<400> 38

Met Lys Leu Lys Asn Pro Asp Lys His Gln Ser Leu Ser Ser Asn Ala
 1 5 10 15

Lys Val Asp Lys Ile Ala Thr Asp Ser Leu Lys Asn Glu Thr Asp Ile
 20 25 30

Glu Leu Lys Asn Met Asn Asn Glu Asp Tyr Leu Arg Met Ser Glu His
 35 40 45

Glu Ser Ile Asp Pro Phe Val Ser Ala Ser Thr Ile Gln Thr Gly Ile
 50 55 60

Gly Ile Ala Gly Lys Ile Leu Gly Thr Leu Gly Val Pro Phe Ala Gly
 65 70 75 80

Gln Ile Ala Ser Leu Tyr Ser Phe Ile Leu Gly Glu Leu Trp Pro Lys
 85 90 95

Gly Lys Ser Gln Trp Glu Ile Phe Met Glu His Val Glu Glu Ile Ile
 100 105 110

Asn Gln Lys Ile Leu Thr Tyr Ala Arg Asn Lys Ala Leu Ser Asp Leu
 115 120 125

Arg Gly Leu Gly Asp Ala Leu Ala Val Tyr His Glu Ser Leu Glu Ser
 130 135 140

Trp Val Lys Asn Arg Asn Asn Thr Arg Ala Arg Ser Val Val Lys Asn
 145 150 155 160

Gln Tyr Ile Ala Leu Glu Leu Met Phe Val Gln Lys Leu Pro Ser Phe
 165 170 175

Ala Val Ser Gly Glu Glu Val Pro Leu Leu Pro Ile Tyr Ala Gln Ala
 180 185 190

Ala Asn Leu His Leu Leu Leu Leu Arg Asp Ala Ser Ile Phe Gly Lys
 195 200 205

Glu Trp Gly Leu Ser Ala Ser Glu Ile Ser Thr Phe Tyr Asn Arg Gln
 210 215 220

Val Glu Arg Thr Arg Asp Tyr Ser Asp His Cys Val Lys Trp Tyr Asn
 225 230 235 240

Thr Gly Leu Asn Asn Leu Arg Gly Thr Asn Ala Lys Ser Trp Val Arg
 245 250 255

Tyr Asn Gln Phe Arg Lys Asp Met Thr Leu Met Val Leu Asp Leu Val
 260 265 270

Ala Leu Phe Pro Ser Tyr Asp Thr Leu Val Tyr Pro Ile Lys Thr Thr
 275 280 285

Ser Gln Leu Thr Arg Glu Val Tyr Thr Asp Ala Ile Gly Thr Val His
 290 295 300

Pro Ser Gly Ala Val Ala Ser Thr Thr Trp Tyr Asn Asn Asn Ala Pro
 305 310 315 320

Ser Phe Ser Thr Ile Glu Ala Ala Val Val Arg Asn Pro His Leu Leu
 325 330 335

Asp Phe Leu Glu Gln Val Thr Ile Tyr Ser Leu Leu Ser Arg Trp Ser

340	345	350
Asn Thr Gln Tyr Met Asn Met Trp Gly Gly His Lys Leu Glu Phe Arg		
355	360	365
Thr Ile Gly Gly Thr Leu Asn Thr Ser Thr Gln Gly Ser Thr Asn Thr		
370	375	380
Ser Ile Asn Pro Val Thr Leu Pro Phe Thr Ser Arg Asp Val Tyr Arg		
385	390	400
Thr Glu Ser Leu Ala Gly Leu Asn Leu Phe Leu Thr Gln Pro Val Asn		
405	410	415
Gly Val Pro Arg Val Asp Phe His Trp Lys Phe Val Thr His Pro Ile		
420	425	430
Ala Ser Asp Asn Phe Tyr Tyr Pro Gly Tyr Ala Gly Ile Gly Thr Gln		
435	440	445
Leu Gln Asp Ser Glu Asn Glu Leu Pro Pro Glu Thr Thr Gly Gln Pro		
450	455	460
Asn Tyr Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ser		
465	470	475
Ala Ser His Val Lys Ala Leu Val Tyr Ser Trp Thr His Arg Ser Ala		
485	490	495
Asp Arg Thr Asn Thr Ile Asn Ser Asp Ser Ile Thr Gln Ile Pro Leu		
500	505	510
Val Lys Ala Phe Asn Leu Pro Ser Gly Ala Ser Val Val Arg Gly Pro		
515	520	525
Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Thr Phe		
530	535	540
Gly Asp Ile Arg Val Asn Ile Asn Pro Pro Phe Ala Gln Arg Tyr Arg		
545	550	555
Leu Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu Glu Phe His Thr Ser		
565	570	575

Ile Asn Gly Lys Ala Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Asn
580 585 590

Arg Gly Glu Asp Leu Asp Tyr Lys Ala Phe Arg Thr Val Gly Phe Thr
595 600 605

Thr Pro Phe Ser Phe Ser Asn Ala Gln Ser Thr Phe Thr Ile Gly Ala
610 615 620

Trp Asn Phe Ser Leu Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe
625 630 635 640

Val Pro Val Glu Val Thr Tyr Glu Ala Glu His Asp Phe Glu Lys Ala
645 650 655

Gln Glu Glu Val Thr Ala Leu Phe Thr Ser Thr Asn Pro Gly Gly Leu
660 665 670

Lys Thr Asn Val Thr Glu Tyr His Ile Asp Gln Val Ser Asn Leu Val
675 680 685

Glu Ser Leu Ser Asn Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Phe
690 695 700

Glu Ile Val Lys Tyr Ala Lys Gln Leu His Ile Gly Arg Asn Met
705 710 715

<210> 39
<211> 1164
<212> PRT
<213> Bacillus thuringiensis

<400> 39

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

Asn Pro Glu Val Gly Ile Leu Asp Ile Glu Asn Phe Asn Leu Glu Leu
20 25 30

Val Ser Gln Val Ser Val Gly Leu Thr Arg Phe Leu Leu Glu Ala Ser
35 40 45

Ile Pro Gly Ala Gly Phe Ala Leu Gly Leu Phe Asp Ile Ile Trp Gly
50 55 60

Ala	Leu	Gly	Val	Asp	Gln	Trp	Ser	Leu	Phe	Leu	Ala	Gln	Ile	Glu	Gln	65	70	75	80
Leu	Ile	Asn	Glu	Arg	Ile	Thr	Thr	Val	Glu	Arg	Asn	Arg	Ala	Ile	Gln	85	90	95	
Ala	Leu	Ser	Gly	Leu	Ser	Ser	Ser	Tyr	Glu	Val	Tyr	Ile	Glu	Ala	Leu	100	105	110	
Arg	Glu	Trp	Glu	Asn	Asp	Leu	Asp	Asn	Pro	Ala	Ser	Arg	Asp	Arg	Val	115	120	125	
Val	Ala	Arg	Phe	Arg	Ala	Thr	Asp	Asn	Ser	Leu	Ile	Thr	Asp	Ile	Pro	130	135	140	
Leu	Leu	Glu	Ile	Pro	Gly	Phe	Glu	Ile	Ala	Thr	Leu	Ser	Val	Tyr	Thr	145	150	155	160
Gln	Ala	Ala	Asn	Leu	His	Leu	Ala	Leu	Leu	Arg	Asp	Ala	Val	Tyr	Phe	165	170	175	
Gly	Glu	Arg	Trp	Gly	Leu	Thr	Gln	Thr	Asn	Ile	Glu	Asp	Leu	His	Thr	180	185	190	
Arg	Leu	Thr	Arg	Tyr	Ile	Gln	Glu	Tyr	Ser	Asp	His	Cys	Ala	Arg	Trp	195	200	205	
Tyr	Asn	Gln	Gly	Leu	Asn	Asn	Ile	Gly	Gly	Ile	Asn	Thr	Arg	Tyr	Leu	210	215	220	
Asp	Phe	Gln	Arg	Glu	Leu	Thr	Ile	Ser	Val	Leu	Asp	Ile	Val	Ala	Leu	225	230	235	240
Phe	Pro	Asn	Tyr	Asp	Ile	Arg	Thr	Tyr	Pro	Ile	Pro	Thr	Gln	Ser	Gln	245	250	255	
Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Ser	Pro	Val	Val	Ala	Pro	Gly	Val	Asn	260	265	270	
Trp	Ile	Leu	Ser	Ile	Ser	Asn	Val	Leu	Arg	Ala	Pro	His	Leu	Met	Asp	275	280	285	

Phe Phe Asp Arg Ile Ile Ile Tyr Thr Gly Thr Val Arg Ser Thr Pro
290 295 300

His Trp Glu Gly His Glu Val Ile Ser Arg Arg Thr Gly Gln Gly Asn
305 310 315 320

Glu Ile Arg Ser Pro Leu Tyr Gly Val Ala Ala Asn Ala Glu Pro Pro
325 330 335

Val Thr Ile Arg Pro Thr Gly Phe Thr Asp Glu Gln Arg Gln Val Tyr
340 345 350

Arg Val Leu Ser Arg Val Ala Ser Phe Arg Asn Ser Gly Thr Asn Phe
355 360 365

Ser Leu Val Asp Ala Ala Ser Phe Leu Thr Ile Phe Ser Ala Ser Ser
370 375 380

Ile Tyr Arg Asn Gly Phe Gly Phe Asn Ala Asp Thr Ile Asp Glu Ile
385 390 395 400

Pro Ile Glu Gly Thr Asp Pro Tyr Ile Gly Tyr Ser His Arg Leu Cys
405 410 415

His Val Gly Phe Thr Ala Ser Ser Pro Phe Ile Ser Gln Tyr Ala Arg
420 425 430

Ala Pro Val Phe Ser Trp Thr His Arg Ser Ala Thr Phe Thr Asn Thr
435 440 445

Ile Asp Pro Glu Arg Ile Thr Gln Ile Pro Met Val Lys Ala Tyr Asn
450 455 460

Leu His Ala Gly Ala Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly
465 470 475 480

Asp Leu Leu Arg Arg Thr Asn Thr Gly Thr Phe Ala Asp Ile Arg Val
485 490 495

Asn Ile Thr Gly Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr
500 505 510

Ala Ser Thr Thr Asp Leu Gln Phe Phe Thr Arg Ile Asn Gly Thr Ser
515 520 525

Val	Asn	Gln	Gly	Asn	Phe	Gln	Arg	Thr	Met	Asn	Arg	Gly	Asp	Asn	Leu
530						535					540				
Glu	Ser	Gly	Asn	Phe	Arg	Thr	Ala	Gly	Phe	Ser	Thr	Pro	Phe	Ser	Phe
545					550					555					560
Ser	Asn	Ala	Gln	Ser	Thr	Phe	Thr	Leu	Gly	Thr	Gln	Ala	Phe	Ser	Asn
				565					570					575	
Gln	Glu	Val	Tyr	Ile	Asp	Arg	Ile	Glu	Phe	Val	Pro	Ala	Glu	Val	Thr
			580					585					590		
Phe	Glu	Ala	Glu	Ser	Asp	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val	Asn	Ala
		595					600					605			
Leu	Phe	Thr	Ser	Thr	Ser	Gln	Leu	Gly	Leu	Lys	Thr	Asn	Val	Thr	Gly
610						615					620				
Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Ala	Cys	Leu	Ser	Asp	Glu
625					630					635					640
Phe	Cys	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys	His	Ala
				645					650					655	
Lys	Arg	Leu	Ser	Asp	Lys	Arg	Asn	Leu	Leu	Gln	Asp	Pro	Asn	Phe	Arg
			660					665					670		
Gly	Ile	Asn	Arg	Gln	Pro	Asp	His	Gly	Trp	Arg	Gly	Ser	Thr	Asp	Ile
		675					680					685			
Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	Thr	Leu
690						695					700				
Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	Lys	Ile
705					710					715					720
Asp	Glu	Ser	Lys	Leu	Lys	Ala	Tyr	Thr	Arg	Tyr	Gln	Leu	Arg	Gly	Tyr
				725					730					735	
Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Leu	Ile	Arg	Tyr	Asn	Ser
			740					745					750		

Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu
755 760 765

Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg Cys Ala
770 775 780

Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp Gly
785 790 795 800

Glu Lys Cys Val His His Ser His His Phe Ser Leu Asp Ile Asp Val
805 810 815

Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Leu Ile Phe Lys
820 825 830

Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu
835 840 845

Glu Glu Glu Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Arg Ala
850 855 860

Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr Asn
865 870 875 880

Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn
885 890 895

Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His
900 905 910

Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu
915 920 925

Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu
930 935 940

Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr Asp Ala Arg Asn Val Ile
945 950 955 960

Lys Asn Gly Asn Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys Gly
965 970 975

His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Ile
980 985 990

Pro Glu Trp Glu Ala Glu Val Ser Gln Lys Val Arg Val Cys Pro Gly
 995 1000 1005

Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
 1010 1015 1020

Glu Gly Cys Val Thr Ile His Glu Ile Glu Asp Asn Thr Asp Glu
 1025 1030 1035

Leu Lys Phe Ser Asn Cys Val Glu Glu Gly Tyr Pro Asn Asn Thr
 1040 1045 1050

Val Thr Cys Asn Glu Tyr Thr Met Asn Gln Gly Val Gly Glu Cys
 1055 1060 1065

Thr Asp Ala Cys Asn Val Arg Asn Arg Gly Tyr Glu Asp Ala Tyr
 1070 1075 1080

Gly His Asn Pro Ser Thr Pro Val His Tyr Thr Thr Pro Tyr Glu
 1085 1090 1095

Glu Glu Thr Tyr Thr Asp Glu Arg Arg Glu Asn Pro Cys Glu Ala
 1100 1105 1110

Asn Lys Gly Tyr Val Asn Tyr Thr Pro Leu Pro Val Gly Tyr Val
 1115 1120 1125

Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Thr Val Trp Ile
 1130 1135 1140

Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu
 1145 1150 1155

Leu Leu Leu Met Glu Glu
 1160

<210> 40

<211> 1166

<212> PRT

<213> Bacillus thuringiensis

<400> 40

Met Lys Ile Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser

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

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

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

Glu Phe Tyr Tyr Gly Pro Ile Tyr Asn Tyr Asn Ile Val Gly Arg Leu
 275 280 285

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

Met Ile Met Tyr Thr Ser Asp Asn Arg Arg Glu His Tyr Trp Ser Gly
 305 310 315 320

Leu Glu Met Lys Ala Thr Asp Thr Ser Gly Asn Gln Val Ser Phe Pro
 325 330 335

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

Asn Gly Glu Gly Val Tyr Arg Ile Leu Ser Glu Pro Phe Tyr Ser Ser
 355 360 365

Pro Phe Leu Gly Thr Ser Val Leu Gly Ser Arg Gly Glu Glu Phe Ala
 370 375 380

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

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

Val Pro Pro His Arg Gly Tyr Ser His Leu Leu Ser His Val Thr Met
 420 425 430

His Asn Ser Ser Pro Ile Phe His Trp Thr His Arg Ser Ala Thr Pro
 435 440 445

Arg Asn Thr Ile Asp Pro Asp Ser Ile Thr Gln Ile Pro Val Val Lys
 450 455 460

Ala Ser His Leu Ser Gly Gly Ser Val Ile Lys Gly Pro Gly His Thr

465		470		475		480
Gly Gly Asp Leu Ile Ser Leu Pro Val Asn Asn Phe Thr His Phe Arg						
		485		490		495
Ile Pro Phe Gln Ala Asn Thr Pro Gln Arg Tyr Arg Ile Arg Ile Arg						
		500		505		510
Tyr Ala Ala Asp Ser Asp Gly Thr Leu Asp Ser Gly Val Phe Leu Ser						
		515		520		525
Ala Ala Ala Gly Asp Gly Phe Asn Thr Thr Ser Tyr Arg Ala Thr Met						
		530		535		540
Ser Pro Gly Gly Ser Leu Thr Ser Arg Asp Phe Gln Phe Leu Asp Leu						
545		550		555		560
Asn Thr Ser Phe Thr Ser Asp Val Ala Ser Asn Leu Trp Leu His Phe						
		565		570		575
Ile Arg Tyr Ile Arg Pro Gly Asn Leu Tyr Ile Asp Arg Ala Glu Phe						
		580		585		590
Ile Pro Val Asp Ala Thr Phe Glu Ala Gly Tyr Asn Leu Glu Arg Ala						
		595		600		605
Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Lys Gly Leu						
		610		615		620
Gln Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val						
625		630		635		640
Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser						
		645		650		655
Glu Lys Val Lys Gln Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu						
		660		665		670
Gln Asp Ser Asn Phe Arg Gly Ile Asn Arg Glu Gln Asp Arg Gly Trp						
		675		680		685
Arg Gly Ser Thr His Ile Thr Ile Gln Gly Gly Asn Asp Val Phe Lys						
		690		695		700

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

930		935		940
Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr				
945		950	955	960
Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser				
	965		970	975
Cys Trp Asn Val Lys Gly His Val Asp Val Glu Val Gln Asn Asn His				
	980		985	990
Arg Ser Val Leu Val Ile Ala Glu Trp Glu Ala Glu Val Ser Gln Glu				
	995	1000		1005
Val Pro Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala				
1010		1015		1020
Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile				
1025		1030		1035
Glu Asp His Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu Glu Glu				
1040		1045		1050
Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Gly Tyr Thr Ala				
1055		1060		1065
Thr Gln Glu Glu Tyr Lys Asp Ala Tyr Thr Ser Arg Asn Arg Gly				
1070		1075		1080
Tyr Asp Glu Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr				
1085		1090		1095
Ala Ser Val Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Glu				
1100		1105		1110
Asn Pro Cys Glu Met Glu Arg Gly Tyr Thr Pro Leu Pro Val Gly				
1115		1120		1125
Tyr Ile Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Thr Val				
1130		1135		1140
Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Ile Asp Ser				
1145		1150		1155

Val Glu Leu Leu Leu Met Glu Glu
1160 1165

<210> 41

<211> 1181

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Protein BT-0001 đột biến

<400> 41

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro

405																410																415																		
Gln	Asp	Asn	Ser	Val	Pro	Pro	Arg	Ala	Gly	Phe	Ser	His	Arg	Leu	Ser																																			
			420																	425																	430													
His	Val	Thr	Met	Leu	Ser	Gln	Ala	Ala	Gly	Ala	Val	Tyr	Thr	Leu	Arg																																			
			435																	440																	445													
Ala	Pro	Thr	Phe	Ser	Trp	Gln	His	Arg	Ser	Ala	Glu	Phe	Asn	Asn	Ile																																			
			450																	455																	460													
Ile	Ala	Ser	Asp	Ser	Ile	Thr	Gln	Ile	Pro	Ala	Val	Lys	Gly	Asn	Phe																																			
			465																	470																	475													
Leu	Phe	Asn	Gly	Ser	Val	Ile	Ser	Gly	Pro	Gly	Phe	Thr	Gly	Gly	Asp																																			
			485																	490																	495													
Leu	Val	Arg	Leu	Asn	Ser	Ser	Gly	Asn	Asn	Ile	Gln	Asn	Arg	Gly	Tyr																																			
			500																	505																	510													
Ile	Glu	Val	Pro	Ile	His	Phe	Pro	Ser	Thr	Ser	Thr	Arg	Tyr	Arg	Val																																			
			515																	520																	525													
Arg	Val	Arg	Tyr	Ala	Ser	Val	Thr	Pro	Ile	His	Leu	Asn	Val	Asn	Trp																																			
			530																	535																	540													
Gly	Asn	Ser	Ser	Ile	Phe	Ser	Asn	Thr	Val	Pro	Ala	Thr	Ala	Thr	Ser																																			
			545																	550																	555													
Leu	Asp	Asn	Leu	Gln	Ser	Ser	Asp	Phe	Gly	Tyr	Phe	Glu	Ser	Ala	Asn																																			
			565																	570																	575													
Ala	Phe	Thr	Ser	Ser	Leu	Gly	Asn	Ile	Val	Gly	Val	Arg	Asn	Phe	Ser																																			
			580																	585																	590													
Gly	Thr	Ala	Gly	Val	Ile	Ile	Asp	Arg	Phe	Glu	Phe	Ile	Pro	Val	Thr																																			
			595																	600																	605													
Ala	Thr	Leu	Glu	Ala	Glu	Tyr	Asn	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val																																			
			610																	615																	620													
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Leu	Lys	Thr	Asn	Val																																			
			625																	630																	635													

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Cys Leu Ser
645 650 655

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
675 680 685

Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
690 695 700

Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705 710 715 720

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

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

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

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

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

Cys Ala Pro Arg Ile Glu Trp Lys Pro Asp Val Asp Cys Ser Cys Arg
805 810 815

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

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

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

Phe Leu Glu Glu Lys Leu Leu Leu Gly Glu Ala Leu Ala Arg Val Lys

865		870		875		880									
Lys	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp	Glu
			885					890						895	
Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe
			900					905					910		
Val	Asp	Ser	Gln	Tyr	Asn	Arg	Leu	Gln	Thr	Asp	Thr	Asn	Ile	Ala	Met
		915					920					925			
Ile	His	Ala	Ala	Asp	Lys	Arg	Val	His	Arg	Ile	Arg	Glu	Ala	Tyr	Leu
	930					935					940				
Pro	Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala	Ile	Phe	Glu	Glu
945					950					955					960
Leu	Glu	Gly	Leu	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr	Asp	Ala	Arg	Asn
			965						970					975	
Val	Ile	Lys	Asn	Gly	Asp	Phe	Asn	Tyr	Gly	Leu	Ser	Cys	Trp	Asn	Val
			980					985					990		
Lys	Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	His	Arg	Ser	Val	Leu
		995					1000					1005			
Val	Ile	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val	
	1010					1015					1020				
Cys	Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu	
	1025					1030					1035				
Gly	Tyr	Gly	Glu	Gly	Cys	Val	Thr	Ile	His	Glu	Ile	Glu	Asp	Asn	
	1040					1045					1050				
Thr	Asp	Glu	Leu	Lys	Phe	Ser	Asn	Cys	Val	Glu	Glu	Glu	Val	Tyr	
	1055					1060					1065				
Pro	Asn	Asn	Thr	Val	Thr	Cys	Asn	Asp	Tyr	Thr	Ala	Thr	Gln	Glu	
	1070					1075					1080				
Glu	Tyr	Glu	Gly	Thr	Tyr	Thr	Ser	Arg	Asn	Arg	Gly	Tyr	Asp	Gly	
	1085					1090					1095				

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

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

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

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

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

Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 42

<211> 1176

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Protein BT-0003 đột biến

<400> 42

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

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

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

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

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

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

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

Ala Phe Ala Glu Trp Glu Lys Ala Pro Asp Asp Pro Glu Leu Arg Glu
 115 120 125

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

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

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

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

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

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

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

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

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

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

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

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

Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro

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

Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn Phe
565 570 575

Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn Ser
580 585 590

Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val
595 600 605

Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn
610 615 620

Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr
625 630 635 640

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp
645 650 655

Glu Phe Cys Leu Asp Glu Lys Gln Glu Leu Ser Glu Lys Val Lys His
660 665 670

Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe
675 680 685

Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp
690 695 700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
705 710 715 720

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

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

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

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

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys

785		790		795		800									
Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg	Asp
			805						810					815	
Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	Asp
			820					825					830		
Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly	Val	Trp	Val	Ile	Phe
		835					840					845			
Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu	Gly	Asn	Leu	Glu	Phe
	850					855						860			
Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	Ile
865					870					875					880
Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp	Glu	Thr
				885					890					895	
Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe	Val
			900					905					910		
Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr	Asn	Ile	Ala	Met	Ile
		915					920					925			
His	Ala	Ala	Asp	Lys	Arg	Val	His	Ser	Ile	Arg	Glu	Ala	Tyr	Leu	Pro
	930					935					940				
Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala	Ile	Phe	Glu	Glu	Leu
945					950				955						960
Glu	Gly	Arg	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr	Asp	Ala	Arg	Asn	Val
			965						970					975	
Ile	Lys	Asn	Gly	Asp	Phe	Asn	Asn	Gly	Leu	Ser	Cys	Trp	Asn	Val	Lys
			980					985					990		
Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	Gln	Arg	Ser	Val	Leu	Val
		995					1000					1005			
Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val	Cys	
	1010					1015					1020				

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

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

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

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

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

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

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

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

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

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

Met Glu Glu
1175

<210> 43

<211> 1237

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Protein BT-0020 đột biến

<400> 43

Met Asn Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asp Ala Ser Phe
1 5 10 15

Ile Pro Ala Val Ser Asn Glu Ser Val Thr Ile Ser Lys Glu Tyr Ala
 20 25 30

Gln Thr Asn Gln Leu Gln Asn Asn Ser Ile Glu Asp Gly Leu Cys Ile
 35 40 45

Ala Glu Gly Glu Tyr Ile Asp Pro Phe Val Ser Ala Ser Thr Val Gln
 50 55 60

Thr Gly Ile Ser Ile Ala Gly Arg Ile Leu Gly Val Leu Gly Val Pro
 65 70 75 80

Phe Ala Gly Gln Leu Ala Ser Phe Tyr Ser Phe Ile Val Gly Glu Leu
 85 90 95

Trp Pro Lys Gly Arg Asp Gln Trp Glu Ile Phe Met Glu His Val Glu
 100 105 110

Gln Leu Val Arg Gln Gln Ile Thr Ala Asn Ala Arg Asn Thr Ala Leu
 115 120 125

Ala Arg Leu Gln Gly Leu Gly Asn Ser Phe Arg Ala Tyr Gln Gln Ser
 130 135 140

Leu Glu Asp Trp Leu Glu Asn Arg Asn Asp Ala Arg Thr Arg Ser Val
 145 150 155 160

Leu Tyr Thr Gln Tyr Ile Ala Leu Glu Leu Asp Phe Leu Asn Ala Met
 165 170 175

Pro Leu Phe Ala Ile Arg Glu Gln Glu Val Pro Leu Leu Met Val Tyr
 180 185 190

Ala Gln Ala Ala Asn Leu His Leu Leu Leu Leu Arg Asp Ala Ser Leu
 195 200 205

Tyr Gly Arg Glu Phe Gly Leu Thr Ser Gln Glu Ile Gln Arg Tyr Tyr
 210 215 220

Glu Arg Gln Val Glu Arg Thr Arg Asp Tyr Ser Asp His Cys Val Gln
 225 230 235 240

Trp Tyr Asn Thr Gly Leu Asn Asn Leu Arg Gly Thr Asn Ala Glu Ser

245							250							255						
Trp	Val	Arg	Tyr	Asn	Gln	Phe	Arg	Arg	Asp	Leu	Thr	Leu	Gly	Val	Leu					
			260					265					270							
Asp	Leu	Val	Ala	Leu	Phe	Pro	Ser	Tyr	Asp	Thr	Arg	Thr	Tyr	Pro	Ile					
		275					280					285								
Asn	Thr	Ser	Ala	Gln	Leu	Thr	Arg	Glu	Val	Tyr	Thr	Asp	Ala	Ile	Gly					
	290					295					300									
Ala	Thr	Gly	Val	Asn	Met	Ala	Ser	Met	Asn	Trp	Tyr	Asn	Asn	Asn	Ala					
305					310					315					320					
Pro	Ser	Phe	Ser	Ala	Ile	Glu	Thr	Ala	Val	Ile	Arg	Ser	Pro	His	Leu					
				325					330					335						
Leu	Asp	Phe	Leu	Glu	Gln	Leu	Lys	Ile	Phe	Ser	Ala	Ser	Ser	Arg	Trp					
			340					345					350							
Ser	Asn	Thr	Arg	His	Met	Thr	Tyr	Trp	Arg	Gly	His	Thr	Ile	Gln	Ser					
		355					360					365								
Arg	Pro	Ile	Arg	Gly	Ala	Leu	Ile	Thr	Ser	Thr	His	Gly	Asn	Thr	Asn					
	370					375					380									
Thr	Ser	Ile	Asn	Pro	Val	Thr	Phe	Gln	Phe	Pro	Ser	Arg	Asp	Val	Tyr					
385					390					395					400					
Arg	Thr	Glu	Ser	Tyr	Ala	Gly	Val	Leu	Leu	Trp	Gly	Ile	Tyr	Leu	Glu					
				405					410					415						
Pro	Ile	His	Gly	Val	Pro	Thr	Val	Arg	Phe	Asn	Phe	Arg	Asn	Pro	Gln					
			420					425					430							
Asn	Thr	Phe	Glu	Arg	Gly	Thr	Ala	Asn	Tyr	Ser	Gln	Pro	Tyr	Glu	Ser					
		435					440					445								
Pro	Gly	Leu	Gln	Leu	Lys	Asp	Ser	Glu	Thr	Glu	Leu	Pro	Pro	Glu	Thr					
	450					455					460									
Thr	Glu	Arg	Pro	Asn	Tyr	Glu	Ser	Tyr	Ser	His	Arg	Leu	Ser	His	Ile					
465					470					475					480					

Gly Ile Ile Leu Gln Thr Arg Leu Asn Val Pro Val Tyr Ser Trp Thr
 485 490 495

His Arg Ser Ala Asp Arg Thr Asn Thr Ile Gly Pro Asn Arg Ile Thr
 500 505 510

Gln Ile Pro Ala Val Lys Gly Asn Leu Leu Phe Asn Gly Ser Val Ile
 515 520 525

Ser Gly Pro Gly Phe Thr Gly Gly Asp Leu Val Arg Leu Asn Asn Ser
 530 535 540

Gly Asn Asn Ile Gln Asn Arg Gly Tyr Leu Glu Val Pro Ile Gln Phe
 545 550 555 560

Thr Ser Thr Ser Thr Arg Tyr Arg Val Arg Val Arg Tyr Ala Ser Val
 565 570 575

Thr Pro Ile His Leu Ser Val Asn Trp Gly Asn Ser Asn Ile Phe Ser
 580 585 590

Ser Thr Val Pro Ala Thr Ala Ala Ser Leu Asp Asn Leu Gln Ser Arg
 595 600 605

Asp Phe Gly Tyr Phe Glu Ser Thr Asn Ala Phe Thr Ser Val Thr Gly
 610 615 620

Asn Val Val Gly Val Arg Asn Phe Ser Glu Asn Ala Arg Val Ile Ile
 625 630 635 640

Asp Arg Phe Glu Phe Ile Pro Val Thr Ala Thr Phe Glu Ala Glu Tyr
 645 650 655

Asp Leu Glu Arg Ala Gln Glu Ala Val Asn Ala Leu Phe Thr Asn Thr
 660 665 670

Asn Pro Arg Arg Leu Lys Thr Gly Val Thr Asp Tyr His Ile Asp Glu
 675 680 685

Val Ser Asn Leu Val Ala Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu
 690 695 700

Lys Arg Glu Leu Leu Glu Lys Val Lys Tyr Ala Lys Arg Leu Ser Asp

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

Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr Lys Arg Val Tyr
 945 950 955 960

Thr Glu Ala Lys Glu Ala Val Asp Ala Leu Phe Val Asp Ser Gln Tyr
 965 970 975

Asp Arg Leu Gln Ala Asp Thr Asn Ile Gly Met Ile His Ala Ala Asp
 980 985 990

Lys Leu Val His Arg Ile Arg Glu Ala Tyr Leu Ser Glu Leu Ser Val
 995 1000 1005

Ile Pro Gly Val Asn Ala Glu Ile Phe Glu Glu Leu Glu Gly Arg
 1010 1015 1020

Ile Ile Thr Ala Ile Ser Leu Tyr Asp Ala Arg Asn Val Val Lys
 1025 1030 1035

Asn Gly Asp Phe Asn Asn Gly Leu Ala Cys Trp Asn Val Lys Gly
 1040 1045 1050

His Val Asp Val Gln Gln Ser His His Arg Ser Val Leu Val Ile
 1055 1060 1065

Pro Glu Trp Glu Ala Glu Val Ser Gln Ala Val Arg Val Cys Pro
 1070 1075 1080

Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr
 1085 1090 1095

Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp
 1100 1105 1110

Glu Leu Lys Phe Lys Asn Cys Glu Glu Glu Glu Val Tyr Pro Thr
 1115 1120 1125

Asp Thr Gly Thr Cys Asn Asp Tyr Thr Ala His Gln Gly Thr Ala
 1130 1135 1140

Ala Cys Asn Ser Arg Asn Ala Gly Tyr Glu Asp Ala Tyr Glu Val
 1145 1150 1155

Asp Thr Thr Ala Ser Val Asn Tyr Lys Pro Thr Tyr Glu Glu Glu

1160 1165 1170
 Thr Tyr Thr Asp Val Arg Arg Asp Asn His Cys Glu Tyr Asp Arg
 1175 1180 1185
 Gly Tyr Val Asn Tyr Pro Pro Leu Pro Ala Gly Tyr Val Thr Lys
 1190 1195 1200
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Leu
 1205 1210 1215
 Gly Glu Thr Glu Gly Thr Phe Leu Val Asp Ser Ile Glu Leu Leu
 1220 1225 1230
 Leu Met Glu Glu
 1235

 <210> 44
 <211> 715
 <212> PRT
 <213> Trình tự Nhân tạo

 <220>
 <223> Protein BT-0022 đột biến

 <400> 44

 Met Lys Ser Lys Asn Gln Asn Met His Gln Ser Leu Ser Asn Asn Ala
 1 5 10 15

 Thr Val Asp Lys Asn Phe Thr Gly Ser Leu Glu Asn Asn Thr Asn Thr
 20 25 30

 Glu Leu Gln Asn Phe Asn His Glu Gly Ile Glu Pro Phe Val Ser Val
 35 40 45

 Ser Thr Ile Gln Thr Gly Ile Gly Ile Ala Gly Lys Ile Leu Gly Asn
 50 55 60

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

 Leu Gly Glu Leu Trp Pro Lys Gly Lys Ser Gln Trp Glu Ile Phe Met
 85 90 95

 Glu His Val Glu Glu Leu Ile Asn Gln Lys Ile Ser Thr Tyr Ala Arg

				100				105				110			
Asn	Lys	Ala	Leu	Ala	Asp	Leu	Lys	Gly	Leu	Gly	Asp	Ala	Leu	Ala	Val
		115				120						125			
Tyr	His	Glu	Ser	Leu	Glu	Ser	Trp	Ile	Lys	Asn	Arg	Asn	Asn	Thr	Arg
130						135				140					
Thr	Arg	Ser	Val	Val	Lys	Ser	Gln	Tyr	Ile	Thr	Leu	Glu	Leu	Met	Phe
145				150						155				160	
Val	Gln	Ser	Leu	Pro	Ser	Phe	Ala	Val	Ser	Gly	Glu	Glu	Val	Pro	Leu
				165				170						175	
Leu	Pro	Ile	Tyr	Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Leu	Leu	Leu	Arg
		180						185				190			
Asp	Ala	Ser	Ile	Phe	Gly	Lys	Glu	Trp	Gly	Leu	Ser	Asp	Ser	Glu	Ile
		195				200						205			
Ser	Thr	Phe	Tyr	Asn	Arg	Gln	Val	Glu	Arg	Thr	Ser	Asp	Tyr	Ser	Asp
210						215				220					
His	Cys	Thr	Lys	Trp	Phe	Asp	Thr	Gly	Leu	Asn	Arg	Leu	Lys	Gly	Ser
225				230						235				240	
Asn	Ala	Glu	Ile	Trp	Val	Lys	Tyr	Asn	Gln	Phe	Arg	Arg	Asp	Met	Thr
				245				250						255	
Leu	Met	Val	Leu	Asp	Leu	Val	Ala	Leu	Phe	Gln	Ser	Tyr	Asp	Thr	His
		260						265				270			
Met	Tyr	Pro	Ile	Lys	Thr	Thr	Ala	Gln	Leu	Thr	Arg	Glu	Val	Tyr	Thr
		275				280						285			
Asn	Ala	Ile	Gly	Thr	Val	His	Pro	His	Pro	Ser	Phe	Thr	Ser	Thr	Thr
290						295				300					
Trp	Tyr	Asn	Asn	Asn	Ala	Pro	Ser	Phe	Ser	Ala	Ile	Glu	Ala	Ala	Val
305				310						315				320	
Ile	Arg	Ser	Pro	His	Leu	Leu	Asp	Phe	Leu	Glu	Gln	Val	Thr	Ile	Tyr
				325				330						335	

Ser Leu Leu Ser Arg Trp Ser Asn Thr Gln Tyr Met Asn Met Trp Gly
 340 345 350

Gly His Lys Leu Glu Phe Arg Thr Ile Gly Gly Thr Leu Asn Thr Ser
 355 360 365

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

Thr Ser Arg Asp Ile Tyr Arg Thr Glu Ser Leu Ala Gly Leu Asn Leu
 385 390 395 400

Phe Leu Thr Gln Pro Val Asn Gly Val Pro Arg Val Asp Phe His Trp
 405 410 415

Lys Phe Val Thr His Pro Ile Ala Ser Asp Asn Phe Tyr Tyr Pro Gly
 420 425 430

Tyr Ala Gly Ile Gly Thr Gln Leu Gln Asp Ser Glu Asn Glu Leu Pro
 435 440 445

Pro Glu Thr Thr Gly Gln Pro Asn Tyr Glu Ser Tyr Ser His Arg Leu
 450 455 460

Ser His Ile Gly Leu Ile Ser Ala Ser His Val Lys Ala Leu Val Tyr
 465 470 475 480

Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn Thr Ile His Ser Asp
 485 490 495

Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly
 500 505 510

Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg
 515 520 525

Arg Thr Ser Gly Gly Pro Phe Ala Phe Ser Asn Val Asn Leu Asp Trp
 530 535 540

Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr
 545 550 555 560

Asn Leu Arg Met Tyr Val Thr Ile Ala Gly Glu Arg Ile Phe Ala Gly

565
 Gln Phe Asn Lys Thr Met Asn Thr Gly Asp Pro Leu Thr Phe Gln Ser
 580 585 590
 Phe Ser Tyr Ala Thr Ile Asp Thr Ala Phe Thr Phe Pro Thr Lys Ala
 595 600 605
 Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val
 610 615 620
 Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Leu Glu Ala
 625 630 635 640
 Val Thr Asp Leu Glu Arg Ala Gln Lys Ala Val His Glu Leu Phe Thr
 645 650 655
 Ser Thr Asn Pro Gly Gly Leu Lys Thr Asp Val Ala Lys Asp His Tyr
 660 665 670
 Thr Asn Thr Leu Ser Lys Ser Val Gln Ser Val Phe Arg Cys Arg Cys
 675 680 685
 Ser Glu Arg Thr Arg Leu Tyr Arg Trp Gly Tyr Pro Ser Lys Lys Glu
 690 695 700
 Tyr Trp Tyr Ile Trp Gly Tyr Thr Ser Lys Tyr
 705 710 715
 <210> 45
 <211> 1094
 <212> PRT
 <213> Trình tự Nhân tạo
 <220>
 <223> Protein BT-0027 đột biến
 <400> 45
 Met Phe Leu Glu Gln Ile Glu Gln Leu Ile Asp Gln Arg Ile Glu Thr
 1 5 10 15
 Val Glu Arg Asn Arg Ala Ile Gln Thr Leu Ile Gly Leu Ser Asn Ser
 20 25 30
 Tyr Asp Val Tyr Ile Glu Ala Leu Lys Glu Trp Glu Asn Asn Pro Asp

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

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

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

His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp
 740 745 750

Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln
 755 760 765

Asp Gly His Ala Lys Ile Gly Asn Leu Glu Phe Leu Lys Glu Lys Leu
 770 775 780

Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Lys Ala Glu Lys Lys Trp
 785 790 795 800

Arg Asp Lys Arg Asp Lys Leu Glu Trp Glu Thr Asn Val Val Tyr Lys
 805 810 815

Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asp Ser Gln Tyr Ser
 820 825 830

Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys
 835 840 845

Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu Leu Thr Val Ile
 850 855 860

Pro Gly Val Asn Ala Ser Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe
 865 870 875 880

Thr Ala Tyr Ser Leu Tyr Gly Ala Arg Asn Val Ile Lys Asn Gly Asp
 885 890 895

Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Glu Val
 900 905 910

Gln Gln Ile His His Arg Ser Val Leu Val Val Pro Ser Trp Lys Thr
 915 920 925

Glu Val Ser Gln Glu Val Cys Val Cys Pro Gly Arg Gly Tyr Ile Leu
 930 935 940

Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Asn Val Thr Ile
 945 950 955 960

His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu

965 970 975
 Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr
 980 985 990
 Val Asn Gln Glu Glu Tyr Lys Gly Thr Cys Thr Ser Arg Asn Arg Gly
 995 1000 1005
 Tyr Asp Glu Ser Tyr Glu Ser Ser Ser Ser Glu Ser Ala Tyr Tyr
 1010 1015 1020
 Ala Ser Val Tyr Glu Glu Lys Gly Tyr Thr Asp Gly Arg Arg Glu
 1025 1030 1035
 Asn Leu Cys Glu Phe Asn Arg Gly Tyr Gly Asp Tyr Thr Ser Ile
 1040 1045 1050
 Pro Thr Ala Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr
 1055 1060 1065
 Asp Lys Val Trp Ile Glu Leu Gly Glu Thr Glu Gly Ala Phe Ile
 1070 1075 1080
 Leu Asp Ser Val Glu Leu Leu Leu Met Glu Glu
 1085 1090
 <210> 46
 <211> 1169
 <212> PRT
 <213> Trình tự Nhân tạo
 <220>
 <223> Protein BT-0029 đột biến
 <400> 46
 Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15
 Asn Pro Glu Ser Glu Ile Leu Asn Val Ala Ile Phe Ser Ser Glu Gln
 20 25 30
 Val Ala Glu Ile His Leu Lys Ile Thr Arg Leu Ile Leu Glu Asn Phe
 35 40 45
 Leu Pro Gly Gly Ser Phe Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly

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

Phe Leu Asn Arg Leu Ile Ile Tyr Thr Gly Glu Tyr Arg Gly Ile Arg
 290 295 300

His Trp Ala Gly His Glu Val Glu Ser Ser Arg Thr Gly Met Met Thr
 305 310 315 320

Asn Ile Arg Phe Pro Leu Tyr Gly Thr Ala Ala Thr Ala Glu Pro Thr
 325 330 335

Arg Phe Ile Thr Pro Ser Thr Phe Pro Gly Leu Asn Leu Phe Tyr Arg
 340 345 350

Thr Leu Ser Ala Pro Ile Phe Arg Asp Glu Pro Gly Ala Asn Ile Ile
 355 360 365

Ile Arg Tyr Arg Thr Ser Leu Val Glu Gly Val Gly Phe Ile Gln Pro
 370 375 380

Asn Asn Gly Glu Gln Leu Tyr Arg Val Arg Gly Thr Leu Asp Ser Leu
 385 390 395 400

Asp Gln Leu Pro Leu Glu Gly Glu Ser Ser Leu Thr Glu Tyr Ser His
 405 410 415

Arg Leu Cys His Val Arg Phe Ala Gln Ser Leu Arg Asn Ala Glu Pro
 420 425 430

Leu Asp Tyr Ala Arg Val Pro Met Phe Ser Trp Thr His Arg Ser Ala
 435 440 445

Thr Pro Thr Asn Thr Ile Asp Pro Asp Val Ile Thr Gln Ile Pro Leu
 450 455 460

Val Lys Ala Phe Asn Leu His Ser Gly Ala Thr Ile Val Lys Gly Pro
 465 470 475 480

Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Val Gly Ser Phe
 485 490 495

Gly Asp Met Arg Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg
 500 505 510

Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu Gln Phe Tyr Thr Asn

515	520	525
Ile Asn Gly Thr Thr Ile Asn Ile Gly Asn Phe Ser Ser Thr Met Asp 530 535 540		
Ser Gly Asp Asp Leu Gln Tyr Gly Arg Phe Arg Val Ala Gly Phe Thr 545 550 555 560		
Thr Pro Phe Thr Phe Ser Asp Ala Met Ser Thr Phe Thr Ile Gly Ala 565 570 575		
Phe Ser Phe Ser Ser Asn Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe 580 585 590		
Val Pro Ala Glu Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Lys Ala 595 600 605		
Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu 610 615 620		
Lys Thr Asp Val Thr Asp Tyr His Ile Asp Lys Val Ser Asn Leu Val 625 630 635 640		
Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser 645 650 655		
Glu Lys Val Lys His Ala Lys Arg Leu Cys Asp Glu Arg Asn Leu Leu 660 665 670		
Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp 675 680 685		
Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys 690 695 700		
Glu Asn Tyr Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr 705 710 715 720		
Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg 725 730 735		
Tyr Glu Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr 740 745 750		

Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr
755 760 765

Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly
770 775 780

Glu Pro Asn Arg Cys Ala Thr His Leu Glu Trp Asn Pro Asp Leu Asp
785 790 795 800

Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe
805 810 815

Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly
820 825 830

Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu
835 840 845

Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu
850 855 860

Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys
865 870 875 880

Leu Glu Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Lys Ser Val
885 890 895

Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr
900 905 910

Asn Ile Ala Ile Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg
915 920 925

Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala
930 935 940

Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr
945 950 955 960

Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser
965 970 975

Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His

```

          980                      985                      990
Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu
      995                      1000                      1005

Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala
      1010                      1015                      1020

Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
      1025                      1030                      1035

Glu Asp Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu
      1040                      1045                      1050

Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala
      1055                      1060                      1065

Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly
      1070                      1075                      1080

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr
      1085                      1090                      1095

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp
      1100                      1105                      1110

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

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr
      1130                      1135                      1140

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

Val Asp Ser Val Glu Leu Leu Ile Met Glu Glu
      1160                      1165

```

```

<210> 47
<211> 1169
<212> PRT
<213> Trình tự Nhân tạo

```

```

<220>
<223> Protein BT-0020 đột biến

```

<400> 47

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

Asn Pro Lys Ile Glu Met Leu Asp Ile Glu Gly Ile Ser Ser Arg Ser
 20 25 30

Arg Glu Gln Val Ala Glu Ile Ser Leu Gly Leu Thr Arg Phe Leu Leu
 35 40 45

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

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

Ile Glu Gln Leu Ile Asp Gln Arg Ile Glu Ala His Val Arg Asn Gln
 85 90 95

Ala Ile Ser Arg Leu Glu Gly Leu Gly Asp Ser Tyr Glu Val Tyr Ile
 100 105 110

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

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

Ala Ile Pro Ile Leu Arg Glu Gln Gly Phe Glu Ile Pro Leu Leu Thr
 145 150 155 160

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

Val Tyr Phe Gly Gln Arg Trp Gly Leu Asp Thr Ala Thr Val Asn Asn
 180 185 190

His Tyr Asn Arg Leu Ile Asn Leu Ile Asn Thr Tyr Ser Asp His Cys
 195 200 205

Ala Gln Trp Phe Asn Arg Gly Leu Asp Asn Phe Gly Val Val Thr Ala
 210 215 220

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

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

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

915	920	925
Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Gly		
930	935	940
Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr		
945	950	955
Asp Ala Arg Asn Val Ile Lys Asn Ser Asp Phe Asn Asn Gly Leu Ser		
965	970	975
Cys Trp Asn Val Lys Gly His Val Asp Ile Glu Glu Gln Asn Asn His		
980	985	990
Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Lys		
995	1000	1005
Val His Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala		
1010	1015	1020
Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile		
1025	1030	1035
Glu Asp His Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu Glu Asp		
1040	1045	1050
Glu Val Tyr Pro Asn Asn Thr Arg Thr Cys Asn Ala Tyr Pro Ala		
1055	1060	1065
Asp Gln Glu Gly Tyr Glu Gly Ala Cys Thr Ser Arg Asn Arg Gly		
1070	1075	1080
Tyr Asp Glu Val Tyr Gly Asn Thr Pro Ser Leu Pro Ala Asp Tyr		
1085	1090	1095
Ala Pro Ile Tyr Glu Glu Asn Ala Tyr Thr Asp Gly Arg Arg Gly		
1100	1105	1110
Asn Pro Cys Glu Ser Ser Arg Gly Tyr Gly Asp Tyr Thr Pro Leu		
1115	1120	1125
Pro Ala Gly Tyr Glu Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr		
1130	1135	1140

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

Val Asp Ser Val Glu Leu Leu Ile Met Glu Glu
 1160 1165

<210> 48
 <211> 719
 <212> PRT
 <213> Trình tự Nhân tạo

<220>
 <223> Protein BT-0031 đột biến

<400> 48

Met Lys Leu Lys Asn Pro Asp Lys His Gln Ser Leu Ser Ser Asn Ala
 1 5 10 15

Lys Val Asp Lys Ile Ala Thr Asp Ser Leu Lys Asn Glu Thr Asp Ile
 20 25 30

Glu Leu Lys Asn Met Asn Asn Glu Asp Tyr Leu Arg Met Ser Glu His
 35 40 45

Glu Ser Ile Asp Pro Phe Val Ser Ala Ser Thr Ile Gln Thr Gly Ile
 50 55 60

Gly Ile Ala Gly Lys Ile Leu Gly Thr Leu Gly Val Pro Phe Ala Gly
 65 70 75 80

Gln Ile Ala Ser Leu Tyr Ser Phe Ile Leu Gly Glu Leu Trp Pro Lys
 85 90 95

Gly Lys Ser Gln Trp Glu Ile Phe Met Glu His Val Glu Glu Ile Ile
 100 105 110

Asn Gln Lys Ile Leu Thr Tyr Ala Arg Asn Lys Ala Leu Ser Asp Leu
 115 120 125

Arg Gly Leu Gly Asp Ala Leu Ala Val Tyr His Glu Ser Leu Glu Ser
 130 135 140

Trp Val Lys Asn Arg Asn Asn Thr Arg Ala Arg Ser Val Val Lys Asn
 145 150 155 160

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

385		390		395		400									
Thr	Glu	Ser	Leu	Ala	Gly	Leu	Asn	Leu	Phe	Leu	Thr	Gln	Pro	Val	Asn
				405					410					415	
Gly	Val	Pro	Arg	Val	Asp	Phe	His	Trp	Lys	Phe	Val	Thr	His	Pro	Ile
			420					425					430		
Ala	Ser	Asp	Asn	Phe	Tyr	Tyr	Pro	Gly	Tyr	Ala	Gly	Ile	Gly	Thr	Gln
		435					440					445			
Leu	Gln	Asp	Ser	Glu	Asn	Glu	Leu	Pro	Pro	Glu	Thr	Thr	Gly	Gln	Pro
	450					455					460				
Asn	Tyr	Glu	Ser	Tyr	Ser	His	Arg	Leu	Ser	His	Ile	Gly	Leu	Ile	Ser
465					470					475					480
Ala	Ser	His	Val	Lys	Ala	Leu	Val	Tyr	Ser	Trp	Thr	His	Arg	Ser	Ala
				485					490					495	
Asp	Arg	Thr	Asn	Thr	Ile	Asn	Ser	Asp	Ser	Ile	Thr	Gln	Ile	Pro	Leu
			500					505					510		
Val	Lys	Ala	Phe	Asn	Leu	Pro	Ser	Gly	Ala	Ser	Val	Val	Arg	Gly	Pro
		515					520					525			
Gly	Phe	Thr	Gly	Gly	Asp	Ile	Leu	Arg	Arg	Thr	Asn	Thr	Gly	Thr	Phe
	530					535					540				
Gly	Asp	Ile	Arg	Val	Asn	Ile	Asn	Pro	Pro	Phe	Ala	Gln	Arg	Tyr	Arg
545					550					555					560
Leu	Arg	Ile	Arg	Tyr	Ala	Ser	Thr	Thr	Asn	Leu	Glu	Phe	His	Thr	Ser
				565					570					575	
Ile	Asn	Gly	Lys	Ala	Ile	Asn	Gln	Gly	Asn	Phe	Ser	Ala	Thr	Met	Asn
			580					585					590		
Arg	Gly	Glu	Asp	Leu	Asp	Tyr	Lys	Ala	Phe	Arg	Thr	Val	Gly	Phe	Thr
		595					600					605			
Thr	Pro	Phe	Ser	Phe	Ser	Asn	Ala	Gln	Ser	Thr	Phe	Thr	Ile	Gly	Ala
	610					615					620				

Trp Asn Phe Ser Leu Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe
625 630 635 640

Val Pro Val Glu Val Thr Tyr Glu Ala Glu His Asp Phe Glu Lys Ala
645 650 655

Gln Glu Glu Val Thr Ala Leu Phe Thr Ser Thr Asn Pro Gly Gly Leu
660 665 670

Lys Thr Asn Val Thr Glu Tyr His Ile Asp Gln Val Ser Asn Leu Val
675 680 685

Glu Ser Leu Ser Asn Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Phe
690 695 700

Glu Leu Val Lys Tyr Ala Lys Gln Leu His Leu Gly Arg Asn Met
705 710 715

<210> 49

<211> 1164

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Protein BT-0201 đột biến

<400> 49

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

Asn Pro Glu Val Gly Ile Leu Asp Ile Glu Asn Phe Asn Leu Glu Leu
20 25 30

Val Ser Gln Val Ser Val Gly Leu Thr Arg Phe Leu Leu Glu Ala Ser
35 40 45

Ile Pro Gly Ala Gly Phe Ala Leu Gly Leu Phe Asp Ile Ile Trp Gly
50 55 60

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

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

Ala Leu Ser Gly Leu Ser Ser Ser Tyr Glu Val Tyr Ile Glu Ala Leu
 100 105 110

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

Val Ala Arg Phe Arg Ala Thr Asp Asn Ser Leu Ile Thr Asp Ile Pro
 130 135 140

Leu Leu Glu Ile Pro Gly Phe Glu Ile Ala Thr Leu Ser Val Tyr Thr
 145 150 155 160

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

Gly Glu Arg Trp Gly Leu Thr Gln Thr Asn Ile Glu Asp Leu His Thr
 180 185 190

Arg Leu Thr Arg Tyr Ile Gln Glu Tyr Ser Asp His Cys Ala Arg Trp
 195 200 205

Tyr Asn Gln Gly Leu Asn Asn Ile Gly Gly Ile Asn Thr Arg Tyr Leu
 210 215 220

Asp Phe Gln Arg Glu Leu Thr Ile Ser Val Leu Asp Ile Val Ala Leu
 225 230 235 240

Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile Pro Thr Gln Ser Gln
 245 250 255

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

Trp Ile Leu Ser Ile Ser Asn Val Leu Arg Ala Pro His Leu Met Asp
 275 280 285

Phe Phe Asp Arg Ile Ile Ile Tyr Thr Gly Thr Val Arg Ser Thr Pro
 290 295 300

His Trp Glu Gly His Glu Val Ile Ser Arg Arg Thr Gly Gln Gly Asn
 305 310 315 320

Glu Ile Arg Ser Pro Leu Tyr Gly Val Ala Ala Asn Ala Glu Pro Pro

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

Ser Asn Ala Gln Ser Thr Phe Thr Leu Gly Thr Gln Ala Phe Ser Asn
 565 570 575

Gln Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val Thr
 580 585 590

Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala
 595 600 605

Leu Phe Thr Ser Thr Ser Gln Leu Gly Leu Lys Thr Asn Val Thr Gly
 610 615 620

Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser Asp Glu
 625 630 635 640

Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala
 645 650 655

Lys Arg Leu Ser Asp Lys Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg
 660 665 670

Gly Ile Asn Arg Gln Pro Asp His Gly Trp Arg Gly Ser Thr Asp Ile
 675 680 685

Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu
 690 695 700

Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile
 705 710 715 720

Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr
 725 730 735

Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ser
 740 745 750

Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu
 755 760 765

Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg Cys Ala
 770 775 780

Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp Gly

```

785                               790                               795                               800

Glu Lys Cys Val His His Ser His His Phe Ser Leu Asp Ile Asp Val
                        805                               810                               815

Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Leu Ile Phe Lys
                        820                               825                               830

Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu
                        835                               840                               845

Glu Glu Glu Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Arg Ala
                        850                               855                               860

Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr Asn
865                               870                               875                               880

Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn
                        885                               890                               895

Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His
                        900                               905                               910

Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu
                        915                               920                               925

Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu
                        930                               935                               940

Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr Asp Ala Arg Asn Val Ile
945                               950                               955                               960

Lys Asn Gly Asn Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys Gly
                        965                               970                               975

His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Ile
                        980                               985                               990

Pro Glu Trp Glu Ala Glu Val Ser Gln Lys Val Arg Val Cys Pro Gly
                        995                               1000                               1005

Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
1010                               1015                               1020

```

Glu Gly Cys Val Thr Ile His Glu Ile Glu Asp Asn Thr Asp Glu
1025 1030 1035

Leu Lys Phe Ser Asn Cys Val Glu Glu Gly Tyr Pro Asn Asn Thr
1040 1045 1050

Val Thr Cys Asn Glu Tyr Thr Met Asn Gln Gly Val Gly Glu Cys
1055 1060 1065

Thr Asp Ala Cys Asn Val Arg Asn Arg Gly Tyr Glu Asp Ala Tyr
1070 1075 1080

Gly His Asn Pro Ser Thr Pro Val His Tyr Thr Thr Pro Tyr Glu
1085 1090 1095

Glu Glu Thr Tyr Thr Asp Glu Arg Arg Glu Asn Pro Cys Glu Ala
1100 1105 1110

Asn Lys Gly Tyr Val Asn Tyr Thr Pro Leu Pro Val Gly Tyr Val
1115 1120 1125

Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Thr Val Trp Ile
1130 1135 1140

Glu Ile Gly Glu Thr Glu Gly Thr Phe Leu Val Asp Ser Val Glu
1145 1150 1155

Leu Leu Ile Met Glu Glu
1160

<210> 50

<211> 1166

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> Protein BT-0202 đột biến

<400> 50

Met Lys Ile Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

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

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

				260				265				270			
Glu	Phe	Tyr	Tyr	Gly	Pro	Ile	Tyr	Asn	Tyr	Asn	Ile	Val	Gly	Arg	Leu
		275				280						285			
Thr	Glu	Gln	Gln	Leu	Arg	Arg	Pro	His	Leu	Met	Asp	Phe	Phe	Asn	Ser
		290				295				300					
Met	Ile	Met	Tyr	Thr	Ser	Asp	Asn	Arg	Arg	Glu	His	Tyr	Trp	Ser	Gly
305				310						315				320	
Leu	Glu	Met	Lys	Ala	Thr	Asp	Thr	Ser	Gly	Asn	Gln	Val	Ser	Phe	Pro
				325				330						335	
Leu	Ala	Gly	Thr	Arg	Gly	Asn	Ser	Ala	Pro	Pro	Val	Thr	Val	Arg	Asn
		340				345						350			
Asn	Gly	Glu	Gly	Val	Tyr	Arg	Ile	Leu	Ser	Glu	Pro	Phe	Tyr	Ser	Ser
		355				360						365			
Pro	Phe	Leu	Gly	Thr	Ser	Val	Leu	Gly	Ser	Arg	Gly	Glu	Glu	Phe	Ala
		370				375				380					
Phe	Ala	Ser	Asn	Thr	Thr	Thr	Ser	Leu	Pro	Ser	Thr	Ile	Tyr	Arg	Asn
385				390						395				400	
Arg	Gly	Thr	Val	Asp	Ser	Leu	Val	Ser	Ile	Pro	Pro	Gln	Asp	Tyr	Ser
				405				410						415	
Val	Pro	Pro	His	Arg	Gly	Tyr	Ser	His	Leu	Leu	Ser	His	Val	Thr	Met
		420				425						430			
His	Asn	Ser	Ser	Pro	Ile	Phe	His	Trp	Thr	His	Arg	Ser	Ala	Thr	Pro
		435				440						445			
Arg	Asn	Thr	Ile	Asp	Pro	Asp	Ser	Ile	Thr	Gln	Ile	Pro	Val	Val	Lys
		450				455				460					
Ala	Ser	His	Leu	Ser	Gly	Gly	Ser	Val	Ile	Lys	Gly	Pro	Gly	His	Thr
465				470						475				480	
Gly	Gly	Asp	Leu	Ile	Ser	Leu	Pro	Val	Asn	Asn	Phe	Thr	His	Phe	Arg
				485				490						495	

Ile Pro Phe Gln Ala Asn Thr Pro Gln Arg Tyr Arg Ile Arg Ile Arg
 500 505 510

Tyr Ala Ala Asp Ser Asp Gly Thr Leu Asp Ser Gly Val Phe Leu Ser
 515 520 525

Ala Ala Ala Gly Asp Gly Phe Asn Thr Thr Ser Tyr Arg Ala Thr Met
 530 535 540

Ser Pro Gly Gly Ser Leu Thr Ser Arg Asp Phe Gln Phe Leu Asp Leu
 545 550 555 560

Asn Thr Ser Phe Thr Ser Asp Val Ala Ser Asn Leu Trp Leu His Phe
 565 570 575

Ile Arg Tyr Ile Arg Pro Gly Asn Leu Tyr Ile Asp Arg Ala Glu Phe
 580 585 590

Ile Pro Val Asp Ala Thr Phe Glu Ala Gly Tyr Asn Leu Glu Arg Ala
 595 600 605

Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Lys Gly Leu
 610 615 620

Gln Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val
 625 630 635 640

Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser
 645 650 655

Glu Lys Val Lys Gln Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu
 660 665 670

Gln Asp Ser Asn Phe Arg Gly Ile Asn Arg Glu Gln Asp Arg Gly Trp
 675 680 685

Arg Gly Ser Thr His Ile Thr Ile Gln Gly Gly Asn Asp Val Phe Lys
 690 695 700

Glu Asn Phe Val Thr Leu Pro Gly Ala Phe Asp Ala Cys Tyr Pro Thr
 705 710 715 720

Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg

725					730					735					
Tyr	Glu	Leu	Arg	Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Asp	Ile	Tyr
			740					745					750		
Leu	Ile	Arg	Tyr	Asn	Thr	Lys	His	Glu	Thr	Leu	Asn	Val	Pro	Gly	Thr
		755					760					765			
Lys	Ser	Pro	Trp	Ser	Leu	Cys	Thr	Glu	Ser	Pro	Leu	Gly	Lys	Cys	Gly
	770					775					780				
Glu	Pro	Asn	Arg	Cys	Ala	Ser	Gln	Ile	Glu	Trp	Asn	Pro	Asp	Leu	Asp
785					790					795					800
Cys	Ser	Cys	Arg	Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe
				805					810					815	
Ser	Leu	Asp	Ile	Asp	Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asn	Leu	Gly
			820					825					830		
Ile	Trp	Val	Ile	Phe	Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu
		835					840					845			
Gly	Asn	Leu	Glu	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Leu	Gly	Glu	Ala	Leu
	850					855					860				
Ala	Arg	Val	Lys	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys
865					870					875					880
Leu	Gln	Ser	Glu	Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ala	Val
				885					890					895	
Asp	Gly	Leu	Phe	Val	Asp	Ser	Gln	Tyr	Glu	Arg	Leu	Gln	Ala	Asp	Thr
			900					905					910		
Asn	Ile	Ala	Met	Ile	His	Ala	Ala	Asp	Lys	His	Val	His	Gln	Ile	Arg
		915					920					925			
Glu	Val	Tyr	Phe	Pro	Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala
	930					935					940				
Ile	Phe	Glu	Glu	Leu	Glu	Gly	Arg	Ile	Phe	Thr	Ala	Tyr	Ser	Leu	Tyr
945					950					955					960

Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser
 965 970 975

Cys Trp Asn Val Lys Gly His Val Asp Val Glu Val Gln Asn Asn His
 980 985 990

Arg Ser Val Leu Val Ile Ala Glu Trp Glu Ala Glu Val Ser Gln Glu
 995 1000 1005

Val Pro Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala
 1010 1015 1020

Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
 1025 1030 1035

Glu Asp His Thr Asp Glu Leu Lys Phe Arg Asn Cys Glu Glu Glu
 1040 1045 1050

Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Gly Tyr Thr Ala
 1055 1060 1065

Thr Gln Glu Glu Tyr Lys Asp Ala Tyr Thr Ser Arg Asn Arg Gly
 1070 1075 1080

Tyr Asp Glu Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr
 1085 1090 1095

Ala Ser Val Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Glu
 1100 1105 1110

Asn Pro Cys Glu Met Glu Arg Gly Tyr Thr Pro Leu Pro Val Gly
 1115 1120 1125

Tyr Ile Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Thr Val
 1130 1135 1140

Trp Ile Glu Leu Gly Glu Thr Glu Gly Thr Phe Leu Ile Asp Ser
 1145 1150 1155

Val Glu Leu Leu Leu Met Glu Glu
 1160 1165