



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

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

(11)



1-0035801

(51)⁷ **A01H 5/00**; C07K 14/325; A61K 35/742; C07K 14/32; A01N 63/02; A61K 35/66

(13) **B**

(21) 1-2018-04061 (22) 07/02/2017
(86) PCT/US2017/016826 07/02/2017 (87) WO 2017/146899 31/08/2017
(30) 201610105880.8 26/02/2016 CN
(45) 25/05/2023 422 (43) 25/02/2019 371A
(73) SYNGENTA PARTICIPATIONS AG (CH)
Schwarzwaldallee 215, 4058 Basel (CH)
(72) QIANQIAN, Duan (CN); CHUNPING, Luo (CN); CHANGMING, Ning (CN);
XIAOJUN, Wang (CN); FENSHAN, Wang (CN); WENQIAN, Mei (CN).
(74) Công ty TNHH Ban Ca (BANCA)

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến protein diệt loài gây hại và phân tử axit nucleic mã hóa cho chúng, cũng như hợp phần và phương pháp để kiểm soát loài gây hại cây trồng.

Tình trạng kỹ thuật của sáng chế

Bacillus thuringiensis (Bt) là vi khuẩn đất tạo bào tử gram dương đặc trưng bởi khả năng tạo ra thể vùi tinh thể mà gây độc đặc hiệu đối với các bộ và các loài nhất định của các loài gây hại thực vật, bao gồm côn trùng, nhưng vô hại đối với thực vật và các sinh vật không phải là đích khác. Vì lý do này, hợp phần 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 côn trùng gây hại nông nghiệp hoặc vật truyền côn trùng của nhiều bệnh ở người hoặc động vật khác nhau.

Protein tinh thể (Cry) từ *Bacillus thuringiensis* có hoạt tính diệt côn trùng hiệu nghiệm chống lại chủ yếu là côn trùng gây hại là loài gây hại thuộc bộ cánh vảy, bộ hai cánh, và bộ cánh cứng. Các protein này cũng thể hiện hoạt tính chống lại loài gây hại thuộc các Bộ Hymenoptera, Homoptera, Phthiraptera, Mallophaga, và Acari, cũng như là các bộ không xương sống khác chẳng hạn 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 từ 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 delta” được sử dụng thay thế lẫn nhau với thuật ngữ

“protein Cry.” Danh pháp hiện nay đối với protein và gen Cry dựa trên độ tương đồng trình tự axit amin chứ không phải là độ đặc hiệu đích côn trùng (Crickmore *et al.* (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 *et al.* (2010, Bioeng. Bugs, 1:31-50), độc tố Cry có thể vẫn được chia thành sáu lớp chính theo độ đặc hiệu vật chủ côn trùng của chúng và bao gồm: nhóm 1—loài gây hại thuộc bộ cánh vảy ví dụ như, Cry1, Cry9 và Cry15); nhóm 2—bộ cánh vảy và bộ hai cánh (ví dụ như, 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). Các độc tố Cry1I, Cry2, Cry3, Cry10 và Cry11 (73–82kDa) là duy nhất bởi vì chúng dường như là dạng cắt tự nhiên của các protein Cry1 và Cry4 lớn hơn (130–140kDa).

Protein Cry là phân tử protein hình cầu mà tích lũy dưới dạng tiền độc tố ở dạng kết tinh trong giai đoạn hình thành bào tử của *Bt*. Sau khi bị loài gây hại ăn, các tinh thể này thường được hòa tan để giải phóng tiền độc tố, mà có kích thước nằm trong khoảng, ví dụ, từ 130-140kDa đối với nhiều protein Cry có hoạt tính đối với bộ cánh vảy, chẳng hạn như Cry1 và Cry9, và 60-80 kDa đối với protein Cry3 có hoạt tính đối với bộ cánh cứng và protein Cry2 có hoạt tính đối với bộ cánh vảy/bộ hai cánh. Sau khi các tinh thể này được hòa tan bởi côn trùng nhạy cảm tiền độc tố đã được giải phóng được xử lý bằng proteaza trong ruột côn trùng, ví dụ trypsin và chymotrypsin, để tạo ra độc tố protein Cry lõi kháng proteaza. Sự xử lý phân giải protein này gồm có sự loại bỏ axit amin từ các vùng khác nhau của các tiền độc tố Cry khác nhau. Ví dụ, tiền độc tố Cry nặng 130-140kDa thường được hoạt hóa thông qua sự loại bỏ bằng cách phân giải protein của peptit đầu tận cùng N có 25-30 axit amin và xấp xỉ một nửa protein còn lại từ đầu tận cùng C tạo ra độc tố Cry trưởng thành nặng xấp xỉ 60-70kDa. Tiền độc tố nặng 60-80kDa, ví dụ như Cry2 và Cry3, cũng được xử lý nhưng không đến cùng mức độ như tiền độc tố lớn hơn. Tiền độc tố nhỏ hơn thường có số lượng axit amin được loại bỏ từ đầu tận cùng N bằng

hoặc lớn hơn so với tiền độc tố lớn hơn nhưng có số lượng axit amin được loại bỏ từ đầu tận cùng C ít hơn. Ví dụ, sự hoạt hóa bằng cách phân giải protein của các thành viên họ Cry2 thường gồm có sự loại bỏ xấp xỉ 40-50 axit amin ở đầu tận cùng N. Nhiều protein Cry khá độc đối với côn trùng đích đặc hiệu, nhưng nhiều protein Cry có phổ hoạt tính hẹp.

Protein Cry thường có năm miền trình tự bảo toàn, và ba miền cấu trúc bảo toàn (xem, ví dụ, de Maagd *et al.* (2001) Trends Genetics 17:193-199). Miền cấu trúc bảo toàn thứ nhất, gọi là miền I, thường gồm có bảy cấu trúc xoắn alpha và tham gia vào sự cài xen màng và tạo lỗ rỗng. Miền II thường gồm có ba tấm beta được bố trí trong cấu hình chìa khóa Greek, và miền III thường gồm có hai tấm beta đối song trong cấu tạo 'cuộn lỏng' (de Maagd *et al.*, 2001, ở trên). Các miền II và III tham gia vào sự nhận diện thụ thể và liên kết, và do đó được coi là các phần quyết định tính đặc hiệu độc tố.

Nhiều cây trồng có giá trị thương mại, bao gồm các cây nông nghiệp thông thường, nhạy cảm đối với sự tấn công bởi các loài gây hại cây trồng bao gồm các côn trùng và giun tròn gây hại, 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 15 đến 20 phần trăm lượng hạt thu hoạch được ở Trung Quốc bị thiệt hại mỗi năm do bệnh và các loài gây hại côn trùng. Ngoài ra, khoảng \$8 triệu bị thiệt hại mỗi năm chỉ riêng ở Mỹ do sự phá hoại của loài gây hại không xương sống bao gồm côn trùng. Côn trùng gây hại cũng là gánh nặng đối với người trồng rau và cây ăn quả, đối với những người sản xuất hoa trang trí, và đối với người làm vườn tại nhà.

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 trồng chuyển gen thể hiện là tạo ra công cụ khác để kiểm soát các loài côn trùng gây hại 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 côn trùng gây hại phát triển khả năng kháng các phương pháp kiểm soát này. Điều này đã được làm giảm bớt một phần bởi các hoạt động quản lý tính kháng khác nhau, nhưng vẫn có nhu cầu phát triển tác nhân kiểm soát loài gây hại mới và hữu hiệu mà mang lại lợi ích kinh tế cho nông dân và chấp nhận được về mặt môi trường. Đặc biệt có nhu cầu là các tác nhân kiểm soát mà có thể nhắm đích vào phổ rộng hơn của các côn trùng gây hại quan trọng về mặt kinh tế và kiểm soát một cách hữu hiệu các chủng côn trùng mà kháng hoặc có thể trở nên kháng với các tác nhân kiểm soát côn trùng hiện có.

Bản chất kỹ thuật của sáng chế

Xem xét đến các nhu cầu này, mục đích của sáng chế là đề xuất tác nhân kiểm soát loài gây hại mới bằng cách đề xuất thể phân lập *Bacillus thuringiensis* (Bt) mới cũng như là gen và protein diệt loài gây hại mới mà có thể được sử dụng để kiểm soát nhiều loài gây hại thực vật.

Sáng chế đề xuất các hợp phần và phương pháp tạo ra hoạt tính diệt loài gây hại cho vi khuẩn, cây, tế bào, mô và hạt của cây. Cụ thể là, gen khảm có chứa polynucleotit mới mà mã hóa cho protein Cry được phân lập từ Bt và các trình tự về cơ bản giống chúng, mà sự biểu hiện của chúng tạo ra protein có độc tính đối với các loài côn trùng gây hại quan trọng về mặt kinh tế, cụ thể là côn trùng gây hại mà phá hoại cây trồng, được đề xuất. Sáng chế còn đề cập đến protein Cry mới thu được từ sự biểu hiện của polynucleotit, và đề cập đến hợp phần và chế phẩm có chứa protein Cry, mà gây độc đối với côn trùng bằng cách ức chế khả năng của côn trùng gây hại để sống sót, sinh trưởng và sinh sản, hoặc hạn chế sự tổn thất hoặc thiệt hại liên quan đến côn trùng đối với cây trồng. Protein Cry theo sáng chế bao gồm protein Cry tự nhiên và protein Cry đột biến hoặc biến thể mà có một hoặc nhiều sự thế, sự thêm hoặc sự bớt axit amin. Ví dụ về protein Cry đột biến

bao gồm mà không giới hạn ở các protein Cry mà được gây đột biến để có phổ hoạt tính rộng hơn hoặc hoạt tính đặc hiệu cao hơn so với các đối tác protein Cry tự nhiên của chúng, các protein Cry này được gây đột biến để đưa vào epitop để tạo ra kháng thể mà nhận diện phân biệt protein đột biến từ protein tự nhiên hoặc protein Cry đột biến để điều biến sự biểu hiện ở sinh vật chuyển gen. Protein Cry mới theo sáng chế có độ độc cao đối với côn trùng gây hại. Ví dụ, protein Cry theo sáng chế có thể được dùng để kiểm soát một hoặc nhiều côn trùng gây hại quan trọng về mặt kinh tế chẳng hạn như sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu đục quả bông (*Helicoverpa armigera*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), sâu bướm Phương Đông (*Mythimna sepatate*), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu bướm hại đậu nhung (*Anticarsia gemmatilis*), sâu đo đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu ngài đêm hại đậu phương tây (*Richia albicosta*), sâu ăn chồi thuốc lá (*Heliothis virescens*), sâu đục thân có sọc (*Chilo suppressalis*), sâu đục thân màu hồng (*Sesamia calamistis*), sâu cuốn lá lúa (*Cnaphalocrocis medinalis*), và dạng tương tự.

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

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

Sáng chế còn đề cập đến chủng *Bacillus thuringiensis* (Bt) được phân lập mà sản xuất 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 vector 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 đã được thiết kế để biểu hiện ở sinh vật chẳng hạn như thực vật hoặc vi khuẩn hoặc để tạo ra độc tố Cry lai có hoạt tính diệt loài gây hại được tăng cường. Sáng chế còn đề cập đến phương pháp tạo ra protein Cry và phương pháp sử dụng các trình tự polynucleotit và protein Cry, ví dụ trong vi sinh vật để kiểm soát côn trùng hoặc trong cây trồng chuyển gen để mang lại sự bảo vệ khỏi sự gây hại của côn trùng.

Khía cạnh khác của sáng chế bao gồm hợp phần diệt côn trùng và chế phẩm có chứa protein Cry hoặc chủng *Bacillus thuringiensis* theo sáng chế, và phương pháp sử dụng hợp phần hoặc chế phẩm này để kiểm soát quần thể côn trùng, ví dụ bằng cách áp dụng hợp phần hoặc chế phẩm này cho diện tích bị phá hoại bởi côn trùng, hoặc để xử lý phòng ngừa cho diện tích dễ bị tấn công bởi côn trùng hoặc cây trồng để mang lại sự bảo vệ chống lại côn trùng gây hại. Một cách tùy ý, hợp phần hoặc chế phẩm theo sáng chế có thể, ngoài protein Cry hoặc chủng *Bt* theo sáng chế, có chứa tác nhân diệt loài gây hại khác chẳng hạn như chất diệt loài gây hại hóa học để gia tăng hoặc tăng cường khả năng kiểm soát côn trùng của hợp phần hoặc chế phẩm.

Hợp phần và phương pháp theo sáng chế hữu dụng để kiểm soát côn trùng gây hại mà tấn công thực vật, cụ thể là cây trồng. Hợp phần theo sáng chế cũng hữu dụng để tạo ra protein Cry được thay đổi hoặc được cải thiện mà có hoạt tính diệt loài gây hại, hoặc để phát hiện sự có mặt của protein Cry hoặc axit nucleic trong sản phẩm thương mại 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 hình vẽ

Hình 1 là sự sắp xếp thẳng hàng của các trình tự axit amin protein Cry BT-0049 và BT-0211.

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

SEQ ID NO:1 thể hiện trình tự nucleotit mã hóa cho protein BT-0049.

SEQ ID NO:2 thể hiện trình tự nucleotit mã hóa cho protein BT-0063.

SEQ ID NO:3 thể hiện trình tự nucleotit mã hóa cho protein BT-0064.

SEQ ID NO:4 thể hiện trình tự nucleotit mã hóa cho protein BT-0086.

SEQ ID NO:5 thể hiện trình tự nucleotit mã hóa cho protein BT-0498.

SEQ ID NO:6 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0049.

SEQ ID NO:7 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0063.

SEQ ID NO:8 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0064.

SEQ ID NO:9 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0086.

SEQ ID NO:10 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0498.

SEQ ID NO:11 thể hiện trình tự nucleotit mã hóa protein BT-0049 đột biến.

SEQ ID NO:12 thể hiện trình tự nucleotit mã hóa protein BT-0063 đột biến.

SEQ ID NO:13 thể hiện trình tự nucleotit mã hóa protein BT-0064 đột biến.

SEQ ID NO:14 thể hiện trình tự nucleotit mã hóa protein BT-0086 đột biến.

SEQ ID NO:15 thể hiện trình tự nucleotit mã hóa protein BT-0498 đột biến.

SEQ ID NO:16 thể hiện trình tự axit amin của protein BT-0049.

SEQ ID NO:17 thể hiện trình tự axit amin của protein BT-0063.

SEQ ID NO:18 thể hiện trình tự axit amin của protein BT-0064.

SEQ ID NO:19 thể hiện trình tự axit amin của protein BT-0086.

SEQ ID NO:20 thể hiện trình tự axit amin của protein BT-0498.

SEQ ID NO:21 thể hiện trình tự axit amin của protein BT-0049 đột biến.

SEQ ID NO:22 thể hiện trình tự axit amin của protein BT-0063 đột biến.

SEQ ID NO:23 thể hiện trình tự axit amin của protein BT-0064 đột biến.

SEQ ID NO:24 thể hiện trình tự axit amin của protein BT-0086 đột biến.

SEQ ID NO:25 thể hiện trình tự axit amin của protein BT-0498 đột biến.

SEQ ID NO:26 thể hiện trình tự axit amin của protein BT-0211.

SEQ ID NO:27 thể hiện trình tự axit amin duy nhất đối với protein giống BT-0049.

SQE ID NO:28 thể hiện trình tự axit amin duy nhất đối với protein giống BT-0049.

Mô tả chi tiết sáng chế

Phần mô tả này không nhằm làm danh mục chi tiết của tất cả các cách khác nhau để có thể thực hiện sáng chế, hoặc tất cả các dấu hiệu có thể được thêm vào sáng chế này. 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 dự tính rằng theo một số phương án của sáng chế, dấu hiệu bất kỳ hoặc dạng 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 người có hiểu biết trung bình trong lĩnh vực của 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 tường tận 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 người có hiểu biết trung bình trong lĩnh vực mà sáng chế thuộc về. Thuật ngữ được sử dụng trong phần mô tả của sáng chế ở đây chỉ nhằm mục đích mô tả các phương án cụ thể và không nhằm làm giới hạn sáng chế.

Định nghĩa

Như được sử dụng ở đây và ở yêu cầu bảo hộ kèm theo, các mạo từ số ít bao gồm việc đề cập đến số nhiều trừ khi ngữ cảnh rõ ràng chỉ ra nghĩa 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 người có hiểu biết trung bình trong lĩnh vực này, 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 dạng 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 trong khoảng. 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 của giá trị bằng số được cho. Thông thường, thuật ngữ "khoảng" được sử dụng ở đây để thay đổi giá trị bằng số ở 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 khối lượng phân tử) giá trị chính xác (tức là, không có "khoảng") được ưu tiên.

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 khuôn. 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, Cangene, Mississauga, Ontario), hệ lắp Q-Beta Replicaza, hệ khuếch đại trên cơ sở phiên mã (TAS), và khuếch đại xáo trộn sợi (SDA). Xem tài liệu, ví dụ, Diagnostic Molecular Microbiology: Principles and Applications, PERSING et al., Ed., American Society for Microbiology, Washington, D.C. (1993). Sản phẩm 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 dùng để chỉ cấu trúc hoặc phân tử bao gồm hai hoặc nhiều polynucleotit có nguồn gốc khác nhau được ghép thành phân tử axit nucleic đơn lẻ. 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 hòa 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 promotor 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 hòa và polynucleotit mã hóa mà có nguồn gốc từ các nguồn khác nhau, hoặc chứa polynucleotit điều hòa 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 cat-xet biểu hiện chứa polynucleotit theo sáng chế dưới sự kiểm soát của polynucleotit điều hòa, cụ thể là dưới sự kiểm soát của polynucleotit điều hòa 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 trong sinh vật để tạo ra protein.

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

"Kiểm soát" côn trùng có nghĩa là ức chế, thông qua tác dụng gây độc, khả năng của côn trùng gây hại để sống sót, sinh trưởng, ăn, hoặc sinh sản, hoặc hạn chế sự tồn thất hoặc thiệt hại do côn trùng ở cây trồng hoặc bảo vệ tiềm năng sản lượng của cây trồng khi sinh trưởng trong sự có mặt của côn trùng gây hại. "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.

Các thuật ngữ "gồm" hoặc "bao gồm," khi được dùng trong bản mô tả này, xác định sự có mặt của dấu hiệu, số nguyên, bước, hoạt động, yếu tố, hoặc thành phần được chỉ ra, nhưng không loại trừ sự có mặt hoặc sự thêm vào của một hoặc nhiều dấu hiệu, số nguyên, bước, hoạt động, yếu tố, thành phần khác, hoặc nhóm của chúng.

Như được sử dụng ở đây, câu chuyển tiếp "về cơ bản 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 làm thay đổi cốt yếu (các) đặc trưng cơ bản và mới" của sáng chế được yêu cầu bảo hộ. Do đó, thuật ngữ "về cơ bản gồm có" khi được sử dụng trong yêu cầu bảo hộ của sáng chế không dự tính được hiểu là tương đương với "bao gồm."

Trong ngữ cảnh của sáng chế, "việc tương ứng với" hoặc "tương ứng với" có nghĩa là khi trình tự axit amin của các protein Cry biến thể hoặc tương đồng được sắp xếp thẳng hàng với nhau, các axit amin mà "tương ứng với" các vị trí được đánh số nhất định trong protein biến thể hoặc tương đồng là các axit amin mà thẳng hàng với các vị trí này trong protein tham chiếu nhưng không nhất thiết là ở các vị trí đánh số chính xác này tương quan với trình tự axit amin tham chiếu cụ thể theo sáng chế. Ví dụ, nếu SEQ ID NO:16 là trình tự tham chiếu và được sắp xếp thẳng hàng với SEQ ID NO:26, Arg415 của SEQ ID NO:26 "tương ứng với" Gly421 của SEQ ID NO:16.

Như được dùng trong bản mô tả này, thuật ngữ "protein Cry" có nghĩa là protein diệt côn trùng thuộc loại nội độc tố delta tinh thể *Bacillus thuringiensis*. Thuật ngữ "protein Cry" có thể dùng để chỉ dạng tiền độc tố hoặc mảnh hoặc độc tố có hoạt tính diệt côn trùng bất kỳ của chúng.

Việc "phân phối" hợp phần hoặc protein gây độc có nghĩa là hợp phần hoặc protein gây độc tiếp xúc với côn trùng, mà làm thuận lợi cho sự ăn bằng miệng của hợp phần hoặc protein gây độc, dẫn đến tác dụng gây độc và kiểm soát côn trùng. Hợp phần hoặc protein gây độc có thể được phân phối theo nhiều cách đã được công nhận, bao gồm nhưng không giới hạn ở, sự biểu hiện cây trồng chuyển gen, (các) hợp phần protein được tạo chế phẩm, (các) hợp phần protein có thể phun được, chất nền môi, hoặc hệ thống phân phối protein đã được công nhận trong lĩnh vực khác bất kỳ.

Thuật ngữ "miền" dùng để chỉ tập hợp các axit amin được bảo toàn ở các vị trí cụ thể dọc theo sự sắp xếp thẳng hàng của các trình tự của các protein có liên quan về mặt tiến hóa. Trong khi các axit amin ở các vị trí khác có thể thay đổi giữa các thể tương đồng, các axit amin được bảo toàn cao ở các vị trí cụ thể thể hiện 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 các trình tự được sắp xếp thẳng hàng của họ của các thể tương đồng protein, chúng có thể được sử dụng làm dấu hiệu nhận dạng để xác định liệu polypeptit đang xem xét bất kỳ có thuộc vào nhóm polypeptit đã được xác định từ trước hay không.

"Lượng hữu hiệu để kiểm soát côn trùng" có nghĩa là nồng độ của protein gây độc mà ức chế, thông qua tác dụng gây độc, khả năng của côn trùng để sống sót, sinh trưởng, ăn hoặc sinh sản, hoặc hạn chế sự tồn thất hoặc thiệt hại do côn trùng ở cây trồng hoặc bảo vệ tiềm năng sản lượng của cây trồng khi sinh trưởng trong sự có mặt của côn trùng gây hại. "Lượng hữu hiệu để kiểm soát côn trùng" có thể có hoặc không có nghĩa là giết côn trùng, mặc dù tốt hơn nên có nghĩa là giết côn trùng.

"Cat-xet biểu hiện" như được dùng trong bản mô tả này có nghĩa là phân tử axit nucleic có khả năng định hướng sự biểu hiện của ít nhất là một polynucleotit được quan tâm, chẳng hạn như polynucleotit mà mã hóa cho protein Cry theo sáng chế, trong tế bào chủ thích hợp, có chứa promotor được liên kết có điều khiển với polynucleotit được quan tâm mà được liên kết có điều khiển với tín hiệu kết thúc. "Cat-xet biểu hiện" thường còn chứa polynucleotit bổ sung cần để phiên mã đúng polynucleotit quan tâm. Cat-xet biểu hiện có thể còn chứa polynucleotit khác không cần thiết trong việc biểu hiện trực tiếp của polynucleotit quan tâm nhưng có mặt do vị trí giới hạn thích hợp để loại bỏ băng khỏi vật truyền biểu hiện. Cat-xet biểu hiện bao gồm (các) polynucleotit quan tâm 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 băng mà có trong tự nhiên nhưng có thể thu được ở dạng tái tổ hợp hữu dụng để biểu hiện khác loài. Tuy nhiên, thông thường, cat-xet biểu hiện là khác loại với vật chủ, tức là polynucleotit quan tâm trong cat-xet biểu hiện không có trong 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. Sự

biểu hiện của (các) polynucleotit quan tâm trong cat-xet biểu hiện thường là chịu sự kiểm soát của promotor. Trong trường hợp sinh vật đa bào, như cây trồng, promotor có thể còn là đặc hiệu hoặc ưu tiên cho mô, hoặc cơ quan, hoặc giai đoạn phát triển cụ thể. Cat-xet biểu hiện, hoặc mảnh của chúng, còn có thể được gọi là "polynucleotit được chèn" hoặc "polynucleotit cài xen" khi được biến nạp vào cây trồng.

"Gen" được xác định trong bản mô tả này là đơn vị di truyền có chứa một hoặc nhiều polynucleotit mà chiếm giữ vị trí cụ thể trên nhiễm sắc thể hoặc plasmit và chứa chỉ thị di truyền đối với đặc điểm hoặc tính trạng cụ thể ở 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. Ví dụ về proteaza của ruột bao gồm trypsin, mà thường phân cắt peptit ở phía đầu tận cùng C của các gốc lysin (K) hoặc arginin (R), và chymotrypsin, mà thường phân cắt peptit ở phía đầu tận cùng C của phenylalanin (F), tryptophan (W) hoặc tyrosin (Y).

Thuật ngữ "khác loại" khi được dùng để đề cập đến gen hoặc polynucleotit hoặc polypeptit dùng để chỉ gen hoặc polynucleotit hoặc polypeptit mà là hoặc chứa một phần của chúng không ở 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 cũng có thể bao gồm polynucleotit tự nhiên đối với sinh vật mà đã được thay đổi theo một cách nào đó (ví dụ như, bị đột biến, được bổ sung trong nhiều bản sao, được nối với các polynucleotit promotor hoặc vùng tăng cường không tự nhiên, 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 theo hướng có nghĩa (để tạo ra mARN) hoặc theo hướng đối nghĩa (để tạo ra bản phiên mã ARN đối nghĩa mà bổ sung với bản 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 hòa như promotor mà không được phát hiện thấy trong tự nhiên là kết hợp với gen đối với 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 phát hiện 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" dùng để chỉ polynucleotit không kết hợp tự nhiên với tế bào chủ mà nó được đưa vào, bao gồm nhiều bản sao không có trong 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 giảm phân.

Trình tự axit nucleic "đăng mã hóa" với trình tự axit nucleic tham 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 tham chiếu. Ví dụ, SEQ ID NO:6 đăng mã hóa với SEQ ID NO: 1 bởi vì chúng đều mã hóa cho trình tự axit amin được thể hiện bởi SEQ ID NO:16.

Thuật ngữ phân tử axit nucleic "được phân lập", polynucleotit hoặc protein 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 ở dạng tinh chế hoặc có thể tồn tại trong vật chủ tái tổ hợp chẳng hạn như trong vi khuẩn chuyển gen hoặc cây trồng 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 có điều khiển" dùng để chỉ 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, promotơ liên kết có điều khiển với polynucleotit mã hóa hoặc ARN chức năng khi nó có thể ảnh hưởng đến sự 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 promotơ). Polynucleotit mã hóa theo hướng có nghĩa hoặc đối nghĩa có thể liên kết có điều khiển với polynucleotit điều hòa.

Như được sử dụng ở đây "diệt loài gây hại," diệt côn trùng," và tương tự, dùng để chỉ 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 của sinh vật gây hại (ví dụ như, côn trùng gây hại) để sống sót, sinh trưởng, ăn, hoặc sinh sản.

"Cây trồng" là cây trồng bất kỳ ở giai đoạn phát triển bất kỳ, cụ thể là cây trồng 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 cây trồng, cơ quan của cây, hoặc toàn bộ cây.

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

"Nguyên liệu cây" dùng để chỉ lá, thân, rễ, hoa hoặc các bộ phận của hoa, quả, phấn hoa, tế bào trứng, hợp tử, hạt, cành giâm, mê 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 cây trồng" 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 cây trồng bất kỳ trong thử nghiệm ở thực vật hoặc trong nuôi cấy đều được bao gồm. Thuật ngữ này bao gồm, nhưng không giới hạn ở, cây trồng nguyên vẹn, cơ quan cây trồng, hạt cây trồng, môi trường nuôi cấy mô và nhóm bất kỳ của tế bào cây trồng được tổ chức thành các 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 cây trồng 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 cây trồng.

"Polynucleotit" dùng để chỉ polyme gồm có nhiều monome nucleotit liên kết cộng hóa trị trong mạch. "Polynucleotit" này bao gồm ADN, ARN, oligo nucleotit được cải biến (ví dụ, oligonucleotit bao gồm các bazơ mà không cơ bản đối với ARN hoặc ADN sinh học, như oligonucleotit được 2'-O-metyl hóa), và dạng tương tự. Theo một số

phương án, axit nucleic hoặc polynucleotit có thể là sợi đơn, sợi đôi, đa sợi, hoặc các dạng 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 cho polynucleotit bổ sung, ngoài polynucleotit được chỉ rõ bất kỳ.

"Polynucleotit được quan tâm" dùng để chỉ polynucleotit bất kỳ mà, khi được chuyển vào sinh vật, ví dụ như, thực vật, mang lại cho sinh vật đặc điểm mong muốn chẳng hạn như tính kháng côn trùng, tính kháng bệnh, khả năng dung chịu chất diệt cỏ, tính kháng kháng sinh, giá trị dinh dưỡng được cải thiện, hiệu suất trong quy trình công nghiệp được cải thiện, sản xuất enzym hoặc chất chuyển hóa có giá trị thương mại hoặc khả năng sinh sản được thay đổi.

Thuật ngữ "promotor" dùng để chỉ polynucleotit, thường ngược dòng (5') với polynucleotit mã hóa của nó, mà kiểm soát sự biểu hiện của polynucleotit mã hóa bằng cách tạo ra nhận diện cho ARN polymeraza và các yếu tố khác cần để phiên mã đúng.

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

Như được dùng trong bản mô tả này, thuật ngữ "tái tổ hợp" dùng để chỉ dạng axit nucleic (ví dụ như, ADN hoặc ARN) hoặc protein hoặc sinh vật mà không thường được tìm thấy trong tự nhiên và như vậy được tạo ra bởi sự can thiệp của con người. Như được dùng trong bản mô tả này, "phân tử axit nucleic tái tổ hợp" là phân tử axit nucleic có chứa sự kết hợp của các polynucleotit mà không cùng có trong tự nhiên và là kết quả của sự can thiệp của con người, ví dụ như, phân tử axit nucleic mà gồm có sự kết hợp của ít nhất là hai polynucleotit khác loại với nhau, hoặc phân tử axit nucleic mà được tổng hợp nhân tạo và có chứa polynucleotit mà bắt nguồn từ polynucleotit mà thường tồn tại trong tự nhiên, hoặc phân tử axit nucleic mà có chứa gen chuyển được kết hợp nhân tạo vào ADN hệ gen của tế bào chủ và ADN lễ kèm theo của hệ gen của tế bào chủ. Ví dụ về phân tử axit nucleic tái tổ hợp là phân tử ADN thu được từ sự cài xen của gen chuyển vào ADN hệ gen của thực vật, mà cuối cùng có thể dẫn đến sự biểu hiện của phân tử ARN hoặc protein tái tổ hợp trong sinh vật đó. Như được dùng trong bản mô tả này, "thực vật tái tổ hợp" là thực vật mà không thường tồn tại trong tự nhiên, là kết quả của sự can thiệp của

con người, và chứa gen chuyển hoặc phân tử axit nucleic khác loại kết hợp vào hệ gen của nó. Do việc thay đổi hệ gen này, thực vật tái tổ hợp rõ ràng khác với thực vật kiểu đại liên quan.

"Các yếu tố điều hòa" dùng để chỉ các trình tự liên quan đến sự kiểm soát biểu hiện của trình tự nucleotit. Các yếu tố điều hòa bao gồm promotor liên kết có điều khiển với trình tự nucleotit 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 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% tương đồng gốc nucleotit hoặc axit amin, 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, độ tương đồng đáng kể tồn tại ở vùng của các trình tự mà dài ít nhất là khoảng 50 gốc hoặc bazơ, tốt hơn nữa là ở vùng có ít nhất là khoảng 100 gốc hoặc bazơ, và tốt nhất là các trình tự về cơ bản là tương đồng ở ít nhất là khoảng 150 gốc hoặc bazơ. Theo phương án được đặc biệt ưu tiên, các trình tự về cơ bản là tương đồng trên toàn bộ chiều dài của vùng mã hóa. Hơn nữa, trình tự axit nucleic hoặc axit amin về cơ bản là tương đồng thực hiện về cơ bản là cùng chức năng.

Để so sánh trình tự, điển hình là một trình tự đóng vai trò như trình tự tham chiếu mà trình tự thử nghiệm được so sánh với nó. Khi dùng thuật toán so sánh trình tự, các trình tự thử nghiệm và tham 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 tỷ lệ phần trăm tương đồng trình tự của (các) trình tự thử nghiệm so với trình tự tham chiếu, dựa trên các thông số chương trình được chỉ định.

Sự sắp xếp thẳng hàng tối ưu của các trình tự để so sánh có thể được thực hiện, ví dụ như, bằng thuật toán tương đồng cục bộ của Smith & Waterman, Adv. Appl. Math. 2: 482 (1981), bằng thuật toán sắp xếp thẳng hàng tương đồng của Needleman & Wunsch, J.

Mol. Biol. 48: 443 (1970), bằng cách tra cứu phương pháp độ tương tự của Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85: 2444 (1988), bằng cách thực hiện vi tính hóa của các thuật toán này (GAP, BESTFIT, FASTA, và TFASTA trong gói phần mềm di truyền học Wisconsin (Wisconsin Genetics Software Package), Genetics Computer Group, 575 Science Dr., Madison, Wis.), hoặc bằng cách quan sát bằng mắt (thường xem trong tài liệu, Ausubel et al., *infra*).

Một ví dụ về thuật toán mà thích hợp để xác định tỷ lệ phần trăm tương đồng trình tự và tính tương đồng trình tự là thuật toán BLAST, mà được mô tả trong Altschul et al., 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 et al., 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 các trình tự axit amin, chương trình BLASTP sử dụng làm mặc định độ dài từ (W) bằng 3, kỳ vọng (E)

bảng 10, và ma trận điểm BLOSUM62 (xem Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 (1989)).

Ngoài việc tính tỷ lệ phần trăm độ tương đồng trình tự, thuật toán BLAST còn thực hiện phân tích thống kê độ tương tự giữa hai trình tự (xem, ví dụ, Karlin & Altschul, Proc. Natl. 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 tham chiếu nếu tổng xác suất nhỏ nhất khi so sánh trình tự axit nucleic thử nghiệm với trình tự axit nucleic tham 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 tương đồng là hai phân tử lai với nhau dưới các điều kiện nghiêm ngặt. Câu "lai đặc hiệu" dùng để chỉ sự 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 này có mặt trong hỗn hợp phức (ví dụ, tế bào tổng số) ADN hoặc ARN. "Về cơ bản liên kết" dùng để chỉ 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 càng dài lai đặc hiệu ở các nhiệt độ càng cao. Hướng dẫn khái quát về lai axit nucleic được tìm thấy trong kỹ thuật thí nghiệm Tijssen (1993) trong Laboratory Techniques in Biochemistry and Molecular Biology - Hybridization with Nucleic Acid Probes part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid 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 độ 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 dung cụ lọc trong kỹ thuật lai thẩm tách Southern hoặc northern 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 0,15M NaCl ở 72°C trong khoảng 15 phút. Ví dụ của các điều kiện rửa nghiêm ngặt là rửa 0,2× SSC ở 65°C trong 15 phút (xem, Sambrook, dưới đây, đối với mô tả đệm 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× 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× 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 1,0 M ion Na, thường nồng độ khoảng 0,01 đến 1,0M ion Na (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 nhiễu hơn gấp 2× (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 hiệu. 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 tham chiếu theo sáng chế: trình tự nucleotit tham chiếu tốt hơn là lai với trình tự nucleotit tham chiếu trong 7% natri dodecyl sulfat (SDS), 0,5M NaPO₄, 1mM 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), 0,5M NaPO₄, 1mM 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), 0,5M NaPO₄, 1mM 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), 0,5M

NaPO₄, 1mM 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), 0,5M NaPO₄, 1mM 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 thể bảo toàn.

"Tổng hợp" dùng để chỉ trình tự nucleotit có chứa bazơ hoặc dấu hiệu cấu trúc mà không có mặt trong trình tự tự nhiên. Ví dụ, trình tự nhân tạo mã hóa cho protein Cry theo sáng chế mà tương tự gần hơn về hàm lượng G+C và sự phân bố bộ ba mã hóa bình thường của gen thực vật hai lá mầm hoặc một lá mầm được coi là tổng hợp.

Như được dùng trong bản mô tả này, protein Cry mà "gây độc" đối với côn trùng gây hại có nghĩa là protein Cry đóng vai trò làm tác nhân kiểm soát côn trùng theo đường miệng để giết chết côn trùng gây hại, hoặc protein Cry có thể phá vỡ hoặc ngăn cản việc ăn của côn trùng, hoặc gây ra sự ức chế sinh trưởng đối với côn trùng gây hại, cả hai có thể gây ra hoặc không gây ra cái chết của côn trùng. Khi protein Cry theo sáng chế được phân phối đến côn trùng hoặc côn trùng trở nên tiếp xúc bằng miệng với protein Cry, kết quả thường là cái chết của côn trùng, hoặc sự sinh trưởng của côn trùng bị chậm lại, hoặc côn trùng ngừng ăn khi nguồn tạo ra protein gây độc Cry có sẵn 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 sự tích hợp ổn định phân tử ADN vào hệ gen của sinh vật mong muốn.

"Được biến nạp / chuyển gen / tái tổ hợp" dùng để chỉ 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 hệ 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), tymin (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), isoleuxin (Ile; I), leuxin (Leu; L), lysin (Lys; K), methionin (Met; M), phenylalanin (Phe; F), prolin (Pro; P), serin (Ser; S), threonin (Thr; T), tryptophan (Trp; W), tyrosin (Tyr; Y), và valin (Val; V).

Sáng chế đề xuất các hợp phần 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, chẳng hạn như *Bacillus thuringiensis*, mà gây độc đối với côn trùng gây hại và đề cập đến polynucleotit mà có chứa trình tự nucleotit mà mã hóa cho protein Cry, và đề cập đến việc tạo ra và sử dụng polynucleotit và protein Cry để kiểm soát côn trùng gây hại.

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 mảnh độc tố của chúng, trong đó trình tự nucleotit (a) có ít nhất 80% đến ít nhất 99% tương đồng trình tự với bất kỳ trong số SEQ ID NOs:1-5 hoặc trình tự mã hóa độc tố của chúng; hoặc (b) mã hóa protein bao gồm trình tự axit amin mà có ít nhất 80% đến ít nhất 99% tương đồng trình tự với bất kỳ trong số SEQ ID NOs:16-20 hoặc mảnh 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 phương án khác, trình tự nucleotit có chứa SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, hoặc mảnh mã hóa cho độc tố bất kỳ có trình tự bất kỳ trong số các SEQ ID NO:1-5. Theo phương án khác, trình tự nucleotit tổng hợp có chứa SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, hoặc mảnh mã hóa cho độc tố bất kỳ có trình tự bất kỳ trong số các SEQ ID NO:6-15.

Polynucleotit mà là mảnh của polynucleotit mã hóa cho tiền độc tố protein Cry cũng được bao hàm bởi sáng chế. "Mảnh" được dự định là phần của trình tự nucleotit mã hóa cho protein Cry. Mảnh của trình tự nucleotit có thể mã hóa cho phần có hoạt tính sinh học của protein Cry, gọi là "mảnh độc tố," hoặc nó có thể là mảnh mà có thể được dùng làm đoạn dò lai hóa hoặc đoạn môi PCR bằng cách sử dụng phương pháp được bộc lộ dưới đây. Phân tử axit nucleic mà là mảnh của trình tự nucleotit mã hóa cho protein Cry có chứa ít nhất là 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 liên tiếp, hoặc lên đến số lượng của nucleotit có mặt trong trình tự nucleotit mã hóa cho protein Cry nguyên chiều dài được bộc lộ trong bản mô tả này (ví dụ, 2343 nucleotit đối với SEQ ID NO:1) tùy thuộc vào việc sử dụng dự kiến. Các nucleotit "liên tiếp" được dự định là các gốc nucleotit mà liền kề nhau. Một số mảnh của trình tự nucleotit theo sáng chế mã hóa cho mảnh độc tố mà giữ lại hoạt tính sinh học của protein Cry và, do đó, giữ lại hoạt tính diệt côn trùng. "Giữ lại hoạt tính diệt côn trùng" được dự định là việc mảnh này có ít nhất là khoảng 30%, tốt hơn là ít nhất là khoảng 50%, tốt hơn nữa là ít nhất là khoảng 70%, tốt hơn nữa là ít nhất là khoảng 80% của hoạt tính diệt côn trùng của protein Cry. Phương pháp đo hoạt tính diệt côn trùng đã được biết rõ trong lĩnh vực. Xem tài liệu, ví dụ, Czapla and Lang (1990) J. Econ. Entomol. 83:2480-2485; Andrews et al. (1988) Biochem. J. 252:199-206; Marrone et al. (1985) J. of Economic Entomology 78:290-293; và bằng sáng chế Mỹ số 5,743,477, tất cả các tài liệu này được kết hợp ở đây bằng cách viện dẫn toàn bộ.

Mảnh độc tố của protein Cry theo sáng chế mã hóa cho ít nhất là khoảng 15, 25, 30, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, và 450 axit amin liên tiếp, hoặc lên đến tổng số lượng của axit amin có mặt trong protein Cry nguyên chiều dài theo sáng chế (ví dụ, 781 axit amin đối với SEQ ID NO:16).

Theo một số phương án, phân tử axit nucleic theo sáng chế có chứa, về cơ bản gồm có hoặc gồm có trình tự nucleotit mã hóa cho protein Cry có chứa trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% với trình tự bất kỳ trong số các SEQ ID NO:16-20 hoặc mảnh độc tố của chúng. Theo một số phương án khác, trình tự axit amin có chứa, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID

NO:16-20 hoặc mảnh độc tố của chúng. Do đó, theo một số phương án, protein Cry mà đã được hoạt hóa bằng phương thức xử lý phân giải protein, ví dụ, bằng proteaza được điều chế từ ruột côn trùng, có thể được xác định đặc điểm và các axit amin ở đầu tận cùng N hoặc đầu tận cùng C của mảnh độc tố đã hoạt hóa được xác định. Theo khía cạnh này của sáng chế, người có hiểu biết trung bình trong lĩnh vực này có thể xác định được rằng, ví dụ, mảnh độc tố của SEQ ID NO:17 có khả năng có chứa các axit amin từ khoảng 27-603 hoặc khoảng 27-624 hoặc khoảng 27-629 của SEQ ID NO:17, hoặc mảnh độc tố của SEQ ID NO:18 có khả năng có chứa các axit amin từ khoảng 43-623 hoặc khoảng 44-623 của SEQ ID NO:18, hoặc mảnh độc tố của SEQ ID NO:19 có khả năng có chứa các axit amin từ khoảng 42-640 hoặc khoảng 43-640, hoặc mảnh độc tố của SEQ ID NO:20 có khả năng nhất là có chứa các axit amin từ khoảng 27-603 hoặc khoảng 27-622 của SEQ ID NO:20, hoặc biến thể protein Cry mảnh độc tố được sản xuất bằng cách đưa vào hoặc loại trừ các vị trí xử lý proteaza ở các vị trí thích hợp trong trình tự mã hóa để cho phép, hoặc loại trừ, sự phân cắt phân giải protein của protein biến thể lớn hơn bởi proteaza của côn trùng, thực vật hoặc vi sinh vật cũng nằm trong phạm vi của sáng chế. Kết quả cuối cùng của thao tác này được hiểu là sự tạo ra của phân tử mảnh độc tố có cùng hoạt tính hoặc hoạt tính tốt hơn so với protein tiền độc tố Cry nguyên vẹn.

Theo một số phương án của sáng chế, sáng chế đề xuất gen khảm mà bao gồm promotor khác loại được liên kết có điều khiển với polynucleotit bao gồm, về cơ bản gồm có hoặc gồm có trình tự nucleotit mà mã hóa cho protein Cry gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó trình tự nucleotit (a) có độ tương đồng trình tự từ ít nhất là 80% (ví dụ như, 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 là 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) với trình tự bất kỳ trong số các SEQ ID NO:1-5, hoặc mảnh mã hóa cho độc tố của chúng; hoặc (b) mã hóa cho protein có chứa trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% (ví dụ như, 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 là 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%)

với trình tự bất kỳ trong số các SEQ ID NO:16-20, hoặc mảnh độc tố của chúng; hoặc (c) là trình tự tổng hợp của (a) hoặc (b) mà đã được tối ưu hóa bộ ba mã hóa để biểu hiện ở sinh vật chuyển gen.

Theo các phương án khác, promotor khác loại là promotor có thể biểu hiện ở thực vật. Ví dụ, mà không làm giới hạn sáng chế, promotor có thể biểu hiện ở thực vật có thể được chọn từ nhóm của các promotor gồm có các promotor ubiquitin, virus cecum yellow, TrpA ở ngô, OsMADS 6, histon H3 ở ngô, UTR 5' gen 9 thực khuẩn thể T3, sucroza synthetaza 1 ở ngô, rượu dehydrogenaza 1 ở ngô, phức hệ thu nhận ánh sáng ở ngô, protein sốc nhiệt ở ngô, mtl ở ngô, RuBP carboxylaza tiểu đơn vị nhỏ ở đậu Hà Lan, actin ở lúa, cyclophilin ở lúa, mannopin syntaza plasmit Ti, nopaline syntaza plasmit Ti, chalcon isomerase ở dạ yên thảo, protein giàu glyxin 1 ở đậu, patatin ở khoai tây, lectin, RuBP carboxylaza tiểu đơn vị nhỏ CaMV 35S và S-E9.

Theo các phương án khác, protein mã hóa bởi gen khả năng gây độc đối với một hoặc nhiều côn trùng cánh vảy gây hại được chọn từ nhóm gồm có sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu đục quả bông (*Helicoverpa armigera*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), sâu bướm Phương Đông (*Mythimna separata*), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu bướm hại đậu nhưng (*Anticarsia gemmatilis*), sâu đo đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu ngài đêm hại đậu phương tây (*Richia albicosta*), sâu ăn chồi thuốc lá (*Heliothis virescens*), sâu đục thân có sọc (*Chilo suppressalis*), sâu đục thân màu hồng (*Sesamia calamistis*) và sâu cuốn lá lúa (*Cnaphalocrocis medinalis*).

Theo phương án khác, polynucleotit bao gồm, về cơ bản gồm có hoặc gồm có trình tự nucleotit mà có độ tương đồng trình tự từ ít nhất là 85% đến ít nhất là 99% với SEQ ID NO:1, hoặc mảnh mã hóa cho độc tố của chúng, hoặc có độ tương đồng trình tự từ ít nhất là 95% đến ít nhất là 99% với SEQ ID NO:2, hoặc mảnh mã hóa cho độc tố của chúng, hoặc có độ tương đồng trình tự từ ít nhất là 90% đến ít nhất là 99% với SEQ ID NO:3, hoặc mảnh mã hóa cho độc tố của chúng, hoặc có độ tương đồng trình tự từ ít nhất là 90%

đến ít nhất là 99% với SEQ ID NO:4, hoặc mảnh mã hóa cho độc tố của chúng, hoặc có độ tương đồng trình tự từ ít nhất là 95% đến ít nhất là 99% với SEQ ID NO:5, hoặc mảnh mã hóa cho độc tố của chúng. Theo phương án khác, polynucleotit có chứa, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID NO:1-5, hoặc mảnh mã hóa cho độc tố của chúng.

Theo phương án khác, polynucleotit bao gồm, về cơ bản gồm có hoặc gồm có trình tự nucleotit mà mã hóa cho protein bao gồm, về cơ bản gồm có hoặc gồm có trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% với trình tự bất kỳ trong số các SEQ ID NO:16-20, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:16, hoặc mảnh độc tố của chúng.

Theo phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:17, hoặc mảnh độc tố của chúng.

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

hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:18, hoặc mảnh độc tố của chúng.

Theo phương án khác, trình tự axit amin có độ tương đồng trình tự ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:19, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:20, hoặc mảnh độc tố của chúng.

Theo một số phương án, gen khảm theo sáng chế bao gồm polynucleotit bao gồm, về cơ bản gồm có hoặc gồm có trình tự tổng hợp của trình tự nucleotit mà có ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với trình tự bất kỳ trong số các SEQ ID NO:6-15, hoặc mảnh mã hóa cho độc tố của chúng, trong đó trình tự tổng hợp đã được tối ưu hóa bộ ba mã hóa để biểu hiện là sinh vật 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 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 NOS:16-25, hoặc mảnh độc tố 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 trồng chuyển gen.

Theo một số phương án, sáng chế đề xuất polynucleotit tổng hợp bao gồm, về cơ bản gồm có hoặc gồm có trình tự nucleotit mà mã hóa cho protein mà gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó trình tự nucleotit này có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với trình tự bất kỳ trong số các SEQ ID NO:6-15, hoặc mảnh mã hóa cho độc tố của chúng.

Theo phương án khác, sáng chế đề xuất polynucleotit tổng hợp bao gồm, về cơ bản gồm có hoặc gồm có trình tự nucleotit mà mã hóa cho protein mà gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó trình tự nucleotit này mã hóa cho trình tự axit amin mà có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là

99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với trình tự bất kỳ trong số các SEQ ID NO:16-25, hoặc mảnh độc tố của chúng.

Protein Cry theo sáng chế có thể được phân lập từ các chủng *Bacillus thuringiensis* (*Bt*) nhất định chẳng hạn như YN171-1, GX078-2 và GX435-1. Sẽ nhận ra rằng protein Cry theo sáng chế cũng có thể được phân lập từ các chủng *Bt* khác và rằng các chủng *Bt* này có thể được phân lập bằng các kỹ thuật tiêu chuẩn và được thử nghiệm về độc tính đối với loài gây hại thuộc bộ cánh vảy theo sáng chế. Thông thường các chủng *Bt* có thể được phân lập từ mẫu môi trường bất kỳ, bao gồm đất, thực vật, côn trùng, bụi nâng hạt, và nguyên liệu mẫu khác, v.v., bằng phương pháp đã biết trong lĩnh vực. Xem tài liệu, ví dụ, Travers et al. (1987) Appl. Environ. Microbiol. 53:1263-1266; Saleh et al. (1969) Can J. Microbiol. 15:1101-1104; DeLucca et al. (1981) Can J. Microbiol. 27:865-870; và Norris, et al. (1981) "The genera *Bacillus* and *Sporolactobacillus*," In Starr et al. (eds.), The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria, Vol. II, Springer-Verlog Berlin Heidelberg. Sau khi phân lập, các chủng *Bt* có thể được thử nghiệm về độc tính đối với loài gây hại thuộc bộ cánh vảy và protein Cry được bao hàm bởi sáng chế có thể được xác định. Do đó, theo một số phương án, sáng chế đề xuất chủng *Bacillus thuringiensis* (*Bt*) được phân lập mà sản xuất protein Cry hoặc protein Cry tái tổ hợp có chứa, về cơ bản gồm có hoặc gồm có trình tự axit amin có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% so với trình tự bất kỳ trong số các SEQ ID NO: 16-25. Theo phương án khác, chủng *Bt* được chọn từ nhóm gồm có YN171-1, GX078-2 và GX435-1. Theo phương án khác nữa, protein Cry hoặc protein Cry tái tổ hợp có chứa, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID NO:16-25.

Theo một số phương án, sáng chế đề xuất protein Cry được phân lập một cách tùy ý mà gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó protein có chứa, về cơ bản gồm có hoặc gồm có (a) trình tự axit amin mà có từ độ tương đồng trình tự ít nhất là 80% đến độ tương đồng trình tự ít nhất là 99% với trình tự axit amin được thể hiện bởi trình tự bất kỳ trong số các SEQ ID NO:16-25, hoặc mảnh độc tố của chúng; hoặc (b) trình tự axit amin mà được mã hóa bởi trình tự nucleotit mà có từ độ tương đồng trình tự ít nhất là

80% đến độ tương đồng trình tự ít nhất là 99% với trình tự nucleotit được thể hiện bởi trình tự bất kỳ trong số các SEQ ID NO:6-15, hoặc mảnh mã hóa cho độc tố của chúng.

Theo phương án khác, protein Cry được phân lập một cách tùy ý có chứa, về cơ bản gồm có hoặc gồm có trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% với trình tự bất kỳ trong số các SEQ ID NO:16-25, hoặc mảnh độc tố của chúng. Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:16, hoặc mảnh độc tố của chúng.

Trình tự axit amin có 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 SEQ ID NO:17, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:18, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:19, hoặc mảnh độc tố của chúng.

Theo các phương án khác, trình tự axit amin có 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 SEQ ID NO:20, hoặc mảnh độc tố của chúng.

Theo một số phương án, trình tự axit amin bao gồm, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID NO:16-25, hoặc mảnh độc tố của chúng. Theo phương án khác, trình tự axit amin được mã hóa bởi trình tự nucleotit bao gồm, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID NO:6-15, hoặc mảnh mã hóa cho độc tố của chúng.

Theo phương án khác, protein Cry theo sáng chế gây độc đối với côn trùng cánh vảy gây hại được chọn từ nhóm gồm có sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu đục quả bông (*Helicoverpa armigera*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), sâu bướm Phương Đông (*Mythimna sepatate*), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu bướm hại đậu nhung (*Anticarsia gemmatalis*), sâu đo đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu ngài đêm hại đậu phương tây (*Richia albicosta*), sâu ăn chồi thuốc lá (*Heliothis virescens*), sâu đục thân có sọc (*Chilo suppressalis*), sâu đục thân màu hồng (*Sesamia calamistis*) và sâu cuốn lá lúa (*Cnaphalocrocis medinalis*).

Theo một số phương án, sáng chế bao hàm protein Cry tái tổ hợp mà gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó protein Cry tái tổ hợp bao gồm, về cơ bản gồm có hoặc gồm có (a) trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% với trình tự axit amin được thể hiện bởi trình tự bất kỳ trong số các SEQ ID NO:21-25, hoặc mảnh độc tố của chúng; hoặc (b) trình tự axit amin mà được mã hóa bởi trình tự nucleotit mà có độ tương đồng trình tự từ 80% đến ít nhất là 99% với trình tự nucleotit được thể hiện bởi trình tự bất kỳ trong số các SEQ ID NO:11-15, hoặc mảnh mã hóa cho độc tố của chúng.

Theo phương án khác, protein Cry tái tổ hợp bao gồm, về cơ bản gồm có hoặc gồm có trình tự axit amin mà có độ tương đồng trình tự từ ít nhất là 80% đến ít nhất là 99% với trình tự bất kỳ trong số các SEQ ID NO:21-25, hoặc mảnh độc tố của chúng. Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:21, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:22, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:23, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:24, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, trình tự axit amin có độ tương đồng trình tự ít nhất là 80%, hoặc ít nhất là 81%, hoặc ít nhất là 82%, hoặc ít nhất là 83%, hoặc ít nhất là 84%, hoặc ít nhất là 85%, hoặc ít nhất là 86%, hoặc ít nhất là 87%, hoặc ít nhất là 88%, hoặc ít nhất là 89%, hoặc ít nhất là 90%, hoặc ít nhất là 91%, hoặc ít nhất là 92%, hoặc ít nhất là 93%, hoặc ít nhất là 94%, hoặc ít nhất là 95%, hoặc ít nhất là 96%, hoặc ít nhất là 97%, hoặc ít nhất là 98%, hoặc ít nhất là 99%, hoặc ít nhất là 99,1%, hoặc ít nhất là 99,2%, hoặc ít nhất là 99,3%, hoặc ít nhất là 99,4%, hoặc ít nhất là 99,5% hoặc ít nhất là 99,6%, hoặc ít nhất là 99,7%, hoặc ít nhất là 99,8%, hoặc ít nhất là 99,9% với SEQ ID NO:25, hoặc mảnh độc tố của chúng.

Theo phương án khác nữa, protein Cry tái tổ hợp bao gồm, về cơ bản gồm có hoặc gồm có trình tự axit amin của trình tự bất kỳ trong số các SEQ ID NO:21-25, hoặc mảnh

độc tố của chúng. Theo phương án khác, protein Cry tái tổ hợp được mã hóa bởi trình tự nucleotit mà có chứa, về cơ bản gồm có hoặc gồm có trình tự bất kỳ trong số các SEQ ID NO:11-15, hoặc mảnh mã hóa cho độc tố của chúng.

Kháng thể mà được tạo ra trong đáp ứng với thử thách miễn dịch bởi BT-0049, BT-0063, BT-0064, BT-0086 và BT-0498 tự nhiên hoặc đột biến và dạng tương tự hoặc protein Cry có liên quan cũng được bao hàm bởi sáng chế. Kháng thể này có thể được tạo ra bằng cách sử dụng các kỹ thuật miễn dịch học tiêu chuẩn để sản xuất kháng huyết thanh đa dòng và, nếu muốn, làm bất tử tế bào sản xuất kháng thể của vật chủ được gây miễn dịch để có nguồn sản xuất kháng thể đơn dòng. Kỹ thuật sản xuất kháng thể đối với chất bất kỳ được quan tâm đã được biết rõ, ví dụ như, như trong tài liệu Harlow and Lane (1988. *Antibodies a laboratory manual*. pp. 726. Cold Spring Harbor Laboratory) và như trong tài liệu Goding (*Monoclonal Antibodies: Principles & practice*. 1986. Academic Press, Inc., Orlando, FL). Sáng chế bao hàm protein diệt côn trùng mà tương tác chéo với kháng thể, cụ thể là kháng thể đơn dòng, được tạo ra để chống lại một hoặc nhiều protein diệt côn trùng Cry theo sáng chế.

Kháng thể được tạo ra trong sáng chế cũng hữu dụng trong thử nghiệm miễn dịch để xác định hàm lượng hoặc sự có mặt của BT-0049, BT-0063, BT-0064, BT-0086 và BT-0498 tự nhiên hoặc đột biến hoặc protein Cry có liên quan trong mẫu sinh học. Các thử nghiệm này cũng hữu dụng trong sự sản xuất có kiểm soát chất lượng của hợp phần có chứa một hoặc nhiều protein Cry theo sáng chế hoặc các protein gây độc có liên quan. Ngoài ra, kháng thể có thể được sử dụng để đánh giá hiệu quả của sự sản xuất tái tổ hợp của một hoặc nhiều protein Cry theo sáng chế hoặc protein có liên quan, cũng như là để sàng lọc thư viện biểu hiện đối với sự có mặt của trình tự nucleotit mã hóa cho một hoặc nhiều protein Cry theo sáng chế hoặc các trình tự mã hóa cho protein có liên quan. Kháng thể cũng hữu dụng làm phối tử ái lực để tinh chế hoặc phân lập một hoặc nhiều protein bất kỳ theo sáng chế và các protein có liên quan. Protein Cry theo sáng chế và protein có chứa các epitop khác nguyên có liên quan có thể thu được bằng cách biểu hiện quá mức nguyên chiều dài hoặc một phần chiều dài của trình tự mã hóa cho tất cả hoặc một phần của protein Cry theo sáng chế hoặc protein có liên quan trong tế bào chủ được ưu tiên.

Hiểu rằng các trình tự ADN 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 này có thể được thay đổi bằng nhiều cách khác nhau bao gồm thế, loại bỏ, cắt cụt, và chèn axit amin của một hoặc nhiều axit amin nêu ở bất kỳ trong số các SEQ ID NOs:16-20, 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 hơn sự thế, loại bỏ 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 mã hóa protein. Điều này cũng có thể được thực hiện bằng một trong một vài dạng của sự gây đột biến hoặc trong sự tiến hóa có 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 của sáng chế, trình tự nucleotit được thể hiện bởi các SEQ ID NO: 1-5 được biến đổi để đưa vào sự thế axit amin trong protein được mã hóa. Theo phương án khác, protein đột biến thu được được mã hóa bởi polynucleotit đột biến tổng hợp có chứa trình tự nucleotit được thể hiện bởi trình tự bất kỳ trong số các SEQ ID NO:11-15. Theo các phương án khác, protein đột biến bao gồm, về cơ bản gồm có hoặc gồm có trình tự axit amin được đại diện bởi trình tự bất kỳ trong số các SEQ ID NO:21-25.

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 hợp phần theo sáng chế. Ví dụ, người ta có thể biểu hiện protein Cry trong tế bào chủ mà thể hiện tỉ lệ kết hợp nhằm bazơ cao trong quá trình sao chép ADN, chẳng hạn 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 vật truyền), nuôi cấy các đột biến protein Cry trong chủng không 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 tài liệu, ví dụ Marrone et al. (1985) *J. of Economic Entomology* 78:290-293. Các thử nghiệm này có thể bao gồm bước cho thực vật tiếp xúc với một hoặc nhiều loài gây hại và xác định khả năng của thực vật để sống sót hoặc gây chết loài gây hại. Ví dụ về các đột biến mà dẫn đến độc tính tăng lên được tìm thấy trong tài liệu Schnepf et al. (1998) *Microbiol. Mol. Biol. Rev.* 62:775-806.

Theo cách khác, các thay đổi có thể được tạo ra đối với trình tự axit amin theo sáng chế ở đầu tận cùng amino hoặc carboxy mà về cơ bản không ảnh hưởng đến hoạt tính. Thay đổi này có thể bao gồm sự cài xen, 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. Một cách 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 sự dung hợp protein. Các protein dung hợp này thường được sử dụng để (1) làm tăng biểu hiện của protein quan tâm (2) đưa vào miền 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 đích 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 sự thay thế axit amin trong protein Cry tự nhiên tạo ra protein Cry đột biến có miền 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 thay thế 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 miền không bảo toàn bên ngoài miền 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.

Các trình tự nucleotit và axit amin biến thể theo sáng chế cũng bao hàm các trình tự được tạo dẫn xuất từ các quy trình gây đột biến và gây tái tổ hợp chẳng hạn như sự xáo trộn 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 miền 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, bằng cách sử dụng nghiên cứu này, mô típ trình tự mã hóa miền quan tâm có thể được xáo trộn 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 quan tâm được cải thiện, như tăng hoạt tính diệt côn trùng. Các chiến lược để xáo trộn ADN này là đã biết trong lĩnh vực. Xem tài liệu, ví dụ, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer et al. (1997) *Nature Biotech.* 15:436-438; Moore et al. (1997) *J. Mol. Biol.* 272:336-347; Zhang et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer et al. (1998) *Nature* 391:288-291; và bằng sáng chế Mỹ số 5,605,793 và 5,837,458.

Sự hoán đổi hoặc xáo trộn miền là cơ chế khác để tạo ra các protein Cry thay đổi theo sáng chế. Các miền 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 để tạo ra protein tái tổ hợp và thử nghiệm chúng về hoạt tính diệt loài gây hại đã được biết rõ trong lĩnh vực (xem, ví dụ, Naimov et al. (2001) *Appl. Environ. Microbiol.* 67:5328-5330; de Maagd et al. (1996) *Appl. Environ. Microbiol.* 62:1537-1543; Ge et al. (1991) *J. Biol. Chem.* 266:17954-17958; Schnepf et al. (1990) *J. Biol. Chem.* 265:20923-20930; Rang et al. (1999) *Appl. Environ. Microbiol.* 65:2918-2925).

Theo một số phương án, sáng chế đề xuất vật truyền tái tổ hợp chứa polynucleotit, phân tử axit nucleic, cat-xet biểu hiện hoặc gen khảm theo sáng chế. Theo các phương án khác, vật truyền còn được xác định là plasmit, cosmit, phagemit, nhiễm sắc thể nhân tạo, vật truyền thể thực khuẩn hoặc virus. Một số vật truyền 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 đề cập đến băng có thể biểu hiện được thiết kế để biểu hiện polynucleotit và phân tử axit nucleic theo sáng chế. Như được sử dụng ở đây, "cat-xet biểu hiện" nghĩa là phân tử axit nucleic có ít nhất trình tự kiểm soát liên kết có điều khiển với trình tự nucleotit quan tâm. Theo phương thức này, ví dụ, các promotor thực vật được liên kết có điều khiển với trình tự nucleotit cần được biểu hiện được cung cấp trong cat-xet biểu hiện để biểu hiện ở cây trồng, phần cây trồng hoặc tế bào cây trồng.

Cat-xet biểu hiện có chứa polynucleotit được quan tâm có thể là khảm, có nghĩa là ít nhất là một trong số các thành phần của nó là khác loại đối với ít nhất là một trong số các thành phần khác của nó. Cat-xet biểu hiện có thể còn là băng mà có trong tự nhiên nhưng thu được dưới dạng tái tổ hợp hữu dụng cho sự biểu hiện khác loại. Thông thường, tuy nhiên, cat-xet biểu hiện khác loại so với vật chủ, tức là, trình tự axit nucleic cụ thể của cat-xet biểu hiện không có trong 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 promotor liên kết có điều khiển với trình tự nucleotit theo sáng chế, cat-xet biểu hiện theo sáng chế còn có thể bao gồm trình tự điều hòa khác. Như được sử dụng ở đây, "trình tự điều hòa" 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 hòa 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 polyadenyl.

Theo một số phương án, cat-xet biểu hiện theo sáng chế cũng có thể bao gồm polynucleotit mà mã hóa cho các tính trạng mong muốn khác ngoài protein Cry theo sáng chế. Cat-xet biểu hiện này có chứa các tính trạng xếp chồng có thể được sử dụng để tạo ra

cây trồng, phần cây trồng hoặc tế bào cây trồng có kiểu hình mong muốn với các tính trạng xếp chồng (tức là, sự xếp chồng phân tử). Các dạng kết hợp xếp chồng này ở cây trồng cũng có thể được tạo ra bằng phương pháp khác bao gồm, nhưng không giới hạn ở, lai chéo cây trồng bằng phương pháp thông thường bất kỳ. Nếu được chồng lên bằng các cây biến nạp gen, trình tự nucleotit quan tâm có thể kết hợp ở thời điểm và trật tự bất kỳ. Ví dụ, cây trồng 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. Trình tự nucleotit khác có thể được đưa vào một cách đồng thời trong quy trình đồng biến nạp với trình tự nucleotit, phân tử axit nucleic, cấu trúc axit nucleic, hoặc hợp phân theo sáng chế, được cung cấp bằng dạng kết hợp bất kỳ của các cat-xet biểu hiện. Ví dụ, nếu hai trình tự nucleotit được đưa vào, chúng có thể kết hợp trong các băng riêng rẽ (*trans*) hoặc có thể kết hợp trong cùng băng (*cis*). Sự biểu hiện của polynucleotit có thể được thực hiện bởi cùng một promotor hoặc các promotor khác nhau. Còn nhận thấy rằng polynucleotit có thể được xếp chồng ở vị trí hệ gen mong muốn bằng cách sử dụng hệ tái tổ hợp đặc hiệu vị trí. 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 cũng có thể bao gồm trình tự mã hóa khác đối với một hoặc nhiều polypeptit hoặc các phân tử ARN sợi kép (dsARN) được quan tâm đối với các tính trạng nông nghiệp mà chủ yếu là có lợi đối với công ty hạt giống, người trồng hoặc người phát triển hạt. Polypeptit quan tâm có thể là polypeptit bất kỳ được mã hóa bởi trình tự nucleotit mong muốn. Ví dụ không làm giới hạn sáng chế về polypeptit được quan tâm mà thích hợp để sản xuất ở cây trồng bao gồm các polypeptit dẫn đến các tính trạng quan trọng trong nông nghiệp chẳng hạn như tính kháng chất diệt cỏ (đôi khi được đề cập đến dưới dạng "khả năng dung chịu chất diệt cỏ"), tính kháng virus, tính kháng tác nhân gây bệnh là vi khuẩn, tính kháng côn trùng, tính kháng giun tròn, hoặc tính kháng nấm. Xem, ví dụ, bằng 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 quan tâm (ví dụ, chất chỉ thị chọn lọc, màu bao hạt, v.v.). Các polypeptit quan tâm khác nhau,

cũng như phương pháp để đưa các polypeptit này vào cây trồng, được mô tả, ví dụ, trong bằng 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/.

Polynucleotit tạo ra tính kháng/khả năng chịu chất diệt cỏ ức chế điểm phát triển hoặc mô phân sinh, như imidazalinon hoặc sulfonylure có thể thích hợp theo một số phương án của sáng chế. Polynucleotit ví dụ trong loại này mã hóa đối với enzym ALS và AHAS đột biến như mô tả, ví dụ, trong bằng sáng chế Mỹ số 5,767,366 và 5,928,937. Bằng 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 aryloxyphenoxypromanoic axit. Sức đề kháng được tạo ra bởi axetyl coenzym A carboxylaza (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 pyridinooxy hoặc axit phenoxy propionic và xyclohexon cũng là thích hợp. Xem, công bố đơn sáng chế châu Âu 0 242 246. Cũng xem, bằng sáng chế Mỹ số 5,879,903, 5,276,268 và 5,561,236.

Polynucleotit thích hợp khác bao gồm các polynucleotit mã hóa để 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 cho tính kháng chất diệt cỏ bao gồm các

polynucleotit mã hóa cho tính kháng axit 2,2-đicloropropionic, sethoxydim, haloxyfop, chất diệt cỏ imidazolinon, chất diệt cỏ sulfonylure, chất diệt cỏ triazolopyrimidin, chất diệt cỏ s-triazin và bromoxynil. Cũng thích hợp là polynucleotit tạo ra khả năng kháng protox enzym, hoặc mà 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 bằng 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). Đã nhận ra rằng hàm lượng của sự sản xuất polypeptit diệt loài gây hại ở cây trồng cần để kiểm soát côn trùng hoặc loài gây hại khác có thể thay đổi tùy thuộc vào cây trồng, kiểu loài gây hại, các yếu tố môi trường và dạng 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 có chứa trình tự nucleotit mã hóa cho protein Cry *Bacillus thuringiensis* (Bt) từ một vài loài phụ đã được tách dòng và các dòng tái tổ hợp đã được phát hiện là gây độc đối với ấu trùng côn trùng thuộc loài gây hại thuộc bộ cánh vảy, bộ hai cánh và bộ cánh cứng. Ví dụ về protein diệt côn trùng Bt này bao gồm protein Cry chẳng hạn như *Cry1Aa*, *Cry1Ab*, *Cry1Ac*, *Cry1B*, *Cry1C*, *Cry1D*, *Cry1Ea*, *Cry1Fa*, *Cry3A*, *Cry9A*, *Cry9B*, *Cry9C*, và dạng tương tự, cũng như là protein diệt côn trùng ở thực vật chẳng hạn như *Vip1*, *Vip2*, *Vip3*, và dạng tương tự. Danh sách đầy đủ của protein có nguồn gốc từ Bt có thể được tìm thấy trên trang mạng lưới toàn cầu ở cơ sở dữ liệu danh pháp độc tố *Bacillus thuringiensis* được duy trì bởi University of Sussex (cũng xem trong tài liệu, Crickmore *et al.* (1998) *Microbiol. Mol. Biol. Rev.* 62:807-813).

Polypeptit mà thích hợp để sản xuất ở cây trồng còn bao gồm polypeptit mà cải thiện hoặc theo cách khác làm thuận lợi cho sự biến đổi của cây trồng hoặc phần cây trồng được thu hoạch thành sản phẩm hữu dụng trên thị trường, bao gồm, ví dụ, hàm lượng hoặc sự phân bố carbohydrat tăng lên hoặc được thay đổi, các tính chất lên men

được cải thiện, hàm lượng dầu tăng lên, hàm lượng protein tăng lên, khả năng tiêu hóa được cải thiện, và tăng lên hàm lượng thành phần hỗ trợ dinh dưỡng, ví dụ như, hàm lượng phytosterol tăng lên, hàm lượng tocopherol tăng lên, hàm lượng stanol tăng lên hoặc hàm lượng vitamin tăng lên. Các polypeptit quan tâm 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" dùng để chỉ polypeptit quan tâm có thể đóng góp trực tiếp hoặc gián tiếp vào sự tồn tại của đặc điểm quan tâm (ví dụ, làm tăng thoái biến xenluloza bằng cách sử dụng enzym xenlulaza khác loại).

Theo một số phương án, polypeptit góp phần vào khả năng tiêu hóa được cải thiện đối với thực phẩm hoặc thức 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 là nấm và vi khuẩn đã được xác định và được xác định đặc điểm (xem tài liệu, ví dụ như, bằng sáng chế Mỹ số 5,437,992; Coughlin *et al.* (1993) "Proceedings of the Second TRICEL Symposium on *Trichoderma reesei* Cellulases and Other Hydrolases" Espoo; Souminen and Reinikainen, eds. (1993) *Foundation for Biotechnical and Industrial Fermentation Research* 8:125-135; công bố bằng sáng chế Mỹ số 2005/0208178; và công bố đơn PCT số WO 03/16654). Cụ thể là, ba xylanaza đặc hiệu (XYL-I, XYL-II, và XYL-III) đã được xác định trong tài liệu *T. reesei* (Tenkanen *et al.* (1992) *Enzyme Microb. Technol.* 14:566; Torronen *et al.* (1992) *Bio/Technology* 10:1461; và Xu *et al.* (1998) *Appl. Microbiol. Biotechnol.* 49:718).

Theo phương án khác, polypeptit hữu dụng đối với sáng chế có thể là enzym làm thoái hóa polysacarit. 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 làm giảm polysaccarit bao gồm: enzym làm giảm 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 khử phân nhánh tinh bột, như a) isoamylaza (EC 3.2.1.68), pullulanaza (EC 3.2.1.41), và tương tự; b) xenlulaza như exo-1,4-3-cellobiohydrolaza (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 ở 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 cellobiohydrolaza (CBH) (EC 3.2.1.91). Theo một phương án, enzym cellobiohydrolaza 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), glucoza oxidaza, pectinaza, xylanaza, transglucosidaza, alpha 1,6 glucosidaza (ví dụ, E.C. 3.2.1.20); esteraza như axit esteraza ferulic (EC 3.1.1.73) và axetyl xylan esteraza (EC 3.1.1.72); và cutinaza (ví dụ E.C. 3.1.1.74).

Các phân tử ARN sợi kép hữu dụng với sáng chế bao gồm, nhưng không giới hạn ở các phân tử ARN sợi kép mà kìm hãm gen côn trùng đích. Như được sử dụng trong bản mô tả này các từ "sự kìm hãm gen", khi dùng cùng nhau, được dự định là dùng để chỉ phương pháp bất kỳ trong số các phương pháp đã biết rõ để làm giảm hàm lượng của protein được tạo ra là kết quả của sự phiên mã gen thành mARN và sự dịch mã tiếp theo của mARN. Sự kìm hãm gen cũng được dự định có nghĩa là sự giảm đi của sự biểu hiện protein từ gen hoặc trình tự mã hóa bao gồm sự kìm hãm gen sau dịch mã và sự kìm hãm phiên mã. Sự kìm hãm gen sau dịch mã được điều hòa qua trung gian 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 kìm hãm và ARN sợi kép tương ứng được dùng để kìm hãm, và dùng để chỉ sự giảm đáng kể và có thể đo được của lượng mARN sẵn có trong tế bào để liên kết bởi ribosom. ARN được phiên mã có thể theo hướng có nghĩa để ảnh hưởng đến cái gọi là sự đồng kìm hãm, theo hướng đối nghĩa để ảnh hưởng đến cái gọi là sự kìm hãm đối nghĩa, hoặc theo cả hai hướng tạo ra dsARN để ảnh hưởng đến cái gọi là sự cản trở ARN (ARNi). Sự kìm hãm phiên mã được điều hòa qua trung gian sự có mặt trong tế bào của dsARN, tác nhân kìm hãm gen, thể hiện độ tương đồng trình tự đáng kể đối với trình tự ADN promoter hoặc phần bổ sung của chúng để ảnh hưởng đến cái gọi là sự kìm hãm qua promoter. Sự kìm hãm gen có thể hữu hiệu chống lại gen cây trồng tự nhiên có liên quan đến tính trạng, ví dụ như, để tạo ra cây trồng thực vật có hàm lượng giảm của protein được mã hóa bởi gen tự nhiên này hoặc có hàm lượng được tăng cường hoặc được giảm đi của chất chuyển hóa bị ảnh hưởng. Sự kìm hãm gen cũng có thể hữu hiệu chống lại gen đích ở loài gây hại thực vật mà có thể ăn hoặc tiếp xúc nguyên liệu cây trồng có chứa tác nhân kìm hãm gen, được thiết kế đặc hiệu để ức chế hoặc kìm hãm sự biểu hiện của một hoặc nhiều trình tự tương đồng hoặc trình tự bổ sung trong tế bào của loài gây hại. Gen được nhắm đích để kìm hãm này có thể mã hóa cho protein thiết yếu, chức năng đã được dự đoán mà được chọn từ nhóm gồm có sự tạo thành cơ, sự tạo thành hormon kích sâu non, sự điều hòa hormon kích sâu non, sự điều hòa và vận chuyển ion, sự tổng hợp enzym tiêu hóa, sự duy trì điện thế màng tế bào, sinh tổng hợp axit amin, sự thoái hóa axit amin, sự tạo thành tinh trùng, sự tổng hợp pheromon, sự thụ cảm pheromon, sự tạo thành anten, sự tạo thành cánh, sự tạo thành chân, sự phát triển và sự biệt hóa, sự tạo thành trứng, sự

trưởng thành ấu trùng, sự tạo thành enzym tiêu hóa, sự tổng hợp máu (haemolymph), sự duy trì máu (haemolymph), sự dẫn truyền thần kinh, sự phân chia tế bào, sự chuyển hóa năng lượng, sự hô hấp, và sự chết theo chương trình.

Theo một số phương án, sáng chế đề xuất tế bào chủ chuyển gen không phải của người bao gồm polynucleotit, phân tử axit nucleic, gen khảm, cat-xet biểu hiện hoặc vật truyền tải 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 các chi *Bacillus*, *Brevibacillus*, *Clostridium*, *Xenorhabdus*, *Photorhabdus*, *Pasteuria*, *Escherichia*, *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Salmonella*, *Pasteurella*, *Xanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Methylophilus*, *Agrobacterium*, *Acetobacter*, *Lactobacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, hoặc *Alcaligenes*. Do đó, ví dụ, để làm các tác nhân kiểm soát côn trùng sinh học, protein Cry theo sáng chế có thể được tạo ra bằng cách biểu hiện gen khảm mã hóa cho protein Cry theo sáng chế ở tế bào vi khuẩn. Ví dụ, theo một số phương án, sáng chế đề xuất tế bào *Bacillus thuringiensis* có chứa gen khảm theo sáng chế.

Theo phương án khác, sáng chế đề xuất tế bào cây trồng 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 một số lượng các tế bào cây hai lá mầm hoặc tế bào cây một lá mầm biểu hiện protein Cry theo sáng chế được mã hóa bởi gen khảm theo sáng chế. Theo các phương án khác nhiều trong số các tế bào được để cạnh để tạo ra con đường vô bào và phát triển dưới ánh sáng mặt trời tự nhiên.

Theo phương án khác theo sáng chế, protein diệt côn trùng Cry theo sáng chế được biểu hiện ở sinh vật bậc cao, ví dụ, thực vật. Trong trường hợp này, cây trồng chuyển gen

biểu hiện lượng hữu hiệu của protein diệt côn trùng bảo vệ bản thân chúng khỏi loài gây hại thực vật chẳng hạn như côn trùng gây hại. Khi côn trùng bắt đầu ăn trên cây trồng chuyển gen này, nó ăn protein diệt côn trùng Cry đã được biểu hiện. Điều này có thể ngăn cản côn trùng khỏi hút thêm vào mô cây trồng hoặc có thể còn gây hại hoặc giết côn trùng. Polynucleotit theo sáng chế được chèn vào cat-xet biểu hiện, mà sau đó được tích hợp ổn định trong hệ gen của cây. Theo phương án khác, polynucleotit được bao gồm trong virus tự sao chép không gây 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, anh đào, đào, xuân đào, mơ, dâu tây, nho, mâm xôi, dâu, dứa, 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 vào 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 trồng chuyển gen, do đó gây ra sự sinh tổng hợp của protein Cry tương ứng, ở dạng tiền độc tố hoặc độc tố trưởng thành, trong cây trồng chuyển gen. Bằng cách này, cây trồng 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 trồng chuyển gen, trình tự nucleotit theo sáng chế có thể cần biến đổi và tối ưu hóa. 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 trồng chuyển gen có thể tạo ra từ trình tự nucleotit của vi khuẩn có bộ ba mã hóa mà không được ưu tiên trong cây trồng. Đã biết trong lĩnh vực rằng các sinh vật sống có các ưu tiên cụ thể đối với việc sử dụng bộ

ba mã hóa, và bộ ba mã hóa 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 hàm lượng GC ít nhất khoảng 35%, 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 hóa 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 hóa 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 *et al.* 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 bằng sáng chế Mỹ số 5,625,136; 5,500,365 và 6,013,523.

Theo một số phương án, sáng chế đề xuất các trình tự hoặc polynucleotit mã hóa tổng hợp được tạo ra theo quy trình được bộc lộ trong bằng sáng chế Mỹ số 5,625,136, được kết hợp ở đây bằng cách viện dẫn. Trong quy trình này, bộ ba mã hóa được ưu tiên của ngô, tức là, bộ ba mã hóa đơn mà mã hóa thường xuyên nhất axit amin ở ngô, được sử dụng. Bộ ba mã hóa 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 hóa của ngô cho 28 gen từ cây ngô được tìm thấy trong Murray *et al.*, Nucleic Acids 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 được tối ưu hóa bộ ba mã hóa của ngô được đại diện bởi trình tự bất kỳ trong số các SEQ ID NO: 6-15. 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 hóa hoặc tổng hợp. Tức là, polynucleotit

có thể có chứa trình tự nucleotit mà là trình tự tự nhiên một phần và trình tự được tối ưu hóa bộ ba mã hóa một phần.

Để khởi đầu dịch mã hiệu quả, các trình tự liên kế 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 trình tự liên ứng thích hợp đối với thực vật (NAR 15:6643-6653 (1987)) và Clonotech đề xuất chất khởi đầu dịch mã liên ứng khác (danh mục 1993/1994, trang 210). Các trình tự 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).

Các trình tự mã hóa cho protein Cry mới theo sáng chế, dưới dạng trình tự tự nhiên của chúng hoặc dưới dạng trình tự tổng hợp như được mô tả ở trên, có thể được dung hợp có điều khiển với nhiều promotor để biểu hiện ở thực vật bao gồm các promotor cấu trúc, cảm ứng, được điều hòa tạm thời, được điều hòa tiến triển, được điều hòa 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 promotor sẽ thay đổi tùy thuộc vào yêu cầu thời gian và không gian, và còn tùy thuộc vào loài đích. Do đó, sự biểu hiện của trình tự nucleotit theo sáng chế ở lá, ở cuống hoặc thân, ở bông, ở cụm hoa (ví dụ như chùm, chùy, lõi, v.v.), ở rễ, hoặc cây 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 promotor từ các cây hai lá mầm đã thể hiện hoạt động trong các cây một lá mầm và ngược lại, lý tưởng là các promotor của cây hai lá mầm được chọn để biểu hiện trong cây hai lá mầm và các promotor của cây một lá mầm để biểu hiện trong cây một lá mầm. Tuy nhiên, không có hạn chế đối với nguồn gốc của promotor được chọn; chỉ cần chúng hoạt động trong việc dẫn hướng biểu hiện của trình tự nucleotit trong tế bào mong muốn.

Các promotor cấu trúc thích hợp bao gồm, ví dụ, promotor CaMV 35S (SEQ ID NO:1546; Odell et al., Nature 313:810-812, 1985); promotor Arabidopsis At6669 (SEQ ID NO:1652; xem công bố đơn PCT Số W004081173A2); Ubi 1 ở ngô (Christensen et al.,

Plant Mol. Biol. 18:675-689, 1992); actin ở lúa (McElroy et al., Plant Cell 2:163-171, 1990); pEMU (Last et al., Theor. Appl. Genet. 81:581-588, 1991); CaMV 19S (Nilsson et al., Physiol. Plant 100:456-462, 1997); GOS2 (de Pater et al., Plant J November; 2(6):837-44, 1992); ubiquitin (Christensen et al., Plant Mol. Biol. 18: 675-689, 1992); xyclophilin ở lúa (Bucholz et al., Plant Mol Biol. 25(5):837-43, 1994); histon H3 ngô (Lepetit et al., Mol. Gen. Genet. 231: 276-285, 1992); Actin 2 (An et al., Plant J. 10(1):107-121, 1996), promotor CT2 đỉnh rễ cấu trúc (SEQ ID NO:1535; cũng xem công bố đơn PCT Số IL/2005/000627) và Super MAS tổng hợp (Ni et al., The Plant Journal 7: 661-76, 1995). Các promotor cấu trúc khác bao gồm các promotor trong bằng 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.

Promotor 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 sự biểu hiện ở rễ, ruột cây, lá hoặc phần hoa. Các promotor đặc hiệu mô thích hợp bao gồm, nhưng không giới hạn ở, các promotor đặc hiệu lá [chẳng hạn như được mô tả, ví dụ, bởi Yamamoto et al., Plant J. 12:255-265, 1997; Kwon et al., Plant Physiol. 105:357-67, 1994; Yamamoto et al., Plant Cell Physiol. 35:773-778, 1994; Gotor et al., Plant J. 3:509-18, 1993; Orozco et al., Plant Mol. Biol. 23:1129-1138, 1993; và Matsuoka et al., Proc. Natl. Acad. Sci. USA 90:9586-9590, 1993], các promotor ưu tiên hạt [ví dụ như, từ gen đặc hiệu hạt (Simon, et al., Plant Mol. Biol. 5. 191, 1985; Scofield, et al., J. Biol. Chem. 262: 12202, 1987; Baszczynski, et al., Plant Mol. Biol. 14: 633, 1990), Brazil Nut albumin (Pearson' et al., Plant Mol. Biol. 18: 235-245, 1992), legumin (Ellis, et al. Plant Mol. Biol. 10: 203-214, 1988), Glutelin (rice) (Takaiwa, et al., Mol. Gen. Genet. 208: 15-22, 1986; Takaiwa, et al., FEBS Letts. 221: 43-47, 1987), Zein (Matzke et al., Plant Mol Biol, 143).323-32 1990), napA (Stalberg, et al., Planta 199: 515-519, 1996), SPA lúa mì (Albanietal, Plant Cell, 9: 171-184, 1997), oleosin hướng dương (Cummins, etal., Plant Mol. Biol. 19: 873-876, 1992)], các promotor đặc hiệu nội nhũ [ví dụ như, LMW và HMW lúa mì, glutenin-1 (Mol Gen Genet 216:81-90, 1989; NAR 17:461-2), a, b và g gliadin lúa mì (EMB03:1409-15, 1984), promotor ltr1 lúa mạch, B1, C, D hordein lúa mạch (Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60,

1996), Barley DOF (Mena et al., *The Plant Journal*, 116(1): 53-62, 1998), Biz2 (EP99106056.7), Promotor tổng hợp (Vicente-Carbajosa et al., *Plant J.* 13: 629-640, 1998), prolamin NRP33 lúa, globulin lúa Glb-1 (Wu et al., *Plant Cell Physiology* 39(8) 885-889, 1998), alpha-globulin lúa REB/OHP-1 (Nakase et al. *Plant Mol. Biol.* 33: 513-S22, 1997), ADP-glucoza PP lúa (*Trans Res* 6:157-68, 1997), họ gen ESR ngô (*Plant J* 12:235-46, 1997), gamma-kafirin lúa miễn (*Plant Mol. Biol* 32:1029-35, 1996)], các promotor đặc hiệu phôi [ví dụ như, OSH1 lúa (Sato et al., *Proc. Nati. Acad. Sci. USA*, 93: 8117-8122), KNOX (Postma-Haarsma of al, *Plant Mol. Biol.* 39:257-71, 1999), oleosin lúa (Wu et at, *J. Biochem.*, 123:386, 1998)], các promotor đặc hiệu hoa [ví dụ như, AtPRP4, chalone syntaza (chsA) (Van der Meer, et al., *Plant Mol. Biol.* 15, 95-109, 1990), LAT52 (Twell et al., *Mol. Gen Genet.* 217:240-245; 1989), apetala-3, các SEQ ID NO thực vật [ví dụ như, các promotor OsmADS (đơn sáng chế Mỹ số 2007/0006344)].

Trình tự nucleotit theo sáng chế còn có thể được biểu hiện dưới sự điều biến của promotor 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 sáng chế châu Âu số EP 0 332 104 và bằng sáng chế Mỹ số 5,614,395. Theo một phương án, promotor điều biến hóa học là promotor PR-1a của thuốc lá.

Nhóm promotor khác hữu dụng theo sáng chế là loại mà cảm ứng cuộn. Nhiều promotor 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, promotor 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 côn trùng gây hại xâm chiếm. Ví dụ về các promotor thuộc loại này bao gồm các promotor được mô tả bởi Stanford *et al.* *Mol. Gen. Genet.* 215:200-208 (1989), Xu *et al.* *Plant Molec. Biol.* 22:573-588 (1993), Logemann *et al.* *Plant Cell* 1:151-158 (1989), Rohrmeier & Lehle, *Plant Molec. Biol.* 22:783-792 (1993), Firek *et al.* *Plant Molec. Biol.* 22:129-142 (1993), và Warner *et al.* *Plant J.* 3:191-201 (1993).

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

Bên cạnh việc lựa chọn promotor 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 có điều khiển 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 băng 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. Các trình tự này đã được xác định đặc điểm mà gây ra sự nhắm đích của sản phẩm gen vào các khoang tế bào khác. Các trình tự đầu tận cùng amino có thể chịu trách nhiệm đối với việc nhắm đích protein được quan tâm vào khoang tế bào bất kỳ, chẳng hạn như, không bào, ty thể, peroxisom, thể protein, lưới nội chất, lục lạp, hạt tinh bột, lạp bột, con

đường vô bào hoặc thành tế bào của thực vật (ví dụ như Unger et. al. Plant Molec. Biol. 13: 411-418 (1989); Rogers et. al. (1985) Proc. Natl. Acad. Sci. USA 82: 6512-651; bằng sáng chế Mỹ số 7,102,057; công bố đơn sáng chế quốc tế số WO 2005/096704, tất cả chúng được kết hợp tại đây bằng cách viện dẫn. Một cách tùy ý, trình tự tín hiệu có thể là trình tự tín hiệu đầu tận cùng N từ sấp, trình tự tín hiệu đầu tận cùng N từ gamma-zein, miền liên kết tinh bột, miền liên kết tinh bột đầu tận cùng C, trình tự nhắm đích lục lạp, mà nhập protein trưởng thành vào lục lạp (Comai et. al. (1988) J. Biol. Chem. 263: 15104-15109; van den Broeck, et. al. (1985) Nature 313: 358-363; bằng sáng chế Mỹ số 5,639,949) hoặc trình tự tín hiệu tiết từ tế bào aloron (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Ngoài ra, các trình tự đầu tận cùng amino kết hợp với các trình tự đầu tận cùng carboxy chịu trách nhiệm đối với việc nhắm đích không bào của sản phẩm gen (Shinshi et. al. (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 promotor cùng nguồn của chúng, mà còn cùng với promotor khác loại sao cho ảnh hưởng đến mục tiêu nhắm đích tế bào cụ thể dưới sự điều biến phiên mã của promotor mà có mẫu biểu hiện khác của promotor từ đó tạo ra tín hiệu nhắm đích.

Biến nạp cây trồng

Quy trình để biến nạp cây trồng là đã biết và quen thuộc trong lĩnh vực và được mô tả qua các tài liệu tham khảo. Ví dụ không làm giới hạn sáng chế về phương pháp để biến nạp cây trồng bao gồm sự biến nạp thông qua sự chuyển axit nucleic qua trung gian vi khuẩn (ví dụ, thông qua *Agrobacterium*), sự chuyển axit nucleic qua trung gian virus, sự chuyển axit nucleic qua trung gian sợi tinh thể silic cacbua hoặc axit nucleic, sự chuyển axit nucleic qua trung gian liposom, vi tiêm, bắn vi hạt, sự biến nạp qua trung gian canxi-

phosphat, sự biến nạp qua trung gian xyclodextrin, đục lỗ điện, sự biến nạp qua trung gian hạt nano, siêu âm, sự thâm nhiễm, sự hấp thụ axit nucleic qua trung gian PEG, cũng như là các cơ chế điện, hóa, lý (cơ học) hoặc sinh học khác bất kỳ mà gây ra sự đưa axit nucleic vào trong tế bào cây trồng, bao gồm dạng kết hợp bất kỳ của chúng. Hướng dẫn chung về các phương pháp biến nạp thực vật khác nhau đã biết trong lĩnh vực bao gồm Miki *et al.* ("Procedures for Introducing Foreign DNA into Plants" in *Methods in Plant Molecular Biology and Biotechnology*, Glick, B.R. and Thompson, J.E., Eds. (CRC Press, Inc., Boca Raton, 1993), pages 67-88) và Rakowoczy-Trojanowska (*Cell. Mol. Biol. Lett.* 7:849-858 (2002)).

Đối với sự biến nạp qua trung gian *Agrobacterium*, các vật truyền hai kỳ hoặc vật truyền mang ít nhất một trình tự viền T-ADN thích hợp, trong khi đó đối với chuyển gen trực tiếp (ví dụ, bắn hạt và dạng tương tự) vật truyền bất kỳ là thích hợp và ADN thẳng chỉ chứa cấu trúc quan tâm có thể được sử dụng. Trong trường hợp chuyển gen trực tiếp, biến nạp với các loại ADN đơn hoặc đồng biến nạp có thể được sử dụng (Schocher *et al.*, *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 thu với chất chỉ thị chọn lọc có thể là lựa chọn dương (phosphomanoza isomeraza), tạo ra khả năng kháng kháng sinh (kanamycin, hygromycin hoặc methotrexat) hoặc chất diệt cỏ (glyphosat hoặc glufosinat). Tuy nhiên, việc lựa chọn chất chỉ thị chọn lọc không thuộc phạm vi sáng chế.

Sự biến nạp qua trung gian *Agrobacterium* là phương pháp thường được sử dụng để biến nạp thực vật bởi vì hiệu quả biến nạp cao của nó và bởi vì khả năng sử dụng rộng của nó với nhiều loài khác nhau. Sự biến nạp qua trung gian *Agrobacterium* thường gồm có sự chuyển của vật truyền nhị thể mang ADN ngoại lai được quan tâm vào chủng *Agrobacterium* thích hợp mà có thể phụ thuộc vào phần bổ sung của gen *vir* được mang bởi chủng vật chủ *Agrobacterium* trên plasmit Ti cư trú hoặc trên nhiễm sắc thể (Uknes *et al.* (1993) *Plant Cell* 5:159-169). Việc chuyển vật truyền nhị thể tái tổ hợp vào *Agrobacterium* có thể được thực hiện bằng quy trình giao phối ba bên sử dụng *Escherichia coli* mang vật truyền nhị thể tái tổ hợp, chủng *E. coli* hỗ trợ mang plasmit mà có khả năng di chuyển vật truyền nhị thể tái tổ hợp đến chủng *Agrobacterium* đích. Theo

cách khác, vật truyền nhĩ thể tái tổ hợp có thể được chuyển vào *Agrobacterium* bằng cách biến nạp axit nucleic (Höfgen & Willmitzer (1988) *Nucleic Acids Res.* 16:9877).

Cây hai lá mầm cũng như là cây một lá mầm có thể được biến nạp bằng cách sử dụng *Agrobacterium*. Phương pháp biến nạp qua trung gian *Agrobacterium* ở lúa bao gồm các phương pháp đã biết để biến nạp lúa, chẳng hạn như các phương pháp được mô tả trong tài liệu bất kỳ trong số các tài liệu sau: đơn sáng chế Châu Âu EP 1198985 A1, Aldemita and Hodges (Planta 199: 612-617, 1996); Chan et al. (Plant Mol Biol 22 (3): 491-506, 1993), Hiei et al. (Plant J 6 (2): 271-282, 1994), mà bản boc lộ được kết hợp ở đây để tham khảo như thể được nêu đầy đủ. Trong trường hợp biến nạp ngô, phương pháp được ưu tiên là phương pháp như được mô tả trong tài liệu Ishida et al. (Nat. Biotechnol 14(6): 745-50, 1996) hoặc Frame et al. (Plant Physiol 129(1): 13-22, 2002), mà bản boc lộ được kết hợp ở đây để tham khảo như thể được nêu đầy đủ. Phương pháp này được mô tả thêm bằng cách lấy ví dụ trong tài liệu B. Jenes et al., Techniques for Gene Transfer, trong: Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung and R. Wu, Academic Press (1993) 128-143 và trong tài liệu Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225). Axit nucleic hoặc cấu trúc cần được biểu hiện tốt hơn là được tách dòng vào vật truyền, mà thích hợp để biến nạp *Agrobacterium tumefaciens*, ví dụ pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). *Agrobacteria* được biến nạp bằng vật truyền này sau đó có thể được sử dụng theo phương thức đã biết để biến nạp thực vật, chẳng hạn như thực vật dùng làm mẫu, như *Arabidopsis* hoặc cây trồng chẳng hạn như, bằng cách lấy làm ví dụ, cây thuốc lá, ví dụ bằng cách nhúng lá đã nghiền thô hoặc lá đã băm trong dung dịch *agrobacteria* và sau đó nuôi cấy chúng trong môi trường thích hợp. Sự biến nạp của thực vật bằng phương thức *Agrobacterium tumefaciens* được mô tả, ví dụ, bởi Hagen và Willmitzer trong tài liệu 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 tài liệu Transgenic Plants, Vol. 1, Engineering and Utilization, eds. S. D. Kung and 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 mô cây 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 đầu 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 tro ở mô và tế bào cây. Xem, ví dụ, bằng 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 tro 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 tro được sử dụng, vật truyền có thể được đưa vào tế bào bằng cách phủ các hạt bằng các vật truyền chứa axit nucleic quan tâm. Tùy ý, tế bào hoặc thể bào có thể bao quanh bởi vật truyền mà vật truyền được 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 cây trồng.

Theo phương án khác, polynucleotit theo sáng chế có thể được biến nạp trực tiếp vào hệ 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 sự kiểm soát của promoter đơn. Công nghệ biến nạp lặp thể được mô tả bao quát trong bằng sáng chế Mỹ số 5,451,513, 5,545,817, và 5,545,818, trong công bố đơn PCT số WO 95/16783, và trong McBride et al. (1994) *Proc. Natl. Acad. Sci. USA* 91, 7301-7305. Kỹ thuật cơ sở để biến nạp lặp thể liên quan đến việc đưa các miền ADN lặp thể đã tách dòng giới hạn hai đầu một chất chỉ thị chọn lọc cùng với gen quan tâm vào mô đích thích hợp, ví dụ, dùng súng bắn gen hoặc biến nạp tế bào trần (như biến nạp qua trung gian canxi clorua hoặc PEG). Các miền giới hạn hai đầu 1-1,5kb, được gọi là các trình tự đích, tạo điều kiện cho tái tổ hợp tương đồng với hệ gen lặp thể và do đó cho phép thay thế hoặc sửa đổi các miền cụ thể của hệ gen lặp thể. Ban đầu, các đột biến điểm trong các gen rARN và rps12 16S lục lạp mang lại tính kháng đối với spectinomycin hoặc streptomycin có thể được sử dụng làm chất chỉ thị chọn lọc để biến nạp (Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., and 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 vật truyền định hướng lặp thể để đưa vào các gen lạp (Staub, J.M., and Maliga, P. (1993) *EMBO J.* 12, 601-606). Sự tăng lên đáng

kể ở tần số biến nạp có thể đạt được bằng cách thay thế gen kháng kháng sinh rARN hoặc r-protein lặn bằng chất chỉ thị chọn lọc trội, gen *aadA* vi khuẩn mã hóa cho enzym khử độc spectinomycin aminoglycosit- 3'- adenyltransferaza (Svab, Z., and Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917). Trước đây, chỉ thị 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 chất 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 hệ gen lặp thể vòng có trong mỗi tế bào của cây, lợi dụng số lượng rất lớn bản sao lợi thể hơn các gen được biểu hiện trong nhân để cho phép các mức biểu hiện có thể dễ dàng vượt quá 10% tổng protein ở cây trồng hoà tan. Theo một phương án, polynucleotit theo sáng chế có thể được chèn vào vật truyền định hướng lặp thể và được biến nạp vào hệ gen lặp thể của vật chủ thực vật mong muốn. Do đó, thực vật đồng nhất đối với các hệ 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 đối với cây trồng, tế bào cây trồng hoặc môi trường nuôi cấy mô cây trồng được biến nạp, chuyển gen là thông thường trong lĩnh vực và có thể được sử dụng trong phương pháp theo sáng chế được đề xuất trong bản mô tả này. Ví dụ, vật truyền tái tổ hợp theo sáng chế cũng có thể bao gồm cat-xet biểu hiện có chứa trình tự nucleotit đối với chất chỉ thị chọn lọc, mà có thể được sử dụng để chọn lọc cây trồng, phần cây trồng hoặc tế bào cây trồng được biến nạp. Như được dùng trong bản mô tả này, "chất chỉ thị chọn lọc" có nghĩa là trình tự nucleotit mà khi được biểu hiện truyền kiểu hình khác biệt cho cây trồng, phần cây trồng hoặc tế bào cây trồng biểu hiện chất chỉ thị và do đó cho phép cây trồng, phần cây trồng hoặc tế bào cây trồng được biến nạp này được phân biệt với cây trồng, phần cây trồng hoặc tế bào cây trồng mà không có chất chỉ thị. Trình tự nucleotit có thể mã hóa hoặc là chất 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 chất chỉ thị chọn lọc thích hợp là đã biết trong lĩnh vực và có thể được sử dụng trong cat-xet biểu hiện mô tả ở đây.

Ví dụ về chất chỉ thị chọn lọc bao gồm, nhưng không giới hạn ở, trình tự nucleotit mã hóa cho *neo* hoặc *nptII*, mà mang lại tính kháng đối với kanamycin, G418, và dạng tương tự (Potrykus *et al.* (1985) *Mol. Gen. Genet.* 199:183-188); trình tự nucleotit mã hóa cho *bar*, mà mang lại tính kháng đối với phosphinothrixin; trình tự nucleotit mã hóa cho 5-enolpyruvylshikimat-3-phosphat (EPSP) syntaza được thay đổi, mà mang lại tính kháng đối với glyphosat (Hinchee *et al.* (1988) *Biotech.* 6:915-922); trình tự nucleotit mã hóa cho nitrilaza chẳng hạn như *bxn* từ *Klebsiella ozaenae* mà mang lại tính kháng đối với bromoxynil (Stalker *et al.* (1988) *Science* 242:419-423); trình tự nucleotit mã hóa cho axetolactat syntaza (ALS) được thay đổi mà mang lại tính kháng đối với imidazolinon, sulfonylure hoặc hóa chất ức chế ALS khác (đơn sáng chế EP số 154204); trình tự nucleotit mã hóa cho dihydrofolat reductaza (DHFR) kháng methotrexat (Thillet *et al.* (1988) *J. Biol. Chem.* 263:12500-12508); trình tự nucleotit mã hóa cho dalapon dehalogenaza mà mang lại tính kháng đối với dalapon; trình tự nucleotit mã hóa cho manosa-6-phosphat isomeraza (còn được gọi là phosphomanoza isomeraza (PMI)) mà mang lại khả năng chuyển hóa manosa (bằng sáng chế Mỹ số 5,767,378 và 5,994,629); trình tự nucleotit mã hóa cho anthranilat syntaza được thay đổi mà mang lại tính kháng đối với 5-metyl tryptophan; hoặc trình tự nucleotit mã hóa cho *hph* mà mang lại tính kháng đối với hygromycin. Người có hiểu biết trung bình trong lĩnh vực này có khả năng chọn chất chỉ thị chọn lọc thích hợp để dùng trong cat-xet biểu hiện của sáng chế.

Các chất 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 enzym 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 cây trồng (Dellaporta *et al.*, "Molecular cloning of the maize R-nj allele by transposon-tagging with Ac" 263-282 trong: *Chromosome Structure and Function: Impact of New Concepts*, 18th Stadler Genetics Symposium (Gustafson & Appels eds., Plenum Press 1988)); trình tự nucleotit mã hóa cho β -lactamaza, enzym mà các chất nền sinh màu khác nhau đối với

chúng đã được biết đến (ví dụ như, PADAC, xephalosporin sinh màu) (Sutcliffe (1978) *Proc. Natl. Acad. Sci. USA* 75:3737-3741); trình tự nucleotit mã hóa cho *xylE* mà mã hóa cho catechol dioxygenaza (Zukowsky *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:1101-1105); trình tự nucleotit mã hóa cho tyrosinaza, enzym có khả năng oxy hóa tyrosin thành DOPA và dopaquinon, mà đến lượt nó ngưng tụ để tạo thành melanin (Katz *et al.* (1983) *J. Gen. Microbiol.* 129:2703-2714); trình tự nucleotit mã hóa cho β -galactosidaza, enzym mà có các chất nền sinh màu đối với nó; trình tự nucleotit mã hóa cho luxiferaza (lux) mà cho phép phát hiện sự phát quang sinh học (Ow *et al.* (1986) *Science* 234:856-859); trình tự nucleotit mã hóa cho aequorin mà có thể được sử dụng trong sự phát hiện phát quang sinh học nhạy cảm với canxi (Prasher *et al.* (1985) *Biochem. Biophys. Res. Comm.* 126:1259-1268); hoặc trình tự nucleotit mã hóa cho protein huỳnh quang màu xanh lá (Niedz *et al.* (1995) *Plant Cell Reports* 14:403-406). Người có hiểu biết trung bình trong lĩnh vực này có khả năng chọn chất chỉ thị chọn lọc thích hợp để dùng trong cat-xet biểu hiện của sáng chế.

Ngoài ra, như đã biết rõ trong lĩnh vực, cây trồng chuyển gen nguyên vẹn có thể được tái tạo từ tế bào cây trồng, môi trường nuôi cấy mô cây trồng hoặc tế bào trần được nuôi cấy được biến nạp bằng cách sử dụng kỹ thuật bất kỳ trong nhiều kỹ thuật đã biết. Sự tái tạo cây trồng từ tế bào cây trồng, môi trường nuôi cấy mô cây trồng hoặc tế bào trần được nuôi cấy được mô tả, ví dụ, trong tài liệu Evans *et al.* (*Handbook of Plant Cell Cultures*, Vol. 1, MacMilan Publishing Co. New York (1983)); và Vasil I. R. (ed.) (*Cell Culture and Somatic Cell Genetics of Plants*, Acad. Press, Orlando, Vol. I (1984), và Vol. II (1986)).

Ngoài ra, các đặc điểm di truyền được thiết kế cho hạt và cây trồng, phần cây trồng, hoặc tế bào cây trồng chuyển gen theo sáng chế được mô tả ở trên có thể được truyền bằng cách sinh sản hữu tính hoặc sinh trưởng sinh dưỡng và do đó có thể được duy trì và được nhân lên ở cây trồng con cháu. 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 trồng, phần cây trồng hoặc tế bào cây trồng theo số lượng cách bất kỳ mà đã biết rõ 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 hệ 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 quan tâm 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.

Các phương án khác của sáng chế bao gồm sản phẩm thu hoạch được tạo ra từ cây trồng chuyển gen hoặc phần của chúng theo sáng chế, cũng như là sản phẩm đã được xử lý được 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 phương án khác, sáng chế đề xuất chất chiết từ hạt chuyển gen hoặc cây trồng chuyển gen theo sáng chế, trong đó chất chiết có chứa phân tử axit nucleic, polynucleotit, trình tự nucleotit hoặc protein gây 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 et al., *Food, Agric. Environ.* 2(1):84-89 (2004); Guidet, *Nucleic Acids Res.* 22(9): 1772-1773 (1994); Lipton et al., *Food Agric. Immun.* 12:153-164 (2000)).

Hợp phần diệt côn trùng

Theo một số phương án, sáng chế đề xuất hợp phần diệt côn trùng bao gồm protein Cry theo sáng chế trong chất mang nông dụng. Như được dùng trong bản mô tả này "chất mang nông dụng" có thể bao gồm nguyên liệu hữu cơ hoặc vô cơ, tự nhiên hoặc tổng hợp

mà được kết hợp với protein Cry hoạt tính để làm thuận lợi cho việc áp dụng nó cho hoặc trong cây trồng, hoặc phần của chúng. 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, 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 hợp phần nông nghiệp. Các hợp phần 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, hợp phần 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 phương án khác, cây trồng sản xuất protein Cry theo sáng chế ở thực vật là chất mang dùng trong nông nghiệp của protein Cry được biểu hiện.

Theo phương án khác, hợp phần diệt côn trùng có chứa 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 sản xuất protein Cry theo sáng chế. Hợp phần diệt côn trùng này có thể được điều chế bằng cách làm mất nước, đông khô, đông nhất hóa, chiết, lọc, ly tâm, lắng, hoặc cô môi trường nuôi cấy của *Bacillus thuringiensis* (*Bt*). Môi trường nuôi cấy *Bt* này có thể là các chủng *Bt* có trong tự nhiên YN171-1, GX078-2 và GX435-1 được mô tả dưới đây trong các ví dụ hoặc môi trường nuôi cấy *Bt* chuyển gen. Theo các phương án bổ sung, hợp phần bao gồm từ khoảng 1% đến khoảng 99% theo khối lượng của protein Cry theo sáng chế.

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

khác, chế phẩm bao gồm tác nhân kiểm soát loài gây hại thứ hai khi cây trồng chuyển gen có chứa 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 gồm có chất diệt loài gây hại hóa học, chẳng hạn 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 serin và loại cystein), lectin, alpha -amylaza, peroxidaza, cholesterol oxidaza và phân tử ARN sợi kép (dsARN).

Theo phương án khác, tác nhân kiểm soát loài gây hại thứ hai là chất diệt loài gây hại hóa học được chọn từ nhóm gồm có pyrethroid, carbamat, neonicotinoid, chất phong bế kênh natri thần kinh, lacton vòng lớn diệt côn trùng, chất đối kháng axit gamma-aminobutyric (GABA), ure diệt côn trùng và chất giả hormon kích sâu non. Theo phương án khác, chất diệt loài gây hại hóa học được chọn từ nhóm gồm có abamectin, axephat, axetamiprid, amidoflomet (S-1955), avermectin, azadirachtin, azinphos-methyl, bifenthrin, bifenazat, buprofezin, carbofuran, chlorfenapyr, chlorfluazuron, chlorpyrifos, chlorpyrifos-methyl, chromafenozit, clothianidin, xyfluthrin, beta-xyfluthrin, xyhalothrin, lambda-xyhalothrin, xypmethrin, xyromazin, deltamethrin, diafenthiuron, diazinon, diflubenzuron, dimethoat, diofenolan, emamectin, endosulfan, esfenvalerat, ethiprol, fenothicarb, fenoxycarb, fenpropathrin, fenproximat, fenvalerat, fipronil, flonicamid, fluxythrinate, tau-fluvalinat, flufenoxuron, fonophos, halofenozit, hexaflumuron, imidacloprid, indoxacarb, isofenphos, lufenuron, malathion, metaldehyd, methamidophos, methidathion, methomyl, methopren, methoxychlor, monocrotophos, methoxyfenozit, nithiazin, novaluron, noviflumuron (XDE-007), oxamyl, parathion, parathion-methyl, permethrin, phorat, phosalon, phosmet, phosphamidon, pirimicarb, profenofos, pymetrozin, pyridalyl, pyriproxyfen, rotenon, spinosad, spiromesifin (BSN 2060), sulprofos, tebufenozit, teflubenzuron, tefluthrin, terbufos, tetrachlorvinphos, thiacloprid, thiamethoxam, thiodicarb, thiosultap-natri, tralomethrin, trichlorfon và triflumuron, aldicarb, oxamyl, fenamiphos, amitraz, chinomethionat, clorobenzilat, xyhexatin, dicofol, dienochlor, etoxazol, fenazaquin, fenbutatin oxit, fenpropathrin,

fenpyroximat, hexythiazox, propargit, pyridaben và tebufenpyrad. Theo phương án khác nữa, chất diệt loài gây hại hóa học được chọn từ nhóm gồm có xypermethrin, xyhalothrin, xyfluthrin và beta-xyfluthrin, esfenvalerat, fenvalerat, tralomethrin, fenothicarb, methomyl, oxamyl, thiodicarb, clothianidin, imidacloprid, thiacloprid, indoxacarb, spinosad, abamectin, avermectin, emamectin, endosulfan, ethiprol, fipronil, flufenoxuron, triflumuron, điofenolan, pyriproxyfen, pymetrozin 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 phương án khác, tác nhân kiểm soát loài gây hại thứ hai là protein Cry được chọn từ nhóm gồm có 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 phương án khác, tác nhân kiểm soát loài gây hại thứ hai là protein diệt côn trùng ở thực vật Vip3 được chọn từ nhóm gồm có 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 trồng chuyển gen. Việc đồng biểu hiện của nhiều hơn một gốc diệt loài gây hại trong cùng cây trồng 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 trồng thứ hai, bố mẹ 2, có thể được thiết kế di truyền để biểu hiện của 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 phương án khác, sáng chế đề xuất cây trồng chuyển gen xếp chồng kháng với sự phá hoại cây trồng bởi loài gây hại có chứa trình tự ADN mã hóa cho dsARN để kìm hãm gen thiết yếu ở loài gây hại đích và trình tự ADN mã hóa cho protein Cry theo sáng chế thể hiện hoạt tính sinh học chống lại loài gây hại đích. Đã có báo cáo rằng dsARN không hữu hiệu để chống lại loài gây hại thuộc bộ cánh vảy nhất định (Rajagopol et al. 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 khi loài gây hại đích là loài gây hại thuộc bộ cánh vảy, protein Cry theo sáng chế tác động làm giảm tạm thời độ pH trung tràng mà đáp ứng làm ổn định dsARN được tiêu hóa đồng thời mang lại dsARN hữu hiệu trong việc làm câm gen đích.

Ngoài việc đề xuất hợp phần, sáng chế đề xuất phương pháp sản xuất protein Cry gây độc đối với loài gây hại thuộc bộ cánh vảy. Phương pháp này bao gồm bước, nuôi cấy tế bào chủ không phải ở người chuyển gen mà có chứa polynucleotit hoặc gen khảm hoặc phân tử axit nucleic hoặc vật truyền tải tổ hợp theo sáng chế trong điều kiện trong đó tế bào chủ sản xuất protein gây độc đối với loài gây hại thuộc bộ cánh vảy. Theo một số phương án, tế bào chủ chuyển gen không phải của người là tế bào của cây trồng. Theo một số phương án khác, tế bào cây trồng 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 nấm men.

Theo phương án khác của phương pháp này, côn trùng cánh vảy gây hại được chọn từ nhóm gồm có sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu đục quả bông (*Helicoverpa armigera*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), sâu bướm Phương Đông (*Mythimna sepatate*), sâu đục thân

ngô châu Âu (*Ostrinia nubilalis*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu bướm hại đậu nhung (*Anticarsia gemmatilis*), sâu đo đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu ngài đêm hại đậu phương tây (*Richia albicosta*), sâu ăn chồi thuốc lá (*Heliothis virescens*), sâu đục thân có sọc (*Chilo suppressalis*), sâu đục thân màu hồng (*Sesamia calamistis*) và sâu cuốn lá lúa (*Cnaphalocrocis medinalis*), và dạng kết hợp bất kỳ của chúng.

Theo các phương án khác của phương pháp, gen khảm bao gồm bất kỳ trong số các SEQ ID NO:1-5. Theo các phương án khác nữa, protein được tạo ra bao gồm trình tự axit amin nêu trong trình tự bất kỳ trong số các SEQ ID NO: 16-25.

Theo một số phương án của phương pháp, gen khảm có chứa trình tự nucleotit mà được tối ưu hóa bộ ba mã hóa để biểu hiện ở thực vật. Theo các phương án khác, gen khảm bao gồm bất kỳ trong số các SEQ ID NOs:6-10. 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 NOs:16-25.

Theo phương án khác, sáng chế đề xuất phương pháp sản xuất cây trồng chuyển gen kháng loài gây hại (ví dụ như, kháng côn trùng), bao gồm: bước đưa polynucleotit, gen khảm, vật truyền tái tổ hợp, cat-xet biểu hiện hoặc phân tử axit nucleic theo sáng chế có chứa trình tự nucleotit mà mã hóa cho protein Cry theo sáng chế vào cây trồng, trong đó trình tự nucleotit được biểu hiện trong cây trồng, bằng cách đó mang lại cho cây trồng tính kháng đối với côn trùng cánh vảy gây hại, và tạo ra cây trồng chuyển gen kháng côn trùng. Theo một số phương án, cây trồng chuyển gen kháng loài gây hại kháng với côn trùng cánh vảy gây hại thuộc chi *Ostrinia* khi so sánh với cây trồng đối chứng thiếu polynucleotit, gen khảm, vật truyền tái tổ hợp, cat-xet biểu hiện hoặc phân tử axit nucleic theo sáng chế. Theo phương án khác, côn trùng thuộc chi *Ostrinia* là sâu đục thân ngô châu Á (*Ostrinia furnacalis*) hoặc sâu đục thân ngô châu Âu (*Ostrinia nubilalis*). Theo một số phương án, việc đưa vào đạt được bằng cách biến nạp cây trồng. Theo các phương án khác, việc đưa vào đạt được bằng cách lai cây thứ nhất chứa gen khảm, vật truyền tái tổ hợp, cat-xet biểu hiện hoặc phân tử axit nucleic theo sáng chế với cây thứ hai khác.

Theo một số phương án, cây trồng chuyển gen theo sáng chế mà kháng với ít nhất là sâu đục thân ngô châu Á (*Ostrinia furnacalis*) hoặc sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) hoặc kháng thêm với ít nhất là một côn trùng cánh vảy gây hại khác, trong đó côn trùng cánh vảy gây hại khác bao gồm, nhưng không giới hạn ở, sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu bướm hại đậu nhưng (*Anticarsia gemmatilis*), sâu đo đậu tương (*Chrysodeixis includes*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu ngài đêm hại đậu phương tây (*Richia albicosta*), sâu ăn chồi thuốc lá (*Heliothis virescens*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân có sọc (*Chilo suppressalis*), sâu đục thân màu hồng (*Sesamia calamistis*) hoặc sâu cuốn lá lúa (*Cnaphalocrocis medinalis*), và dạng kết hợp bất kỳ của chúng.

Theo phương án khác, phương pháp kiểm soát côn trùng cánh vảy gây hại chẳng hạn như sâu đục thân ngô châu Á (*Ostrinia furnacalis*) hoặc sâu đục thân ngô châu Âu (*Ostrinia nubilalis*) được đề xuất, phương án này có chứa bước phân phối cho côn trùng lượng hữu hiệu 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. Các cách để phân phối protein theo đường miệng đến côn trùng bao gồm, nhưng không giới hạn ở, cung cấp protein (1) ở cây trồng chuyển gen, trong đó côn trùng ăn (ăn vào bụng) một hoặc nhiều phần của cây trồng chuyển gen, bằng cách đó tiêu hóa polypeptit mà được biểu hiện trong cây trồng chuyển gen; (2) trong (các) hợp phần protein được tạo chế phẩm mà có thể được áp dụng cho hoặc được kết hợp vào, ví dụ, môi trường sinh trưởng côn trùng; (3) trong (các) hợp phần protein mà có thể được áp dụng cho bề mặt, ví dụ, được phun, lên trên bề mặt của phần cây trồng, mà sau đó được ăn vào bụng bởi côn trùng khi côn trùng ăn một hoặc nhiều phần cây trồng được phun; (4) chất nền môi; hoặc (5) hệ thống phân phối protein khác bất kỳ đã biết 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 trồng 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 hợp phần bao gồm các protein Cry theo sáng chế. Phân phối hợp phần theo sáng chế lên bề mặt cây có thể được thực hiện bằng cách sử dụng phương pháp bất kỳ đã biết đối với người có hiểu biết trung bình trong lĩnh vực này để áp dụng hợp chất, hợp phần, 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, rễ, 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 người có hiểu biết trung bình trong lĩnh vực này.

Theo một số phương án, sáng chế bao hàm phương pháp cung cấp cho nông dân phương thức kiểm soát loài gây hại thuộc bộ cánh vảy, phương pháp này bao gồm bước cung ứng hoặc bán cho nông dân nguyên liệu cây trồng chẳng hạn như hạt, nguyên liệu cây trồng này có chứa polynucleotit, gen khảm, cat-xet biểu hiện hoặc vật truyền tái tổ hợp có khả năng biểu hiện protein Cry theo sáng chế ở cây trồng sinh trưởng từ hạt này, như được mô tả ở trên.

Các phương án của 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 của 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. Người có hiểu biết trung bình trong lĩnh vực này 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 chủng *Bacillus thuringiensis* (Bt) được phân lập từ mẫu đất thu được ở Trung Quốc. Mẫu đất được tạo huyền phù trong môi trường lỏng LB+2,5M natri axetat sau đó là

xử lý nhiệt ở 70°C trong thời gian 20 phút. Sau đó một microlit huyền phù được phết lên đĩa aga T3 + penixilin và ủ ở nhiệt độ 28°C đến khi tạo thành khuẩn lạc. Các khuẩn lạc có hình thái giống *Bacillus* được nhặt ra từ đĩa và được vạch lại trên đĩa aga T3 + penixilin đến khi chúng hình thành bào tử, thường là trong xấp xỉ ba ngày. Các chủng *Bt* được nhận diện bằng cách nhuộm màu môi trường nuôi cấy bằng xanh Coomassie/axit axetic và làm hiển thị bằng kính hiển vi. Sau khi hình thành bào tử, môi trường nuôi cấy được tạo huyền phù trong chất đệm kiểm và được thử nghiệm về hoạt tính chống lại các loài cánh vảy quan tâm. Các huyền phù được thử nghiệm trong thử nghiệm sinh học nhiễm bẹ mặt, trong đó dịch lỏng được phủ lên khẩu phần ăn nhân tạo nhiều loài. Mỗi thể phân lập được sàng lọc chống lại ít nhất là ba loài côn trùng cánh vảy gây hại, bao gồm sâu đục quả bông (*Helicoverpa armigera*), sâu ngài đêm màu đen (*Agrotis ipsilon*) và sâu đục thân ngô châu Á (*Ostrinia furnacalis*) với cỡ mẫu là 12 ấu trùng mới nở. Thời gian của mỗi thử nghiệm là khoảng 7 ngày ở nhiệt độ 27°C; các đĩa được tính điểm về tỷ lệ chết cũng như là sự ức chế sinh 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, ba chủng *Bt*, YN171-1, GX078-2 và GX435-1, được chọn để phân tích thêm.

Ví dụ 2. Lắp ráp và phân tích hệ 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 bằng cách 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/nổ: 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 Braxin, 28-57 Mỹ, Vương Quốc Anh và Thụy Sĩ) được nối lại sử dụng NEB Quick Ligation™ như được mô tả bởi nhà cung cấp (New England Biolabs, Inc. Ipswich, MA). Việc nối được làm sạch bằng cách 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 50µl được trộn với 45µl hỗn hợp hạt 75% (25% hạt AMPure cộng thêm dung dịch NaCl/PEG 75% TekNova cat #

P4136). Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi bề mặt thu được được chuyển vào giếng mới và 45µl hỗn hợp hạt 50% (50% hạt AMPure cộng với 50% dung dịch NaCl/PEG TekNova cat # P4136) được bổ sung. Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi thu được được loại ra và các hạt được rửa bằng 80% etanol. 25µl 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 1,5ml. 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, 45 giây]; 12 x [98°C, 15 giây, 60°C, 30 giây, 72°C, 30 giây]; [72°C, 1 phút]. Mỗi hỗn hợp phản ứng chứa: 5µl thư viện ADN, 1µl đoạn môi thông thường Bioscience (25µM), 18µl nước vô trùng, 1µl đoạn môi được đánh số thứ tự Bioscience (25µM), 25µl 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 xác định trình tự các đầu bắt cặp (PE) (100 chu kỳ mỗi lần đọc; 12-24 thư viện mỗi lần) trên hệ xác định trình tự HiSeq 2500 sử dụng nguyên tắc xác định trình tự của nhà sản xuất (Illumina, Inc., San Diego, CA).

Công cụ phân tích vi tính *Bacillus* được phát triển để nhận dạng và xác định đặc điểm gen giống Cry được sử dụng để sắp xếp thứ tự ưu tiên của sự dẫn đầu để thử nghiệm thêm trong phòng thí nghiệm.

Sự lắp ráp và phân tích hệ gen cũng như là phân tích thư viện hệ gen được mô tả ở dẫn đến sự nhận diện của ba gen giống Cry trong các chủng *Bt* có độc tính đối với ít nhất là một côn trùng cánh vảy gây hại, ví dụ, sâu đục thân ngô châu Á (*Ostrinia furnacalis*). Việc xác định đặc điểm của gen giống Cry và protein được thể hiện trong bảng 1.

Bảng 1. Gen/protein Cry được xác định trong các chủng *Bacillus thuringiensis*

Chủng	Protein/ Tên gen	Thành viên họ Cry gần nhất	% Độ tương đồng	Khối lượng phân tử (kD)	Nucleotit SEQ ID NO:	Axit amin SEQ ID NO:
YN171	BT0049	Cry20Ba	68%	89,5	1	16
GX078	BT0063	Cry9Db1	91%	132,2	2	17
GX078	BT0064	Cry2Ae	75%	69,8	3	18
GX078	BT0086	Cry2Ah	85%	71,4	4	19
GX435	BT0498	Cry1Ga	89%	132,1	5	20

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 trong chủng *Bacillus thuringiensis* (*Bt*) không có tinh thể không có hoạt tính diệt côn trùng nên có thể quan sát được thông qua vật truyền con thoi được thiết kế pCIB5634', được thiết kế để biểu hiện ở cả *E. coli* và *Bt*. Vật truyền pCIB5634' có chứa promoter Cry1Ac mà điều khiển sự biểu hiện của gen Cry *Bt* đã tách dòng và chất chỉ thị tính kháng erythromyxin. Cat-xet biểu hiện có chứa trình tự mã hóa Cry được quan tâm được biến nạp vào chủng *Bt* vật chủ thông qua đục lỗ điện và các chủng *Bt* chuyển gen được chọn trên đĩa aga có chứa erythromyxin. Các chủng *Bt* chuyển gen được chọn được cho sinh trưởng đến pha tạo bào tử trong môi trường T3 ở nhiệt độ 28°C trong thời gian 4-5 ngày. Các viên tế bào được thu và rửa đi rửa lại trước khi hòa tan trong đệm cacbonat độ pH cao (50mM) chứa 2mM DTT.

Biểu hiện *E. coli*. Protein Cry được biểu hiện ở các chủng *E. coli* bằng cách sử dụng các vật truyền pET28a hoặc pET29a (Merck KGaA, Darmstadt, Đức). Cấu trúc được biến nạp bằng cách đục lỗ điện và các dòng *E. coli* chuyển gen được chọn trên đĩa

aga có chứa kanamycin. Các chủng *E. coli* chuyển gen được chọn được cho sinh trưởng và sự biểu hiện protein Cry được gây ra bằng cách sử dụng sự cảm ứng IPTG ở nhiệt độ 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 Microfluidics LV-1.

Phân tích biểu hiện. Sau đó dịch tan tế bào thu được từ các chủng *Bt* hoặc *E. coli* chuyển gen được phân loại thông qua ly tâm và các mẫu được phân tích về độ tinh sạch thông qua SDS-PAGE và điện di đồ bằng cách sử dụng hệ thống BioRad Experion (Biorad, Hercules, CA). Tổng nồng độ protein được xác định qua thử nghiệm Bradford hoặc Thermo 660. Sau đó protein Cry đã tinh chế được thử nghiệm trong các thử nghiệm sinh học được mô tả dưới đây.

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

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 châu Á (ACB; *Ostrinia furnacalis*); sâu đục quả bông (CBW; *Helicoverpa armigera*); sâu ngài đêm màu đen (BCW; *Agrotis ipsilon*); bướm đêm hại đào màu vàng (YPM; *Conogethes punctiferalis*); sâu bướm Phương Đông (OAW, *Mythimna separate*); sâu đục thân có sọc (SSB; *Chilo suppressalis*); sâu đục thân ngô châu Âu (ECB; *Ostrinia nubilalis*); sâu xanh mùa thu (FAW; *Spodoptera frugiperda*); sâu đục trái trên cây ngô (CEW; *Helicoverpa zea*); sâu đục thân mía (SCB; *Diatraea saccharalis*); sâu đo đậu tương (SBL; *Pseudoplusia includens*) và sâu ăn chồi thuốc lá (TBW; *Heliothis virescens*).

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 đĩa 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 lộ ra với chủng *Bacillus* kiểu đại có hoạt tính mạnh và phổ rộng. Nhóm đối chứng âm bao gồm ấu trùng được lộ ra với 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 xấp xỉ 120 giờ và tính điểm tương đối so với đối chứng.

Kết quả được thể hiện trong bảng 2, trong đó "-" có nghĩa là không có hoạt tính so với nhóm đối chứng, "+/-" có nghĩa là hoạt tính 0-10% so với dấu kiểm tra (loại này cũng bao gồm tỉ lệ chết 0% với sự ức chế sinh trưởng ấu trùng mạnh), "+" có nghĩa là hoạt tính 10-25% so với dấu kiểm tra, "++" có nghĩa là hoạt tính 25-75% so với dấu kiểm tra, và "+++" có nghĩa là hoạt tính 75-100% so với dấu kiểm tra. Chữ trong ngoặc đơn chỉ ra sự kìm hãm ấu trùng quan sát được (sự ức chế sinh trưởng) như sau: "(l)" có nghĩa là kìm hãm nhẹ; "(m)" có nghĩa là kìm hãm trung bình và "(s)" có nghĩa là kìm hãm mạnh. Ký hiệu "nt" trong bảng 2 có nghĩa là protein được chỉ ra không được thử nghiệm chống lại loài gây hại cụ thể đó.

Bảng 2. Kết quả của thử nghiệm sinh học với protein Cry

Côn trùng	Các protein BT				
	0049	0063	0064	0086	0498
ACB	++(l)	+++	+++	-	++(m)
ECB	-	+++ (s)	-(s)	++	+++
BCW	+	+/-	+	+/-	không kiểm tra
CBW	-	-	-	-	+(l)
OAW	không kiểm tra	+++	++(m)	+	không kiểm tra
YPM	-	+(s)	+++	-(l)	++(m)
SCB	-	-	-(s)	không kiểm	+/-

				tra	
SWCB	-	+/-	-	không kiểm tra	+/-
FAW	-	-	-	+/-	-
CEW	+/-	-	-	-	+/-
SBL	-	+	++	không kiểm tra	không kiểm tra

Bướm đêm hại đào màu vàng (YPM; *D. punctiferalis*; đồng nghĩa với *Conogethes punctiferalis*) có thể gây thiệt hại nghiêm trọng cho nhiều cây trồng khác nhau, bao gồm ngô, cao lương hạt, đậu tương, bông, đào và cây cam chanh, dẫn đến thiệt hại kinh tế nặng nề. Nó phân bố rộng từ Ấn Độ và Pakistan qua Tây Nam Á đến Úc và thâm nhập vào các quốc gia khác qua việc chuyên chở trái cây, bao gồm Anh và Hà Lan. Bướm đêm hại đào màu vàng còn được gọi là ngài đốm hại ngô ở Úc và sâu bướm hại hạt thầu dầu ở Ấn Độ. YPM được kiểm soát chủ yếu bằng chất diệt vật gây hại hóa học nên có nhu cầu về các sản phẩm kiểm soát sinh học. Có một số ít các báo cáo về sự nhạy cảm YPM đối với protein Cry Bt. Do đó, phát hiện về protein Cry mới có độc tính cao hơn đối với YPM là cần thiết để mang lại phương thức hữu dụng hơn để kiểm soát các loài vật gây hại phá hoại này ở Trung Quốc và các quốc gia khác. Kết quả được thể hiện ở trên chỉ ra rằng BT0064, thuộc họ Cry2A, rất hoạt tính để chống lại YPM và gây ra sự ức chế sinh trưởng đáng kể.

Ngạc nhiên là, BT-0049, mà có sự ăn khớp trình tự gần nhất với các thành viên của họ Cry20, có hoạt tính chống lại các côn trùng cánh vảy gây hại nhất định, cụ thể là sâu đục thân ngô châu Á (*Ostrinia furnacalis*). Điều này được tin là báo cáo đầu tiên về protein liên quan đến Cry20 có hoạt tính chống lại côn trùng cánh vảy gây hại. Protein Cry20 được biết đến trong lĩnh vực về hoạt tính chống lại côn trùng hai cánh, ví dụ muỗi

(Xem tài liệu ví dụ, Ibrahim *et al. supra.*, Konecka *et al.* 2014. Biol. Lett. 5:83-92; và Gama & Nakagoshi, 2014. Int. J. Curr. Microbiol. App. Sci. 3:179-197). Protein Cry Bt khác được phân lập như mô tả ở trên, BT-0211 được thiết kế (SEQ ID NO:26), cũng như là thành viên của họ Cry20, không có hoạt tính chống lại sâu đục thân ngô châu Á (*Ostrinia furnacalis*) khi được thử nghiệm trong cùng điều kiện như mô tả ở trên. Điều này là chỉ thị khác của phổ hoạt tính duy nhất của protein BT-0049 Cry. Bảng 3 thể hiện ma trận tỉ lệ phần trăm tương đồng giữa BT-0049, BT-0211 và các protein Cry20 đã biết. Các số ở trong ngoặc đơn sau tên của protein Cry20 đã biết là mã số truy cập GenBank. Dựa trên độ tương đồng được thể hiện trong ma trận này và phương pháp gọi tên protein Cry mới theo Crickmore *et al.* (1998. Microbiol. Molec. Biol. Rev. 62:807-813), BT-0211 có khả năng là Cry20Bb2 và BT-0049 có khả năng là protein Cry20 mới mà sẽ được chỉ định là Cry20Ca1.

Bảng 3. Ma trận tỉ lệ phần trăm độ tương đồng đối với các thành viên họ Cry20

Protein Cry Bt	0049	0211	Cry20Aa1	Cry20Ba1	Cry20Ba2	Cry20Bb1
BT-0049	-	67	63	68	67	68
BT-0211	67	-	67	71	71	96
Cry20Aa1(AAB93476)	63	67	-	64	64	68
Cry20Ba1(ACS93601)	68	71	64	-	97	72
Cry20Ba2(KC156694)	67	71	64	97	-	72
Cry20Bb1(KC182372)	68	96	68	72	72	-

Sự sắp xếp thẳng hàng trình tự giữa protein BT-0049 có hoạt tính với côn trùng cánh vảy và protein BT-0211 không có hoạt tính với côn trùng cánh vảy (Hình 1) chỉ ra rằng có hai vùng đáng chú ý của sự khác biệt axit amin có mặt trong protein BT-0049 mà có thể chịu trách nhiệm hoàn toàn hoặc một phần đối với sự khác biệt phổ hoạt tính giữa hai protein này. Vùng thứ nhất là từ axit amin 674 đến axit amin 709 của BT-0049 (SEQ ID NO:16) gồm có trình tự SVYHPDANDAYDEGDDNTYNQNYNSMY

NQNADNTYD (SEQ ID NO:27). Vùng thứ hai là từ axit amin 730 đến axit amin 746 của SEQ ID NO:16 gồm có trình tự GYEQNNYDSDDSDYNNT (SEQ ID NO:28).

Ví dụ 5. Tạo vật truyền của gen để biểu hiện ở thực vật

Trước khi biểu hiện ở cây trồng, các polynucleotit mã hóa cho các protein Cry, BT-0049 (SEQ ID NO:16), BT-0063 (SEQ ID NO:17), BT-0064 (SEQ ID NO:18), BT-0086 (SEQ ID NO:19) hoặc BT-0498 (SEQ ID NO:20) hoặc protein Cry đột biến, mBT-0049 (SEQ ID NO:21), mBT-0063 (SEQ ID NO:22), mBT-0064 (SEQ ID NO:23), mBT-0086 (SEQ ID NO:24) hoặc mBT-0498 (SEQ ID NO:25) được tổng hợp trên nền tảng tổng hợp gen tự động (ví dụ, Genscript, Inc., Piscataway, NJ). Đối với ví dụ này, cat-xet biểu hiện thứ nhất được tạo ra có chứa promotor có thể biểu hiện ở thực vật được liên kết có điều khiển với trình tự mã hóa cho protein Cry mà được liên kết có điều khiển với vùng kết thúc và cat-xet biểu hiện thứ hai được tạo ra có chứa promotor có thể biểu hiện ở thực vật được liên kết có điều khiển với gen đánh dấu chọn lọc mà được liên kết có điều khiển với vùng kết thúc. Sự biểu hiện của gen đánh dấu chọn lọc cho phép nhận diện cây trồng chuyển gen trên môi trường chọn lọc. Cả hai cat-xet biểu hiện được tách dòng vào vật truyền thích hợp để biến nạp lúa hoặc ngô qua trung gian *Agrobacterium*.

Ví dụ 6. Sự biểu hiện và hoạt tính của protein Cry ở cây ngô

Biến nạp của các phôi ngô chưa chín được thực hiện chủ yếu như được mô tả trong Negrotto và cộng sự, 2000, Plant Cell Reports 19: 798-803. Tóm lại, chủng *Agrobacterium* LBA4404 (pSB1) có chứa vật truyền biểu hiện được mô tả trong ví dụ 5 được cho sinh trưởng trên YEP (dịch chiết nấm men (5g/l), pepton (10g/l), NaCl (5g/l), 15g/l aga, độ pH 6,8) môi trường rắn trong thời gian 2-4 ngày ở nhiệt độ 28°C. Khoảng $0,8 \times 10^9$ tế bào *Agrobacterium* được tạo huyền phù trong môi trường LS-inf được bổ sung thêm 100µM As. Vi khuẩn được cảm ứng trước trong môi trường này trong khoảng 30-60 phút.

Phôi chưa trưởng thành từ dòng ngô lai cùng dòng được thực hiện từ các tai 8-12 ngày tuổi trong LS-inf lỏng + 100µM As. Phôi được rửa một lần bằng môi trường gây nhiễm mới. Sau đó dung dịch *Agrobacterium* được bổ sung vào và phôi được tạo xoáy trong 30 giây và cho xử lý bằng vi khuẩn trong 5 phút. Sau đó phôi được chuyển hướng

phía hình vảy lên vào môi trường LSAs và nuôi cấy trong tối trong từ hai đến ba ngày. Sau đó, khoảng từ 20 đến 25 phôi mỗi đĩa petri được chuyển vào môi trường LSDc được bổ sung xefotaxim (250mg/l) và bạc nitrat (1,6mg/l) và được nuôi cấy trong tối ở khoảng 28°C trong 10 ngày.

Các phôi chưa trưởng thành, tạo ra các thể sần 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 sần sống sót được chuyển sang môi trường Reg1 được bổ sung manozơ. Sau khi nuôi cấy trong sáng (chế độ 16 giờ sáng/ 8 giờ tối), sau đó mô màu xanh được chuyển vào môi trường Reg2 không có chất điều hòa sinh 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 trồng được thử nghiệm về sự có mặt của gen đánh dấu chọn lọc và gen *cry Bt* bằng 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 trồng chuyển gen biểu hiện BT-0049, BT-0063, BT-0064, BT-0086, BT-0498, mBT-0049, mBT-0063, mBT-0064, mBT-0086 hoặc mBT-0498 được thử nghiệm chống lại một hoặc nhiều côn trùng cánh vảy gây hại chẳng hạn như sâu đục thân ngô châu Á (*Ostrinia furnacalis*). Kết quả của thử nghiệm sinh học mô cây trồng chuyển gen sẽ xác nhận rằng protein Cry theo sáng chế khi biểu hiện ở cây trồng chuyển gen gây độc đối với một hoặc nhiều côn trùng cánh vảy gây hại đích.

YÊU CẦU BẢO HỘ

1. Phân tử axit nucleic bao gồm trình tự nucleotit mã hóa protein Cry của *Bacillus thuringiensis* mà độc đối với loài gây hại thuộc bộ cánh vảy, trong đó protein Cry bao gồm SEQ ID NO: 18.
2. Phân tử axit nucleic theo điểm 1, trong đó trình tự nucleotit này bao gồm SEQ ID NO: 3 hoặc SEQ ID NO: 8.
3. Gen khảm bao gồm phân tử axit nucleic theo điểm bất kỳ trong số các điểm 1 hoặc 2 và promotor khác loại, trong đó promotor khác loại này được liên kết có điều khiển với phân tử axit nucleic.
4. Gen khảm theo điểm 3, trong đó promotor khác loại là promotor có thể biểu hiện ở cây.
5. Gen khảm theo điểm 4, trong đó promotor có thể biểu hiện ở cây được chọn từ nhóm của các promotor gồm có các promotor ubiquitin, virus cestrum yellow, TrpA ở ngô, OsMADS 6, histon H3 ở ngô, UTR 5' gen 9 thực khuẩn thể T3, sucroza synthetaza 1 ở ngô, rượu dehydrogenaza 1 ở ngô, phức hệ thu nhận ánh sáng ở ngô, protein sốc nhiệt ở ngô, mtl ở ngô, RuBP carboxylaza tiểu đơn vị nhỏ ở đậu Hà Lan, actin ở lúa, xyclophilin ở lúa, mannopin syntaza plasmit Ti, nopalin syntaza plasmit Ti, chalcon isomeraza ở dạ yên thảo, protein giàu glyxin 1 ở đậu, patatin ở khoai tây, lectin, RuBP carboxylaza tiểu đơn vị nhỏ CaMV 35S và S-E9.
6. Protein Cry mà gây độc đối với loài gây hại thuộc bộ cánh vảy, trong đó protein Cry có chứa SEQ ID NO: 18.
7. Chủng *Bacillus thuringiensis* được phân lập mà sản xuất protein Cry theo điểm 6.
8. Hợp phần diệt côn trùng bao gồm protein Cry theo điểm 6 và chất mang nông dụng.

9. Hợp phần theo điểm 8, trong đó hợp phần này được điều chế bằng cách làm mất nước, đông khô, đồng nhất hóa, chiết, lọc, ly tâm, lắng, hoặc cô môi trường nuôi cấy của chủng *Bacillus thuringiensis*.

10. Hợp phần theo điểm 9, trong đó chủng *Bacillus thuringiensis* là chủng *Bacillus thuringiensis* theo điểm 7.

11. Hợp phần theo điểm 8, trong đó hợp phần có chứa tế bào vi khuẩn chuyển gen mà sản xuất protein Cry.

12. Vật truyền tái tổ hợp có chứa phân tử axit nucleic theo điểm 1 hoặc gen khảm theo điểm 3.

13. Phương pháp tạo ra cây chuyển gen kháng côn trùng, bao gồm: bước đưa gen khảm theo điểm 4 vào cây, trong đó protein Cry được biểu hiện trong cây, bằng cách đó tạo ra cây chuyển gen kháng côn trùng, và trong đó côn trùng được chọn từ nhóm gồm có sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu bướm Phương Đông (*Mythimna sepatate*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), và sâu đo đậu tương (*Chrysodeixis includes*).

14. Phương pháp theo điểm 13, trong đó bước đưa vào đạt được bằng cách biến nạp cây hoặc bằng cách lai chéo cây thứ nhất bao gồm gen khảm với cây thứ hai khác.

15. Phương pháp theo điểm 13 hoặc 14, trong đó cây là cây lúa hoặc cây ngô.

16. Phương pháp kiểm soát loài gây hại thuộc bộ cánh vảy, bao gồm bước phân phối cho loài gây hại thuộc bộ cánh vảy hoặc môi trường của chúng hợp phần bao gồm lượng hữu hiệu của protein Cry theo điểm 6 hoặc hợp phần diệt côn trùng theo điểm 8, trong đó loài gây hại thuộc bộ cánh vảy được chọn từ nhóm gồm có sâu đục thân ngô châu Á (*Ostrinia furnacalis*), sâu ngài đêm màu đen (*Agrotis ipsilon*), sâu bướm Phương Đông (*Mythimna sepatate*), bướm đêm hại đào màu vàng (*Conogethes punctiferalis*), và sâu đo đậu tương (*Chrysodeixis includes*).

1/1

BT-0049pro	1 mnpynkndheilnatqnnrcmpncysryplsndpqiiscqnnhykdwnlkyenitttssl
BT-0211pro	1 .s.--.n.tmv.spe.--ys.ivnk..yasn.n.em..mn....mag..ef-ap...
BT-0049pro	61 sgiltsgmfnkvmltgvlgvtpiviniildmvrldmgtgddllahteqliqqtlaaq
BT-0211pro	55 .l.s...il.q.it.....k..ei...vqe....rgn..n...v.v.....t.
BT-0049pro	121 yrsaatgaiygisrayenylmffkqwerntpqngqqvesaftvtlctalapqalls
BT-0211pro	115nd..k..r.....n.....s..
BT-0049pro	181 rrgfetlllpnyalaanfhlldrdaviyrnswlpsnvsikswnltlirnaiseyrnhcn
BT-0211pro	175t.....l..tq..ansm.trni.iq..tr..d.....
BT-0049pro	241 hwyndglnrfarssfnvfrfnayrrdmtlsvldfitvftypdiipkatnieltrivy
BT-0211pro	235 y...i.....t.t.....v.....r..rp..v.....
BT-0049pro	301 tnpipprgfprrtpptfnqmenliisgcpsflnqlriyttfthdphnisdfrwagngf
BT-0211pro	295 .d..s.....s.....g.....hvy..y....r.y
BT-0049pro	361 lsngtirtsgalagrriidfsmqnidvfrvnmthdidslsrsyggvhrsdfigniqnng
BT-0211pro	355if.q..st.pwstnip.....i...l.....mq.....ti...
BT-0049pro	421 gtsilynvpldttrfmrtnetaflpgdsglepnernythrlfqvmttyrsnpnarraaf
BT-0211pro	415 r.tlf.hqnv.n.sl.l....v.....q.....t.....
BT-0049pro	481 hawthrsllrrngfradqiiqipavkaintsddsavvsnvgenmmkldnltslsykla
BT-0211pro	475t...m.....t.s.g..r...l.y...i.....sg....v..
BT-0049pro	541 edseasntfriviriyasmnnknlnlilngtqiaslnvegtmqngg-sltldlqyenfkya
BT-0211pro	535 t..a.....v.....d.....v...a.....r.v....e...a..c.d...
BT-0049pro	600 afasnfkigsqsiigifktsnadfildkielipvdifipllkqtqdcnnyngdtiycnq
BT-0211pro	595 t..gd.qm....l....di.....sh.mss.e....dys...n.....
BT-0049pro	660 gydayepnasdsydsvyhpdandaydegddntynqnymsynqnadntyqsynnydpnm
BT-0211pro	655 ..gt.dh.s.nm..-----h.n.k..tqd.
BT-0049pro	720 entyqpsydngyeqnnysdsdsydnntynqntnmyngqynndnpnngctckqcyngny
BT-0211pro	680 dt....d...-----ntd..dsg...sq.--t....n.g..n.f
BT-0049pro	780 pk
BT-0211pro	721 ..

Hình 1

DANH MỤC TRÌNH TỰ

- <110> Syngenta Participations AG
 Mei, Wenqian
 Luo, Chunping
 Duan, Qianqian
 Wang, Fenshan
 Wang, Xiaojun
 Changming, Ning
- <120> Phân tử axit nucleic, protein và hợp phần diệt côn trùng và phương pháp để kiểm soát loài gây hại cây
- <130> 80982-WO-REG-ORG-P-1
- <160> 28
- <170> PatentIn phiên bản 3.5
- <210> 1
 <211> 2343
 <212> ADN
 <213> Bacillus thuringiensis
- <400> 1
 atgaatcctt atcagaataa gaatgacat gaaattttta atgccacaca aaataattgt 60
 cgcatgccta attgttattc aagatatcca ttatcaaag atccacaaat atcctgtcag 120
 aataaccatt ataaagattg gctgaataag tatgaaaata ttactactac atcttcttta 180
 tctggaatth taacttcaat aggaatgttc aataaggtac tcatgttaac cggagtatta 240
 ggcgtcacac caatagttat taacataata ctagatatgg tgagattaat tgatatggga 300
 acaggcgatg atttattagc tcatacagaa caacttattc aacaaacact agcagcacaa 360
 tatagaagtg ctgccactgg agccatatat ggtatatcta gagcatacga aaattatttg 420
 atgtttttta agcaatggga acgtaataga actcctcaa atggacaaca agtagagagt 480
 gctttttacta ctgttaatac tttatgtatt actgccctag ctctcaggc tttactttct 540
 cgtagaggat tcgaaacact tttattacca aactatgcac tagcggcaaa tttccatttg 600
 ttattattaa gagatgccgt tatttataga aacagctggg tgcctagtaa tgtttcaata 660
 aaaagtgtta atcttacaat attaagaaat gccataagtg aatatcgtaa tcattgtaat 720
 cattgggtata atgatgggtt aaatagattt gcccgctcat cttttaatga ttgggttcgt 780
 tttaatgcct atcgtagaga tatgacgtta tccgtattag atttcattac agtggtcccc 840
 acctatgacc caattatata tccaaaagca acaaatattg aattaactag aattgtgtat 900
 acaaatccaa taattccacc tagaggattt cctagaacta atcctcctac ttttaatcag 960

```

atggaaaatt taattatttc cggttgccct agttttttga atcaattgcg gatataataca 1020
actttttatc atgacccgca taatatatca agagattttt gggctggtaa tcaaaatttt 1080
ttaagtaatg gtaccattcg tacttctgga gctcttgccg gaagacgaat cgatttttagt 1140
atgcaaaaata ttgatgtctt cagggtgaat atgactactc atgacattga ttctttatca 1200
agaagttacg gtggagttca tagatctgat tttattggtg taaatataca aaataatcaa 1260
ggaacatcaa tattatataa tgtgcctctc gatactactc gatttatgac tcgcaatgaa 1320
acagcatttt tacccgga tcttggtg gagcccaatg aacgtaatta tactcatagg 1380
ttatttcaag tgatgacaac atatcgttct aatccaaatg ctgtagagc agcattttta 1440
cacgcatgga cgcatagaag tttaagacgc cgaaatggat ttagagcgga tcagattata 1500
caaatacccg ccgtaaaagc aataaatact agtgatgatt ctgctgtagt atcaaatggt 1560
ggagaaaata tgatgaaatt agataactta actacaagtt tctcctataa attaacagcg 1620
gaggattccg aagcatcgaa tacacgtttt atagtgcgta ttcgttatgc tagtatgaac 1680
aataataaat tgaatctaatt tttaaattgt actcagatag catcgctgaa tgtggaaggt 1740
acaatgcaaa atggcgcatc attaacagat cttcaatatg aaaattttta atatgccgca 1800
tttgcaagta atttcaagat aggttctcag tctataatag ggatttttta agagacatct 1860
aatgcagatt ttattctaga taaaattgaa ttgattccag tggatttcat accattatta 1920
aagcaaacac aagattgtaa caactataat caagatacta tttatacatg taaccaagga 1980
tatgatgctt acgaaccgaa tgctagtgat tcgtatgata gcgtttatca tccagatgct 2040
aatgacgcgt atgacgaggg ggacgacaat acttacaatc aaaactataa tagtatgtac 2100
aatcagaatg ctgataatac gtatgaccaa agttacaaca attacgatcc aaacatggag 2160
aatacgtatc aaccaagcta tgacaatggc tatgaacaaa ataactatga tagtgatgat 2220
tcagattata acaatactta caatcaaaat actaataata tgtacaatca agaataataat 2280
aataacgata atccaaacaa tggctgtacg tgcaagcaat gttacaacca gaactatcct 2340
aaa 2343

```

```

<210> 2
<211> 3510
<212> ADN
<213> Bacillus thuringiensis

<400> 2

```

atgaatcgaa ataatcaaaa tgaatatgaa gttatcgatg cccccattg tgggtgtccg	60
tcagaggata tcgtaaaata tcctctgaca gatgatccga atgctggatt gcaaaatatg	120
aactataagg aatatttaca aatgcatggt ggagactata cagatcctct tatcaatcct	180
aacttatctg ttagtggaag agatgtaata caagttggaa ttaatatgt aggaagatta	240
ctaagctttt ttggattccc tttttctagt caatgggtta ctgtatatag ttatctttta	300
aacagcttgt ggccgatga cgagaattct gtatgggatg cttttatgga gagagtagaa	360
gaacttattg atcaaaaaat ctcaacatta gtaaagggtta acgcattgga taacctaact	420
ggattacaaa ataattataa tttatatgta gaagcattag atgagtggct gaatagaccg	480
aatggggcaa gggcagcctt agtttctcag cgattttaca ttttagatag cctatttacc	540
caatttatgc caagctttgg ctctggttcc gggagtcaaa attatgcaac tatattactt	600
gcagtatatg cacaagcagc aaaccttcat ttgttattat taaaagatgt agaaatttat	660
ggagctagat gggggctgaa tcaaactcaa atagatcaat tccattctcg tcaacaaagc	720
cttactcaga cttatacaaa tcattgtgtt actgcgata atgatggatt agcgggaatta	780
agaggcacia gcgttgagag ttggctcaaa tatcatcaat accgtaggga aatgacagta	840
acggcaatgg atttagtggc attattccca tactataatg ttcgacaata tccaaatggg	900
gcaaatccac aacttacacg tgaggatatat acagatccaa tcgtatttaa tctacctgaa	960
cgtccaagtg gcgctttctg cgaaagtttt tatactatcc gagcggctcg agaacgttta	1020
actttttcgc aacttgaaaa tgcaataatt cgtccaccgc gcttgtttga aaggtttcaa	1080
gcttttagga tttatacagg cgaggcgga ctgaatgcaa atagtgtcc aatgaactat	1140
tggattggac attttataag aaatacacgt ttgggtgact caacaacaat tactacaaat	1200
tacggaacag tcaataatcg tttaactaac ttcatctctc ctactaccag tgatgtttat	1260
cagatcaatt caatctcaag taatttagcc gctatttttag gcactttata tggggttact	1320
agagcacaat tctattatgg atcaggaatt atttggctgt atgtcggaca aaatagcgtt	1380
cttgacaaat gtcatacaaa ctataattca atagaagaat taccaaacca aagcgatgaa	1440
cctacagtta gaagtatatag ccatagatta tctcatatca cctcttttaa tttcaatgta	1500
cagcttaata atcctgtaat ttctgctggc aatatgcctg tatatgtgtg gacacatcgc	1560
agtgtggacc ttaataacac gattacttct gatagaatta ctcaattacc agcagtaaag	1620
gcgtttgagc taagcgatgg ggctactatt gtaaaaggcc caggatttac agggggagat	1680

gtaatccgaa gaacaaatca aggtgggttc ggaagtataa gagtgtcggg tactggcccc	1740
ctaactcaac gatatcgcat aaggttccgt tatgcttcga caatagattt tgatttcttt	1800
gtaacacgtg gaggaactac tataaataat tttagattta cacgtacaat gaacagcgga	1860
caggaatcaa gatatgaatc ctatcgctact gtagagtta caactccttt taactttaca	1920
caaagtcaag atataattcg aacatctatc cagggactta gtggaaatgg ggaagtatac	1980
cttgatagaa ttgaaatcat ccctgtaaat ccaacacgag aagcggaaga ggatttagaa	2040
gcagcgaaga aagcgggtggc gagcttggtt acacgtacta gagatggatt gcaggtaa	2100
gtgacagatt atcaagtcga tcaagcggca aatttagtgt cgtgcttatac agatgaacaa	2160
tatgggcatg ataaaaagat gttattggaa gccgtacgcg cggcaaagcg tctcagccgc	2220
gaacgtaatt tgcttcagga cccagatttt aatgctataa atagcacaga agaaaatggc	2280
tggaaggcaa gtaacggcgt tactattagc gagggcggtc cattctttaa aggccgtgca	2340
cttcaattag caagcgcaag agagaattat ccaacataca tttatcaaaa ggtagatgca	2400
tccgtgttaa agccttatac acgctataga ctagatggat ttgtgaagag tagtcaaat	2460
ttagaaattg atctcattca ccatcataaa gtccatcttg taaaaaatgt accagataat	2520
ttagtatctg atacttactc agatggttct tgcagcgga ttaatcgttg tgatgaacag	2580
catcaggtag atatgcagct agatgcggag catcatcaa tggattgctg tgaagcggct	2640
caaacacatg agttttcttc ttatattaat acaggggatc taaattcaag tgtagatcaa	2700
ggtatttggg ttgtattgaa agttcgaaca acagatggtt atgcaacgct gggaaatctt	2760
gaattggtag aagttggacc attatcaggt gaatccctag aacatgaaaa aagaaaaaat	2820
gcggaatgga atgcaggatt aggaagaaag cgtacagaaa cagatcgcgt gtatcaagct	2880
gcgaaacaag caattaatca tctatttgta gactatcaag atcaacaatt aaatccggaa	2940
atagggctaa cggagattaa tgaagtctca aatcttggtg agtcaattac ggatgtatat	3000
agtgatacag tattgcaa	3060
atcgaatc cctggaatt aactacgaga ttacaaaaga gttatccgat	3120
cgattacaac aagcatcgaa tctgtatacg tctcgaaatg ctgtgcaaaa cggcgacttt	3180
gacagtgggt tagatggttg gaacgcaaca actgatacat cggttcagca ggatggcaat	3240
atgcatttct tggttctttc ccattgggat gcacaagttt cccaacaatt gagagtacac	3300
ccgaactgta agtatgtctt acgtgtgaca gcaagaaaag tgggaggcgg cgatgggtac	3360
gtcacaatcc gagatggcgc tcatcaccaa aaaaccctta catttaatgc atgtgattac	

gatgtaaatg gtacgtatgt agatgataat acgtatataa caaaagatgt gatattctac 3420
ccagagacaa aacatatgtg ggtagagggtg agtgaatccg aagggttcatt ctatatagat 3480
agtattgagt ttattgaaac acaagagtaa 3510

<210> 3
<211> 1872
<212> ADN
<213> *Bacillus thuringiensis*

<400> 3
atgaataata tattgaataa cgaaagaact actattttgtg atgcgcataa tgtagtggct 60
catgatccat ttagttttga acataaatca gtagatacca tccaaaaaga atggacggaa 120
tggaagagaa ctgatcatag ttatatacgt tctcctattg ttgggaccgt tgctagtttg 180
ctgctaaaaga aagtggggag tcttatttga aaaaggatat tgagtgagtt atgggggttta 240
atattcccta gtggtagtac aaatctaata caagatattt taagagcgac agaacaattc 300
ataaatcaaa gacttaatgc agacactctt gctcgtataa atgtagaatt ggtagggctt 360
caagccaatg taaggaggtt taatcaacaa gtagataatt ttttaaacc taaccgaaac 420
gctgttcctt tatcaataac ttcttcagtt aatacaatgc agcaaatatt tctaaataga 480
ttaccccagt tccagataca aggatacgag ctgttattat tacctttatt cgcacaagca 540
gccaatatgc atctttcttt tattagagat gttattctta atgcagatga atgggacatt 600
tccccagcaa tgttacgtac atatcgagat cacctaaaaa gatatacaaa cgaatattcc 660
aattattgta taaatacgt tcaactgaa tttgaagggt taaatgttcc tttaaaagat 720
atgttagaat ttagaaccta tatgttttta aatgtatttg aatatgtatc catatggctg 780
ctattttaa ataccagagcct tctggtaact tctggcgcta acttgatgtt ttctggtaca 840
actgtaggtg caaatcaatc ttttactgca agaaattggc cgttttttaa tactcttttc 900
caagttaatt caaattatat attgactgga tttagtggag ctagatttca atcaacotta 960
ctcaattcat ttgttatcac attattgggt gcacgtattc attacatggg aggagtctca 1020
tctggtaatt taggtccagc agctaactat aatgttagtt gtaatttatc atccctata 1080
caattgccat atcgaacatc ttatttagac tcatttatat taagagatgg aaatcgaaca 1140
aatgaatatt ttacggggcg ttttcttgca ttgatactac cttgtgggtg tactggacct 1200
actctgggtg cagctaactc tttcccgaa tattccatac ggaatatttc tggggctggt 1260
ggctgggttta gaaatgaaga tttagaaaga cctttacact ttaatgaaat aagaaatata 1320

gaaggtcaac cgggagtagg tggtggaagt agaatctcta gtttagtatac tgtgcataac 1380
 agaaaaaata atatcgctgc tcatgaaaat ggtactatga ttcatttagc tccagatgat 1440
 tatacaggat tcactatatac accgatacat gcaactcaag tgaataatca aacgcgaaca 1500
 attattttctg aaaaatttgg aaatcaaggt gattccttaa ggtttgaaca aagcaacacg 1560
 acagctcggtt atacgcttag agggaatgga aatagttaca atctttatctt aagagtatct 1620
 tcaataggaa attccactat tcgggttact ataaacggta gagtttatac tgctacaaat 1680
 gttaatacta ctacaaataa cgatggagta cttgataatg gagctcgttt ttcagatatt 1740
 aatatcggtta atgtagtggc aagtgcctaact actaatgtac cattagatat aaatgtgaca 1800
 ttttaacgcca atacaccatt tgagcttatg aatattatgt ttgtgccaac taatcttcca 1860
 ccagtatatatt aa 1872

<210> 4
 <211> 1920
 <212> ADN
 <213> *Bacillus thuringiensis*

<400> 4
 atggatagtg tattgaatag cggaagaatt accaatcgtg atgcgtataa tgtagtggct 60
 cttgatccat ttagttttca acacaaatca ttagatacca tacaaaaaga atggacggag 120
 tggaaaaaaaa ataatcatag tttatatgta gatcctattg ttggaactgt ggctagtttt 180
 ctgttaaaga aagtggggag tcttattgga aaaaggatac tgagtggagt acggaattta 240
 atatttccta gtggcagtac aaatctaata caagatattt taagagagac agaaaaattc 300
 ctgaatcaaa gacttaatac agacactctt gcccggtgta atgctgaatt gacagggatg 360
 caagcaaata tagacgagtt taatcgacaa gtagataatt ttttgaaccc taaccgaaac 420
 gctgttcctt tatcaataac ttcttcagtt aatacaatgc agcaattatt tctaaataga 480
 ttaccccagt tccagatgca aggataccaa ctgttattat tacctttatt tgcacaggca 540
 gccaatctac atctttcttt tattagagat gttattctta atgcagatga atggggaatc 600
 tcagcagcaa cattacgtac gtatcgagat tacctgaaaa attatacaag agagtactct 660
 aactattgta taaatacgta tcaaagtgcg tttagagggt taaacaccag tttacacgat 720
 atgttagaat ttagaaccta tatgttttta aatgtatttg aatatgtctc tatctggctg 780
 ttattttaaat atcaaagtct tctagtatct tccggcgcta atttatatgc aagtggtagt 840

tcaattggta gtgtaccaga gcggacccaa acatatactt cacaagattg gccattctta 900
 aattcccttt tccaagttaa ttcaaattat gtattatctg gttatagtgg tgctagtctt 960
 tttattttctc tccctaatat tgggtctacct cctggcagta ctacaactca agcattgctt 1020
 ggtgcaaggg ttaattattg tggaggaata tcctctggct tcataggagt atctaatttc 1080
 aatcaaaaact ttaattgtag cgcgggtccc tctcctttat caacaccatt tattagaagt 1140
 tggctggatg caggttcaga tcgacagggc gctaatacct ctacaaattg gcaaacagaa 1200
 gcatttcaga caactcaggg ccgaagtgtg ggtgcttttg gagctcgtgg taattcgaac 1260
 tttttccctg attattttat ccgtaatatt tctggagttc ctttagttgt tagaaatgaa 1320
 gatctaagaa gaccactaca ttacaatgaa ataaggaata tagaaagtcc tgcaggtaca 1380
 cctgctggat cacgagctta tatggtatct gtgcataaca gaaaaataa tatttatgcc 1440
 gttcatgaaa acggtactat gattcatttg gcgccggaag attatacagg atttactata 1500
 tcaccgatac atgcaactca agtgaatagt caagctaatac aagcgcgaac atttatttct 1560
 gaaaaatttg gaaatcaggg tgattcctta agggttgaac aaagcacctc aacggttcaa 1620
 tacacattta gagggaatgg acatagttac aatctttatt taagagtatc ttcaatagga 1680
 agttccacta ttcgagttgg tctaaacggc acagtttata ctgttcgaga tgttaatact 1740
 actacaaata atgatggagt acttgataat ggagctcgtt tttcagatat ttatatcggt 1800
 aatgtagtgg caagtgctaa tactaatgta ccattaaata taactatttc atttaacggc 1860
 aatccacaat ttgagcttat gaatattatg tttgtgccaa ctaatcttcc accagtatat 1920

<210> 5
 <211> 3501
 <212> ADN
 <213> *Bacillus thuringiensis*

<400> 5
 atggaggtaa ataaccagaa tcaatgtatc ccoctataact gtttgaataa tcctgaaagt 60
 gagatattta atgctaggaa ttccaatttc caacttgttt ctcaagtcag cgtgggaactt 120
 acacgttttc ttctagaagc cgctattcca ggggctgggt ttgcactggg cctatttgat 180
 atcatttggg gcgctctagg tgttgatcaa tggagtttat ttcttgagca gattgagcaa 240
 ttaatacgac aagaaataac aactgtggaa agaaatagag caactcaagc attagtggga 300
 ctttcggcaa gttatgatat atatattgag gcattaagag aatgggaaaa tgatcctgat 360
 aatccagctt cacaagatag agtggttgca cgttttcgtg caacagataa ttctctaata 420

acagatatatac ctctatttagc aatcccaaac tatgagatag cgactttatc agtctatgtt	480
caagcagcga atctccattt agctttgtta agagatgctg tttactttgg agaaagatgg	540
ggattgacac cagtaaatat tgaagattta tacacaagac tcacgaggaa tatccaacta	600
tattcagacc attgtgtaag atggtataat caaggcttga atagtattgg agcgacaaat	660
acgagatatt tggacttcca aagagattta acaatttctg tcttagatat tgtcgccctt	720
ttcccaaact acgatatccg aacatatcca attccgacac aaagccaatt aacaagggag	780
atttatacat ctcccgctgt tgcagatggg gttaaattggc ttttaagtat tgcgaatgta	840
ttgagagccc ctcatctggg ggactttttt gatcgaataa tcatttatac taatacagtt	900
agaagtactc cacattgggc agggcatgaa gtcatatcga gaagaacagg gcaaggacaa	960
ggaaatgaga tacgatttcc tttatatgga gtggctgcaa atacagaacc accagttact	1020
ataagaccta caggatttac tgatgagcaa cgacaatggg atagagtact atcacgtatt	1080
gtctcgttta gaagttcagg gcaagatttt agtttggttag atgcagtagg attccttact	1140
atatttagtg ccctttcatt ctatagaaat ggctttggat ttaacacgga tactattaat	1200
gaaattccaa ttgagggaa ccatccacta actggatata gccaccgatt atgccatgtt	1260
ggttttgttg cgtcatctcc attcatcagt cagtatgcaa gggttcctgt attttcttgg	1320
acgcaccgta gtgcaaccct taaaaatata attgctccag atgtcattac ccaaataccg	1380
ttagtaaagg ctttcaatct tcattcaggt gccacggttg ttagaggacc aggttttaca	1440
ggtggggata tccttcgaag aacgaatgtt ggtagctttg gagatatgcg tgtaaacatt	1500
actgcaccac tatcaciaag atatcgcgta aggattcgtt atgcttctac gacagattta	1560
caattccata catcaattaa cggaagagcg attaatcagg caaattttcc agcaactatt	1620
gagaataggg ggaatttaca attcggagac ttcagaactg taggttttac tactccattt	1680
actttttcag atgcaaacag cacattcaca ataggtgctt ggagcttctc ttcaggtaac	1740
gaagtttata tagatcgaat tgaatttggt ccggcagaag taacctttga ggcagaatat	1800
aattttaaac aagcacaaca ggcggtgaat gcgctgttta cttcttccaa tcaaattggg	1860
ttaaaaacag atgtgacgga ttatcaaatt gatcaagtgt ccaatttagt tgagtgttta	1920
tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag	1980
cgactcagtg atgagcgcaa tttacttcaa gattcaaact ttagaggcat caatagacaa	2040
ccagaccgtg gttggagagg aagtacagat attatcatcc aaggaggaga tggcgtattc	2100

```

aaagagaatt acgttacact accgggtacc tttgatgagt gctatccaac gtattttatat 2160
caaaaaatag atgagtcgaa attaaaaagcc tatacccggt atcaattaag aggggtatatac 2220
gaggatagtc aagatctaga aatctatttg attcgttata atgcaaagca tgaaacagta 2280
aatgtgccag gatctggctc cttacggccg ctttcagccc aaaatccaat tggaaagtgc 2340
ggagagccga atcgctgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2400
agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgttggg 2460
tgtacagatc taaatgagga cctaggtgta tgggtgatat tcaagattaa gacgcaagat 2520
ggtcattgca gactaggaaa tctagagttt ctcgaagaga aaccactatt aggggaagca 2580
ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgagaaaa attggaattg 2640
gaaacaaata ttgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaaactct 2700
caatatgata gattacaagc ggatacgaat atcgcgatga ttcattgcggc agataaacgc 2760
gttcatagaa ttcgagaagc gtatctgcca gagttatctg taattccggg tgtaaatgcg 2820
ggcatttttg aagaattaga gggacgtatt ttcacagcct actctcaata tgatgcgaga 2880
aatgtcatta aaaacggcga ttttaataat ggtttatcgt gctggaacgt gaaagggcat 2940
gtagatgtag aagaacaaca caaccacgt tccgtccttg ttgtcccga atgggaagca 3000
gaagtgtccc aagaagtctg tgtctgtcca ggtcgtggct atatcctgcg tgttacagcg 3060
tacaaagagg gatattggaga gggcacagta acaatccatg agatcgaaga tcatacagat 3120
gaactgaaat tcagcaactg tgtagaagag gaagtctatc caaacaacac ggtaacgtgt 3180
aatgattata ctgcgactca agaagaatac gggggtgcat acacttcccg taatcgtgga 3240
tatgacgaaa cttatggaag caattcttcc gtatcagcag attatgcgtc agtctatgaa 3300
gaaaaagcgt atacagatgg acgaagagag aatccttgtg aatttaacag aggatatggg 3360
gattacacac cactaccagc tggttatgtg acaaaagaat tagagtactt cccagaaacc 3420
gataagatat ggattgagat cggagaaaac gaaggaacat tcatcgtgga cagcgtggaa 3480
ttactcctta tggaggaata g 3501

```

<210> 6
 <211> 2346
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Bộ ba mã hóa BT-0049 được tối ưu hóa.

<400> 6

atgaacccat accagaacaa gaacgaccac gagatcctca acgccaccca gaacaactgc	60
aggatgccga actgctactc ccgctaccca ctgagcaacg atccgcagat ctctgccag	120
aacaaccact acaaggactg gctgaacaag tacgagaaca tcaccacgac ctccagcctc	180
tccggcatcc tgaccagcat cggcatgttc aacaagggtgc tgatgctcac cggcgtgctc	240
ggcgtcacgc cgatcgtgat caacatcatc ctcgatatgg tcaggctgat cgacatgggc	300
accggcgacg atctcctggc acacaccgag cagctcatcc agcagacgct ggcagcacag	360
tacaggtccg cagcaacggg cgcaatctac ggcatcagcc gcgcgtacga gaactacctg	420
atgttcttca agcagtggga gaggaaccgc accccacaga acggccagca ggtggagtcc	480
gccttcacga ccgtcaacac gctctgcatc accgcactgg caccacaggc actcctgagc	540
aggagggggt tcgagaccct cctgctccca aactacgcac tcgcagcgaa cttccacctg	600
ctcctgctca gggatgccgt gatctacagg aactcctggc tgccgagcaa cgtgtccatc	660
aagtccgtca acctcacgat cctgaggaac gccatcagcg agtaccgcaa cactgcaac	720
cactggtaca acgatggcct caacagggtc gcccgctcca gcttcaacga ctgggtgagg	780
ttcaacgcgt accgcaggga tatgaccctc tccgtgctgg acttcatcac ggtcttcccg	840
acctacgacc caatcatcta cccgaaggcg acgaacatcg agctacccg catcgtctac	900
acgaacccaa tcatcccacc aaggggcttc ccaaggacga acccaccac cttcaaccag	960
atggagaacc tgatcatctc cggctgcccc agcttctctc accagctgag gatctacacg	1020
accttctacc acgatccgca caacatctcc cgcgacttct gggccggcaa ccagaacttc	1080
ctctccaacg gcacgatcag gaccagcggc gcactggcag gcaggaggat cgatttcagc	1140
atgcagaaca tcgacgtggt ccgcgtcaac atgacgaccc acgacatcga ttccctcagc	1200
aggtcctacg gcggcgtgca cagggtccgac ttcatcggcg tcaacatcca gaacaaccag	1260
ggcaccagca tcctctacaa cgtgccgctg gatacgacca ggttcatgac gaggaacgag	1320
accgcgttcc tcccaggcga cagcggcctg gagccaaacg agaggaacta cacgcaccgc	1380
ctgttccagg tcatgacgac ctacagggtc aacccaaacg caaggagggc agcgttctct	1440
cacgcatgga cccacaggag cctgaggagg aggaacggct tccgcgcaga ccagatcatc	1500
cagatcccgg ccgtgaaggc gatcaacacg tccgacgata gcgcgggtgg ctccaacgtc	1560
ggcgagaaca tgatgaagct cgataacctg acgaccagcc tctcctacaa gctgaccgca	1620

gaggactccg aggcaagcaa cacgcgcttc atcgtgagga tccgctacgc ctccatgaac 1680
 aacaacaagc tgaacctcat cctgaacggc acccagatcg cgtccctcaa cgtcgagggc 1740
 accatgcaga acggcggcag cctcacggac ctgcagtacg agaacttcaa gtacgccgcg 1800
 ttcgcctcca acttcaagat cggcagccag tccatcatcg gcatcttcaa ggagaccage 1860
 aacgcggatt tcatcctcga caagatcgag ctgatcccag tggatttcat cccgctgctc 1920
 aagcagacgc aggattgcaa caactacaac caggacacga tctacacctg caaccagggc 1980
 tacgacgcct acgagccaaa cgcgagcgac tcctacgatt cgtctacca cccggacgcc 2040
 aacgatgcgt acgacgaggg cgacgataac acctacaacc agaactacaa ctccatgtac 2100
 aaccagaacg ccgataacac ctacgaccag agctacaaca actacgatcc aaacatggag 2160
 aacacgtacc agccgagcta cgacaacggc tacgagcaga acaactacga ttccgacgat 2220
 agcgactaca acaacactta taaccagaac acgaacaaca tgtacaacca ggagtacaac 2280
 aacaacgaca acccaaaciaa cggctgcacc tgcaagcagt gctacaacca gaactaccgc 2340
 aagtga 2346

<210> 7
 <211> 3510
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Bộ ba mã hóa BT-0063 được tối ưu hóa.

<400> 7
 atgaaccgca acaaccagaa cgagtacgag gtcatcgacg caccacactg cggctgccc 60
 agcgaggata tcgtgaagta cccgctgacc gacgatccaa acgcgggcct ccagaacatg 120
 aactacaagg agtacctgca gatgcacggc ggcgactaca ccgatccgct gatcaacca 180
 aacctctccg tcagcggcaa ggacgtgatc caggtcggca tcaacatcgt gggccgcctc 240
 ctgtccttct tcggcttccc gttctccagc cagtgggtga ccgtctactc ctacctctg 300
 aacagcctct gccagacga tgagaacagc gtctgggacg ctttcatgga gaggggtggag 360
 gagctgatcg atcagaagat ctccaccctc gtgaagggca acgcgctcga caacctgacg 420
 ggcctccaga acaactacaa cctctacgtg gaggcctgg atgagtggct caacaggcca 480
 aacggcgcaa gggcagcact cgtgtcccag cgcttctaca tcctggacag cctcttcacc 540
 cagttcatgc catccttcgg ctccggcagc ggctcccaga actacgccac gatcctcctg 600

gcagtctacg cacaggcagc aaacctccac ctcttgctcc tgaaggacgt ggagatctac	660
ggcgcacgct ggggcctgaa ccagaccag atcgatcagt tccacagcag gcagcagtc	720
ctcaccaga cgtacaccaa cactgctc accgcataca acgacggcct ggcagagctc	780
aggggcacga gcggtggagtc ctggctcaag taccaccagt accgcaggga gatgacggtc	840
accgccatgg atctggtggc gctcttccc tactacaacg tgcgccagta ccaaacggc	900
gcaaaccac agctgacgag ggaggtctac accgaccaa tcgtgttcaa cctcccagag	960
aggccatccg gcgcattctg cgagagcttc tacaccatca ggcagcaag ggagaggctg	1020
acgttcagcc agctcgagaa cgcgatcatc cgcgcgcaa ggctgttcga gcgcttcag	1080
gcactcggca tctacaccgg cgaggcaagg ctgaacgcaa acagcgcgcc gatgaactac	1140
tggatcggcc acttcatccg caacacgagg ctgggcgact ccaccacgat caccacgaac	1200
tacggcacgg tcaacaaccg cctcaccaac ttcatccgc caaccacgtc cgatgtgtac	1260
cagatcaaca gcatctccag caacctcgca gcaatcctgg gcacgctcta cggcgtcacc	1320
agggcccagt tctactacgg ctccggcatc atctggagct acgtcggcca gaactccgtg	1380
ctggcgagtc gccaccagaa ctacaacagc atcgaggagc tccgaacca gtccgacgag	1440
ccaaccgtgc gcagctactc ccacaggctg agccacatca cgtccttcaa cttcaacgtc	1500
cagctcaaca acccggtcac cagcgccggc aacatgccag tctacgtgtg gaccaccgc	1560
agcgtcgacc tgaacaacac gatcacctcc gataggatca cgcagctgcc agcagtgaag	1620
gcattcgagc tctccgacgg cgcaacgata gtcaagggcc caggcttcac cggcggggat	1680
gtgatcagga ggacgaacca gggcggttc ggcagcatcc gcggtgccgt cacgggcca	1740
ctcaccaga ggtacaggat ccgcttcagg tacgccagca ccatcgactt cgatttcttc	1800
gtcacgaggg gcggcaccac gatcaacaac ttccgcttca cgaggaccat gaacagcggc	1860
caggagtccc gctacgagag ctacaggacc gtggagttca ccacgccatt caacttcacc	1920
cagagccagg acatcatcag gacgtccatc cagggcctga gcggcaacgg cgaggtctac	1980
ctcgaccgca tcgagatcat ccggtgaac ccaaccaggg aggcagagga ggatctggag	2040
gcagcgaaga aggcagtggc atccctgttc acgaggacca gggacggcct ccaggtgaac	2100
gtcaccgact accaggtcga tcaggccgcg aacctgggtga gctgcctctc cgacgagcag	2160
tacggccacg ataagaagat gtcctggag gccgtccgcg cagcaaagag gctcagcagg	2220
gagaggaacc tcctgcagga ccagatcttc aacgccatca actccaccga ggagaacggc	2280

tggaaggcgt ccaacggcgt gacgatcagc gagggcggcc cattcttcaa gggcagggca	2340
ctccagctgg caagcgcaag ggagaactac ccgacctaca tctaccagaa ggtcgatgcg	2400
tccgtgctga agccatacac gcgctacagg ctcgacggct tcgtgaagtc cagccagaac	2460
ctggagatcg atctcatcca ccaccacaag gtgcacctgg tcaagaacgt gccggacaac	2520
ctcgtctccg acacctacag cgatggcagc tgctccggca tcaaccgctg cgacgagcag	2580
caccaggtgg acatgcagct ggatgccgag caccacccga tggactgctg cgaggcagca	2640
cagaccacg agttctccag ctacatcaac acgggcgacc tcaactccag cgtcgatcag	2700
ggcatctggg tggctctgaa ggtgaggacc acggatggct acgcgaccct cggcaacctc	2760
gagctggtgg aggtcggccc actgagcggc gagtccctcg agcacgagaa gcgcaagaac	2820
gcagagtggg acgcaggcct gggccgcaag aggacggaga ccgacagggg ctaccaggcc	2880
gcgaagcagg ccatcaacca cctcttcgtg gactaccagg atcagcagct gaacccggag	2940
atcggcctca ccgagatcaa cgaggtctcc aacctggtgg agagcatcac ggacgtctac	3000
agcgataccg tgctccagat ccaggcatc aactacgaga tctacaagga gctgtccgac	3060
aggctccagc aggcattcaa cctgtacacc agcaggaacg cggatgcagaa cggcgacttc	3120
gatagcggcc tcgatggctg gaacgcaacc acggacacgt ccgtccagca ggatggcaac	3180
atgcacttcc tcgtgctgtc cactgggac gcacaggtca gccagcagct gaggggtgcac	3240
ccgaactgca agtatgtgct cagggtgacc gcaaggaagg tcggcggcgg cgacggctac	3300
gtgacgatca gggatggcgc ccaccaccag aagacgctca ccttcaacgc gtgcgactac	3360
gatgtcaacg gcacctacgt ggacgataac acgtacatca ccaaggacgt gatcttctac	3420
ccggagacga agcacatgtg ggtggaggtc agcgagtccg agggctcctt ctacatcgat	3480
agcatcgagt tcatcgagac ccaggagtga	3510

<210> 8

<211> 1872

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Bộ ba mã hóa BT-0064 được tối ưu hóa.

<400> 8

atgaacaaca ttctcaacaa cgagcgcacc acgatctgcg acgcccacaa cgtggctgcg	60
---	----

cacgatccgt tctccttcga gcacaagagc gtcgacacga tccagaagga gtggacggag	120
---	-----

tggaagagga ccgatcactc cctctacgtc agcccaatcg tgggcaccgt cgcaagcctc	180
ctgctcaaga aggtcggctc cctgatcggc aagcgcatcc tcagcgagct gtggggcctc	240
atcttcccgt ccggcagcac gaacctcatg caggacatcc tgagggcgac cgagcagttc	300
atcaaccagc gcctcaacgc cgataccctg gcgaggatca acgtggagct ggtcggcctc	360
caggcaaacg tccgcgagtt caaccagcag gtggacaact tcctcaaccc gaacaggaac	420
gcgggtccac tgtccatcac gtccagcgtg aacaccatgc agcagatctt cctgaaccgc	480
ctcccgagct tccagatcca gggctacgag ctgctcctgc tcccactctt cgcacaggca	540
gcgaacatgc acctctcctt catcaggagc gtgatcctga acgccgacga gtgggatatc	600
agcccggcga tgctccgcac gtacagggat cacctgaagc gctacaccaa cgagtacagc	660
aactactgca tcaacacgta caggaccgag ttcgagggcc tgaacgtccc actcaaggac	720
atgctggagt tccgcacgta catgttcctc aacgtgttcg aatacgtgtc catctggagc	780
ctgttcaagt accagtcctt gctcgtgacc agcggcgcca acctctacgt gtccggcacc	840
acggtcggcg caaaccagag cttcacggcg aggaactggc cgttcctgaa caccctcttc	900
caggtcaact ccaactacat cctgaccggc ttcagcggcg caaggttcca gtccacgctg	960
ctcaacagct tcgtgatcac cctgctcgcc gcgaggatcc actacatggg cggcgtctcc	1020
agcggcaacc tcggcccagc cgcgaactac aacgtgtcct gcaacctctc cagcccgatc	1080
cagctgccat accgcacctc ctacctcgac agcttcatcc tgcgcgatgg caacaggacg	1140
aacgagtact tcaccggcgc cttcctggca ctcatcctgc catgcggcgt cacgggcca	1200
acctgggcg cagcgaactc cttcccgaac tacagcatcc gcaacatctc cggcgcagtg	1260
ggctggttca ggaacgagga cctcgagagg ccgctgcaact tcaacgagat caggaacatc	1320
gagggccagc caggcgtcgg cggcggcagc cgcattctcca gctcgtgtc cgtccacaac	1380
aggaagaaca acatcgtggc ccacgagaac ggcacgatga tccacctggc gccggacgat	1440
tacacgggct tcaccatctc cccaatccac gccaccaggg tgaacaacca gacgcgcacc	1500
atcatcagcg agaagttcgg caaccagggc gatagcctca ggttogagca gtccaacacc	1560
acggcaaggt acaccctgag gggcaacggc aactcctaca acctctacct gcgcgtctcc	1620
agcatcggca acagcacgat ccgctgacc atcaacggca gggctctacac ggccaccaac	1680
gtgaacacca cgaccaacaa cgacggcgtg ctcgataacg gcgcgaggtt ctccgacatc	1740
aacatcggca acgtggtcgc cagcgcgaac acgaacgtcc cgctggatat caacgtgacg	1800

ttcaacgcc aacacccggtt cgagctcatg aacatcatgt tcgtcccaac caacctgccg 1860
ccagtgtact ga 1872

<210> 9
<211> 1923
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Bộ ba mã hóa BT-0086 được tối ưu hóa.

<400> 9
atggacagcg tgctcaactc cggccgcatc accaacaggg acgcgtacaa cgtgggtcgcc 60
ctcgatccgt tcagcttcca gcacaagtcc ctggacacga tccagaagga gtggaccgag 120
tggaagaaga acaaccactc cctctacgtc gatccaatcg tgggcaccgt cgccagcttc 180
ctcctgaaga aggtcggctc cctgatcggc aagcgcatcc tcagcgagct gaggaacctc 240
atcttcccggt ccggcagcac gaacctcatg caggacatcc tgcgcgagac cgagaagtcc 300
ctgaaccagc gcctcaacac cgatacgtc gcaaggggtca acgcagagct gaccggcatg 360
caggcgaacg tcgacgagtt caaccgccag gtggataact tcctcaaccc gaacaggaac 420
gccgtcccac tgtccatcac gtccagcgtg aacaccatgc agcagctctt cctgaacagg 480
ctcccgagct tccagatgca gggctaccag ctctgtctcc tgccactgtt cgcacaggca 540
gcgaacctcc acctgtcctt catccgcgac gtgatcctca acgcagatga gtggggcatc 600
agcgagcaa cgctccgcac ctacagggac tacctgaaga actacacgag ggagtacagc 660
aactactgca tcaacaccta ccagtccggt ttccgcggtc tgaacacgag cctccacgat 720
atgctggagt tcaggacctc catgttcttc aacgtcttcg aatacgtgtc catctggagc 780
ctgttcaagt accagtcctt cctggtctcc agcggcgcaa acctgtacgc aagcggctcc 840
agcatcgggt ccgtgccgga gcgcacccag acgtacacca gccaggactg gccattcctg 900
aacagcctct tccaggtcaa ctccaactac gtgctctccg gctacagcgg cgcatccctc 960
ttcatctccc tgccaaacat cggcctgcca ccaggcagca ccacgaccca ggcactcctg 1020
ggcgcaaggg tcaactactg cggcggcatc tccagcgggt tcatcggcgt gtccaacttc 1080
aaccagaact tcaactgcag cgcaggccca agcccactct ccacgccgtt catcaggagc 1140
tggttgagc caggctccga taggcagggc gcaaacacga gcaccaactg gcagaccgag 1200
gcattccaga cgaccagggt caggctctgc ggcgcattcg gcgcgagggg caacagcaac 1260

ttcttcccg actacttcat ccgcaacatc tccggcgtgc cactcgtggt caggaacgag 1320
 gatctccgca ggccgctgca ctacaacgag atccgcaaca tcgagagccc agcaggcacg 1380
 ccagcaggca gcagggcata catggtgtcc gtccacaaca ggaagaacaa catctacgag 1440
 gtccacgaga acggcaccat gatccacctc gccccggagg actacacggg cttcaccatc 1500
 tccccaatcc acgcgacgca ggtgaacagc caggccaacc aggcgaggac cttcatctcc 1560
 gagaagttcg gcaaccaggg cgatagcctg agggtcgagc agtccacgag caccgtgcag 1620
 tacacgttca ggggcaacgg cactcctac aacctctacc tgcggtctc cagcatcggc 1680
 tccagcacca tcagggtgga cctcaacggc acggtgtaca ccgtccgca cgtgaacacg 1740
 accacgaaca acgatggcgt cctggataac ggcgccaggc tctccgatat ctacatcggc 1800
 aacgtggtcg ccagcgcgaa cacgaacgtg cactcaaca tcaccatcag cttcaacggc 1860
 aaccgcagt tcgagctcat gaacatcatg ttcgtcccaa ccaacctgcc gccagtgtac 1920
 tga 1923

<210> 10
 <211> 3501
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Bộ ba mã hóa BT-0498 được tối ưu hóa.

<400> 10
 atggagggtga acaaccagaa ccagtgcac cgtacaact gcctgaacaa cccagagagc 60
 gagatcttca acgcccga ctccaacttc cagctggtct cccagggtgag cgtcggcctc 120
 accaggttcc tcctggaggc agcaatccca ggcgcaggct tcgcgctggg cctcttcgac 180
 atcatctggg gcgcactggg cgtggatcag tggagcctgt tcctcgagca gatcgagcag 240
 ctcatccgcc aggagatcac cacggtggag cgcaacaggg caaccaggc actggtcggc 300
 ctctccgcca gctacgatat ctacatcgag gcactcaggg agtgggagaa cgaccgggat 360
 aaccagcaa gccaggacag ggtggtcgca cgcttcaggg caacggacaa ctccctgatc 420
 accgatatcc cgctcctggc catcccaaac tacgagatcg caaccctctc cgtctacgtg 480
 caggcagcaa acctgcacct cgcactcctg agggacgcag tgtacttcgg cgagcgtgg 540
 ggctgacgc cagtcaacat cgaggatctc tacacgcgcc tgaccaggaa catccagctc 600
 tacagcgacc actgcgtgag gtggtacaac cagggcctga actccatcgg cgccaccaac 660

acgcgctacc tcgaattcca gagggatctc accatcagcg tgctggacat cgtcgcgctc	720
ttcccaaaact acgatatccg cacgtaccgg atcccaaccc agagccagct gacgagggag	780
atctacacct cccagtggt cgcagacggc gtcaactggc tcctgtccat cgcaaactg	840
ctgagggcac cacacctggt cgacttcttc gatcgcatca tcattctacac caacacggtg	900
aggagcacc cactgggc aggcacgag gtcattctcca ggaggacggg ccagggccag	960
ggcaacgaga tcaggttccc gctctacggc gtggcagcaa acacggagcc accagtcacc	1020
atcaggccaa cgggcttcac cgacgagcag cgccagtggt acaggggtgt gtcccgcatc	1080
gtcagcttca ggtccagcgg ccaggacttc agcctcgtgg atgccgtcgg cttcctgacc	1140
atcttctccg cgctcagctt ctaccgcaac ggcttcgggt tcaacaccga cacgatcaac	1200
gagatcccga tcgagggcac ggatccactg accggctaca gccacagggt ctgccacgtg	1260
ggcttcgtcg catccagccc gttcatctcc cagtacgcgc gcgtgccagt cttctcctgg	1320
accacagga gcgcaaccct gacgaacacc atcgaccag acgtgatcac gcagatccca	1380
ctggtcaagg cttcaacct cactccggc gcaacgggtg tcagggggcc aggttcacc	1440
ggcggcgaca tcctcaggag gacgaacgtg ggcagcttcg gcgatatgcg cgtcaacatc	1500
accgcaccac tgtcccagag gtacaggggt cgcattccgt acgcgagcac cacggacctc	1560
cagttccaca cctccatcaa cggcagggcg atcaaccagg ccaacttccc ggcgacgatc	1620
gagaaccgcg gcaacctgca gttcggcgat ttcaggaccg tgggcttcac cacgccattc	1680
acgttctccg acgccaacag cacgttcacc atcggcgctg ggagcttctc cagcggcaac	1740
gaggtctaca tcgatcgcat cgagttcgtg ccagcagagg tcaccttcga ggcagagtac	1800
aacctcaagc aggccagca ggcggtgaac gccctgttca cgtccagcaa ccagatcggc	1860
ctcaagacgg acgtcaccca ttaccagatc gaccaggtga gcaacctcgt cgagtgcctg	1920
tccgacgagt tctgctgga tgagaagcgc gagctcagcg agaaggtgaa gcacgccaag	1980
aggctcagcg acgagcgcaa cctcctgcag gattccaact tcaggggcat caacaggcag	2040
ccagacaggg gctggagggg ctccaccgac atcatcatcc agggcggcga tggcgtgttc	2100
aaggagaact acgtcacgct gccgggcacc ttcgacgagt gctacccaac ctacctctac	2160
cagaagatcg atgagagcaa gctgaaggcc tacacgaggt accagctccg cggctacatc	2220
gaggactccc aggatctcga gatctacctg atcaggtaca acgcgaagca cgagaccgtg	2280
aacgtcccag gctccggcag cctgaggcca ctacgcgcac agaacccaat cggcaagtgc	2340

ggcgagccaa acaggtgctc accacacctg gagtggaaacc cagacctcga ttgctcctgc 2400
 agggatggcg agaagtgcgc acaccactcc caccacttca gcctggacat cgatgtgggc 2460
 tgcaccgacc tcaacgagga tctgggctgt tgggtcatct tcaagatcaa gacgcaggac 2520
 ggccacgcaa ggctgggcaa cctcgagttc ctggaggaga agccactcct gggcgaggca 2580
 ctcgcaaggg tgaagagggc ggagaagaag tggcgcgata agaggagaa gctggagctc 2640
 gagaccaaca tcgtctacaa ggaggccaag gagagcgtgg acgcgctgtt cgtcaactcc 2700
 cagtacgaca ggctccaggc cgatacgaac atcgcgatga tccacgccgc ggataagcgc 2760
 gtgcaccgca tcagggaggc ctacctcccg gagctgagcg tgatcccagg cgtcaacgca 2820
 ggcatcttcg aggagctcga gggccgcctc ttcaccgcct acagccagta cgacgcgagg 2880
 aacgtgatca agaacggcga tttcaacaac ggctgtcct gctggaacgt caagggccac 2940
 gtggacgtcg aggagcagca caaccacagg agcgtgctcg tgggtcccaga gtgggaggca 3000
 gaggtctccc aggaggtgag ggtctgcccc ggagggggct acatcctgag ggtgaccgca 3060
 tacaaggagg gctacggcga gggcacggtc accatccacg agatcgagga ccacacggat 3120
 gagctcaagt tcagcaactg cgtggaggag gaggtctacc cgaacaacac ggtgacctgc 3180
 aacgattaca cggcaacca ggaggagtac ggcggcgcat acacctccag gaacaggggc 3240
 tacgacgaga cgtacggctc caactccagc gtgagcgcg attacgcctc cgtctacgag 3300
 gagaaggcat acaccgacgg caggagggag aacccatgcg agttcaacag gggctacggc 3360
 gactacaccc cactgccagc aggctacgtg acgaaggagc tcgagtactt cccagagacc 3420
 gataagatct ggatcgagat cggcgagacg gagggcacct tcacgtgga ctccgtcgag 3480
 ctctgctca tggaggagtg a 3501

<210> 11

<211> 2346

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0049 đột biến

<400> 11

atgaacccat accagaacaa gaacgaccac gagatcctca acgccacca gaacaactgc 60

aggatgccga actgctactc ccgctaccca ctacgaacg atccgcagat ctctgccag 120

aacaaccact acaaggactg gctgaacaag tacgagaaca tcaccacgac ctccagcctc	180
tccggcatcc tgaccagcat cggcatgttc aacaagggtgc tgatgctcac cggcgtgctc	240
ggcgtcacgc cgatcgtcat aaacatcatc ctcgatatgg tcaggctgat cgacatgggc	300
accggcgacg atctcctggc acacaccgag cagctcatcc agcagacgct ggacgacag	360
tacagggtccg cagcaacggg cgcaatctac ggcatcagcc gcgcgtacga gaactacctg	420
atgttcttca agcagtggga gaggaaccgc accccacaga acggccagca ggtggagtcc	480
gccttcacga ccgtcaacac gctctgcatc accgcactgg caccacaggc actcctgagc	540
aggaggggct tcgagaccct cctgctccca aactacgcac tcgcagcgaa cttccacctg	600
ctcctgctca gggatgccgt gatctacagg aactcctggc tgccgagcaa cgtgtcaatc	660
aagtccgtca acctcacgat cctgaggaac gccatcagcg agtaccgcaa ccactgcaac	720
cactggtaca acgatggcct caacagggttc gcccgtcca gcttcaacga ctgggtgagg	780
ttcaacgcgt accgcaggga tatgaccctc tccgtgctgg acttcatcac ggtcttcccg	840
acctacgacc caatcatcta cccgaaggca accaacatcg agctcaccg catcgtctac	900
acgaacccaa tcatcccacc aaggggcttc ccaaggacga acccaccaac cttcaaccag	960
atggagaacc tgatcatctc cggctgcccc agcttcctca accagctgag gatctacacg	1020
accttctacc acgatccgca caacatctcc cgcgacttct gggccggcaa ccagaacttc	1080
ctctccaacg gcacgatcag gaccagcggc gcactggcag gcaggaggat cgatttcagc	1140
atgcagaaca tcgacgtgtt ccgcgtcaac atgacgacc acgacatcga ttccctcagc	1200
aggctctacg gcggcgtgca caggctccgac ttcatcggcg tcaacatcca gaacaaccag	1260
ggcaccagca tcctctacaa cgtgccgctg gatacgacca ggttcatgac gaggaacgag	1320
accgcgttcc tcccaggcga cagcggcctg gagccaaacg agaggaaacta cacgcaccgc	1380
ctgttccagg tcatgacgac ctacagggtc aacccaaacg caaggagggc agcgttcttc	1440
cacgcatgga cccacaggag cctgaggagg aggaacggct tccgcgcaga ccagatcatc	1500
cagatcccg cgtgaaggc aatcaacacg tccgacgata gcgcggtggc ctccaacgtc	1560
ggcgagaaca tgatgaagct cgataacctg acgaccagcc tctcctacaa gctgaccgca	1620
gaggactccg aggcaagcaa cacgcgcttc atcgtgagga tccgctacgc ctccatgaac	1680
aacaacaagc tgaacctcat cctgaacggc acccagatcg cgtccctcaa cgtcgagggc	1740
accatgcaga acggcggcag cctcacggac ctgcagtacg agaacttcaa gtacgccgcg	1800

```

ttcgccctcca acttcaagat cggcagccag tccatcatcg gcatcttcaa ggagaccagc 1860
aacgcgggatt tcatcctcga caagatcgag ctgatcccag tggatttcct cccgctgctc 1920
aagcagacgc aggattgcaa caactacaac caggacacgc tctacacctg caaccagggc 1980
tacgacgcct acgagccaaa cgcgagcgac tcctacgatt ccgtctacca cccggacgcc 2040
aacgatgcgt acgacgaggg cgacgataac acctacaacc agaactacaa ctccatgtac 2100
aaccagaacg ccgataacac ctacgaccag agctacaaca actacgatcc aaacatggag 2160
aacacgtacc agccgagcta cgacaacggc tacgagcaga acaactacga ttccgacgat 2220
agcgactaca acaacactta taaccagaac acgaacaaca tgtacaacca ggaatacaac 2280
aacaacgaca acccaaacaa cggctgcacc tgcaagcagt gctacaacca gaactaccgc 2340
aagtga 2346

```

<210> 12

<211> 3510

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0063 đột biến

<400> 12

```

atgaaccgca acaaccagaa cgagtacgag gtcacgacg caccacactg cggctgcccc 60
agcgaggata tcgtgaagta cccgctgacc gacgatccaa acgcgggcct ccagaacatg 120
aactacaagg agtacctgca gatgcacggc ggcgactaca ccgatccgct gatcaaccca 180
aacctctccg tcagcggcaa ggacgtgatc caggtcggaa tcaacatcgt gggccgcctc 240
ctgtccttct tcggcttccc gttctccagc cagtgggtga ccgtctactc ctacctctg 300
aacagcctct ggccagacga tgagaacagc gtctgggacg ccttcatgga gaggggtggag 360
gagctgatcg atcagaagat ctccaccctc gtgaagggca acgcgctcga caacctgacg 420
ggcctccaga acaactacaa cctctacgtg gaggccttgg atgagtggct caacaggcca 480
aacggcgcaa gggcagcact cgtgtcccag cgcttctaca tcctggacag cctcttcacc 540
cagttcatgc catccttcgg ctccggcagc ggctcccaga actacgccac gatcctctg 600
gcagtctacg cacaggcagc aaacctccac ctctgctcc tgaaggacgt ggagatctac 660
ggcgcacgct ggggcctgaa ccagaccagc atcgatcagt tccacagcag gcagcagtcc 720
ctcaccacga cgtacaccaa ccactgcgtc accgcataca acgacggcct ggcagagctc 780

```

aggggacga gcgaggagtc ctgggtcaag taccaccagt accgcaggga gatgacggtc	840
accgccatgg atctggtggc gctcttccca tactacaacg tgcgccagta cccaaacggc	900
gcaaaccac agctgacgag ggaggtctac accgacccaa tcgtgttcaa cctcccagag	960
aggccatccg gcgcattctg cgagagcttc tacaccatca gggcagcaag ggagaggctg	1020
acgttcagcc agctcgagaa cgcgatcatc cgcccgccaa ggctgttcga gcgcttcag	1080
gcactcggca tctacaccgg cgaggcaagg ctgaacgcaa acagcgcgcc gatgaactac	1140
tggatcggcc acttcatccg caacacgagg ctgggagact ccaccacgat caccacgaac	1200
tacggcacgg tcaacaaccg cctcaccaac ttcatccgc caaccacgtc cgatgtgtac	1260
cagatcaaca gcatctccag caacctcgca gcaatcctgg gcacgctcta cggcgtcacc	1320
agggccagtc tctactacgg ctccggcatc atctggagct acgtcggcca gaactccgtg	1380
ctggcgagtc gccaccagaa ctacaacagc atcgaggagc tcccgaacca gtccgacgag	1440
ccaaccgtgc gcagctactc ccacaggctg agccacatca cgtccttcaa cttcaacgtc	1500
cagctcaaca acccggatcat cagcgccggc aacatgccag tctacgtgtg gacccaccgc	1560
agcgtcgacc tgaacaacac gatcacctcc gataggatca cgcagctgcc agcagtgaag	1620
gcattcgagc tctccgacgg cgcaacgata gtcaagggcc caggcttcac cggcgggcat	1680
gtgatcagga ggacgaacca gggcggttc ggcagcatcc gcgtgtccgt cacgggccc	1740
ctcaccacga ggtacaggat ccgcttcagg tacgccagca ccatcgactt cgatttcttc	1800
gtcacgaggg gcggcaccac gatcaacaac ttccgcttca cgaggacat gaacagcggc	1860
caggagtccc gctacgagag ctacaggacc gtggagttca ccacgccatt caacttcacc	1920
cagagccagg acatcatcag gacgtccatc cagggcctga gcggcaacgg cgaggtctac	1980
ctcgaccgca tcgagatcat cccggtgaac ccaaccaggg aggagagga ggatctggag	2040
gcagcgaaga aggagtggtc atccctgttc acgaggacca gggacggcct ccaggtgaac	2100
gtcaccgact accaggtcga tcaggccgag aacctggtga gctgcctctc cgacgagcag	2160
tacggccacg ataagaagat gctcctggag gccgtccgag cagcaaagag gctcagcagg	2220
gagaggaacc tcttcgagga ccagatttc aacgcaatca actccaccga ggagaacggc	2280
tgggaaggcgt ccaacggcgt gacgatcagc gagggcgggc cattcttcaa gggcagggca	2340
ctccagctgg caagcgcaag ggagaactac ccgacctaca tctaccagaa ggtcgatgag	2400
tccgtgctga agccatacac gcgctacagg ctcgacggct tcgtgaagtc cagccagaac	2460

ctggagatcg atctcatcca ccaccacaag gtgcacctgg tcaagaacgt gccggacaac 2520
 ctcgctctccg acacctacag cgatggcagc tgctccggaa tcaaccgctg cgacgagcag 2580
 caccaggtgg acatgcagct ggatgccgag caccacccga tggactgctg cgaggcagca 2640
 cagaccacag agttctccag ctacatcaac acggggcgacc tcaactccag cgtcgatcag 2700
 ggcatctggg tggtcctgaa ggtgaggacc acggatggct acgcgaccct cggcaacctc 2760
 gagctggtgg aggtcggccc actgagcggc gagtccctcg agcacgagaa gcgcaagaac 2820
 gcagagtggg acgcaggcct gggccgcaag aggacggaga ccgacagggg ctaccaggcc 2880
 gcgaagcagg ccatcaacca cctcttcgtg gactaccagg atcagcagct gaaccggag 2940
 atcggcctca ccgagatcaa cgaggtctcc aacctgggtg agagcatcac ggacgtctac 3000
 agcgataccg tgctccagat cccaggaatc aactacgaga tctacaagga gctgtccgac 3060
 aggctccagc aggcattcaa cctgtacacc agcaggaacg cgggtgcagaa cggcgacttc 3120
 gatagcggcc tcgatggctg gaacgcaacc acggacacgt ccgtccagca ggatggcaac 3180
 atgcacttcc tcgtgctgtc ccaactggac gcacaggtca gccagcagct gaggggtgcac 3240
 ccgaactgca agtatgtgct cagggtgacc gcaaggaagg tcggcggcgg cgacggctac 3300
 gtgacgatca gggatggcgc ccaccaccag aagacgctca ccttcaacgc gtgcgactac 3360
 gatgtcaacg gcacctacgt ggacgataac acgtacatca ccaaggacgt gctctttctac 3420
 ccggagacga agcacatgtg ggtggaggtc agcgagtcgg agggctcctt ctacctgat 3480
 agcatcgagt tcacgagac ccaggagtga 3510

<210> 13

<211> 1872

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0064 đột biến

<400> 13

atgaacaaca ttctcaacaa cgagcgcacc acgatctgcg acgcccacaa cgtggctcgcg 60
 cacgatccgt tctccttcga gcacaagagc gtcgacacga tccagaagga gtggacggag 120
 tggaagagga ccgatcactc cctctacgtc agcccaatcg tgggcaccgt cgcaagcctc 180
 ctgctcaaga aggtcggctc cctgatcggc aagcgcattc tcagcgagct gtggggcctc 240

atcttccccgt ccggcagcac gaacctcatg caggacatcc tgagggcgac cgagcagttc	300
ataaaccagc gcctcaacgc cgataccctg gcgaggatca acgtggagct ggtcggcctc	360
caggcaaacg tccgcgagtt caaccagcag gtggacaact tcctcaaccc gaacaggaac	420
gcggtcccccac tgtccatcac gtccagcgtg aacaccatgc agcagatctt cctgaaccgc	480
ctccccgagt tccagatcca gggctacgag ctgctcctgc tcccactctt cgcacaggca	540
gcgaacatgc acctctcctt catcagggac gtgatcctga acgccgacga gtgggatatc	600
agccccggcga tgctccgcac gtacagggat cacctgaagc gctacaccaa cgagtacagc	660
aactactgca taaacacgta caggaccgag ttcgagggcc tgaacgtccc actcaaggac	720
atgctggagt tccgcacata catgttcctc aacgtgttcg aatacgtgtc catctggagc	780
ctgttcaagt accagtcctt gctcgtgacc agcggcgcca acctctacgt gtccggcacc	840
acggtcggcg caaaccagag cttcacggcg aggaactggc cgttcctgaa caccctcttc	900
caggtcaact ccaactacat cctgaccggc ttcagcggcg caaggttcca gtccacgctg	960
ctcaacagct tcgtgatcac cctgctcgcc gcgaggatcc actacatggg cggcgtctcc	1020
agcggcaacc tcggcccagc cgcgaactac aacgtgtcct gcaacctctc cagcccgatc	1080
cagctgccat accgcacctc ctacctcgac agcttcatcc tgcgcgatgg caacaggacg	1140
aacgagtact tcaccggcgc cttcctggca ctcatcctgc catgcggcgt cacgggcca	1200
accctggggc cagcgaactc cttcccgaac tacagcatcc gcaacatctc cggcgcagtg	1260
ggctggttca ggaacgagga cctcgagagg ccgctgcact tcaacgagat caggaacatc	1320
gagggccagc caggcgtcgg cggcggcagc cgcattctcca gcctcgtgtc cgtccacaac	1380
aggaagaaca acatcgtggc ccacgagaac ggcacgatga tccacctggc gccggacgat	1440
tacacgggct tcaccatctc cccaatccac gccaccagg tgaacaacca gacgcgcacc	1500
atcatcagcg agaagttcgg caaccagggc gatagcctca ggttcgagca gtccaacacc	1560
acggcaaggc acaccctgag gggcaacggc aactcctaca acctctacct gcgcgtctcc	1620
agcatcggca acagcacgat ccgcgtgaca atcaacggca gggctctacac ggccaccaac	1680
gtgaacacca cgaccaacaa cgacggcgtg ctcgataacg gcgcgaggtt ctccgacctc	1740
aacatcggca acgtggtcgc cagcgcgaac acgaacgtcc cgtgggatct caacgtgacg	1800
ttcaacgcc aacccccgtt cgagctcatg aacatcatgt tcgtcccaac caacctgccg	1860
ccagtgtact ga	1872

<210> 14
 <211> 1923
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0086 đột biến. Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0086 đột biến

<400> 14

atggacagcg tgctcaactc cggccgcatc accaacaggg acgcgtacaa cgtggtcgcc	60
ctcgatccgt tcagcttcca gcacaagtcc ctggacacga tccagaagga gtggaccgag	120
tggaagaaga acaaccactc cctctacgtc gatccaatcg tgggcaccgt cgccagcttc	180
ctcctgaaga aggtcggctc cctgatcggc aagcgcattc tcagcgagct gaggaacctc	240
atcttcccgt ccggcagcac gaacctcatg caggacatcc tgcgcgagac cgagaagttc	300
ctgaaccagc gcctcaacac cgatacgtc gcaaggggtc acgcagagct gaccggcatg	360
caggcgaacg tcgacgagtt caaccgccag gtggataact tcctcaaccg gaacaggaac	420
gccgtcccac tgtccatcac gtccagcgtg aacaccatgc agcagctctt cctgaacagg	480
ctcccgagcgt tccagatgca gggctaccag ctctgtctcc tgccactggt cgcacaggca	540
gcgaacctcc acctgtcctt catccgcgac gtgatcctca acgcagatga gtggggcatc	600
agcgcagcaa cgctccgcac ctacagggac tacctgaaga actacacgag ggagtacagc	660
aactactgca taaacaccta ccagtccgcg ttccgcggcc tgaacacgag cctccacgat	720
atgctggagt tcaggacata catgttcctc aacgtcttcg aatacgtgtc catctggagc	780
ctgttcaagt accagtcctt cctggtctcc agcggcgcaa acctgtacgc aagcggctcc	840
agcatcggct ccgtgccgga gcgcacccag acgtacacca gccaggactg gccattcctg	900
aacagcctct tccaggtcaa ctccaactac gtgtctctccg gctacagcgg cgcatccctc	960
ttcatctccc tgccaaacat cggcctgcc aacagcagca ccacgaccca ggcactcctg	1020
ggcgcaaggg tcaactactg cggcggcatc tccagcggct tcacggcgt gtccaacttc	1080
aaccagaact tcaactgcag cgcaggccca agcccactct ccacgccgtt catcaggagc	1140
tggtctggacg caggctccga taggcagggc gcaaacacga gcaccaactg gcagaccgag	1200
gcattccaga cgaccaggg caggctcctgc ggcgcattcg gcgcgagggg caacagcaac	1260
ttcttcccg actacttcat ccgcaacatc tccggcgtgc cactcgtggt caggaacgag	1320

gatctccgca ggccgctgca ctacaacgag atccgcaaca tcgagagccc agcaggcacg 1380
ccagcaggca gcagggcata catggtgtcc gtccacaaca ggaagaacaa catctacgcg 1440
gtccacgaga acggcaccat gatccacctc gccccggagg actacacggg cttcaccatc 1500
tccccaatcc acgcgacgca ggtgaacagc caggccaacc aggcgaggac cttcatctcc 1560
gagaagttcg gcaaccaggg cgatagcctg agggtcgagc agtccacgag caccgtgcag 1620
tacacgttca ggggcaacgg cactcctac aacctctacc tgcgcgtctc cagcatcggc 1680
tccagcacca tcaggggtggg cctcaacggc acggtgtaca ccgtccgca cgtgaacacg 1740
accacgaaca acgatggcgt cctggataac ggcgccaggc tctccgatat atacatcggc 1800
aacgtggtcg ccagcgcgaa cacgaacgtg cactcaaca tcaccctcag cttcaacggc 1860
aaccgcagt tcgagctcat gaacctcatg ttcgtcccaa ccaacctgcc gccagtgtac 1920
tga 1923

<210> 15

<211> 3501

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0498 đột biến. Trình tự được tối ưu hóa bộ ba mã hóa mã hóa cho protein BT-0086 đột biến

<400> 15

atggagggtga acaaccagaa ccagtgcac cgtacaact gcctgaacaa cccagagagc 60
gagatcttca acgcccga ctccaacttc cagctggtct cccagggtgag cgtcggcctc 120
accaggttcc tcctggaggc agcaatccca ggcgaggct tcgcgctggg cctcttcgac 180
atcatctggg gcgcactggg cgtggatcag tggagcctgt tcctcgagca gatcgagcag 240
ctcatccgcc aggagatcac cacggtggag cgcaacaggg caaccaggc actggtcggc 300
ctctccgcca gctacgatat ctacatcgag gcactcaggg agtgggagaa cgaccggat 360
aaccagcaa gccaggacag ggtggtcgca cgcttcaggg caacggacaa ctccctgac 420
accgatatcc cgctcctggc catcccaaac tacgagatcg caaccctctc cgtctacgtg 480
caggcagcaa acctgcacct cgcactcctg agggacgcag tgtacttcgg cgagcgctgg 540
ggcctgacgc cagtcaacat cgaggatctc tacacgcgcc tgaccaggaa catccagctc 600
tacagcgacc actgcgtgag gtggtacaac cagggcctga actccatcgg cgccaccaac 660

acgcgctacc	tcgacttcca	gagggatctc	accatcagcg	tgctggacat	cgtcgcgctc	720
ttcccaaact	acgatatccg	cacgtacccg	atcccaaccc	agagccagct	gacgagggag	780
atctacacct	ccccagtggg	cgcagacggc	gtcaactggc	tcctgtccat	cgcaaactg	840
ctgagggcac	cacacctcgt	cgacttcttc	gatcgcatca	tcctctacac	caacacggg	900
aggagcacc	cacactgggc	aggccacgag	gtcatctcca	ggaggacggg	ccagggccag	960
ggcaacgaga	tcaggttccc	gctctacggc	gtggcagcaa	acacggagcc	accagtcacc	1020
atcaggccaa	cgggcttcac	cgacgagcag	cgccagtggg	acaggggtgct	gtcccgcatc	1080
gtcagcttca	ggtccagcgg	ccaggacttc	agcctcgtgg	atgccgtcgg	cttctgacc	1140
atcttctccg	cgctcagctt	ctaccgcaac	ggcttcggct	tcaacaccga	cacgatcaac	1200
gagatcccga	tcgagggcac	ggatccactg	accggctaca	gccacaggct	ctgccacgtg	1260
ggcttcgtcg	catccagccc	gttcatctcc	cagtacgcgc	gcgtgccagt	cttctcctgg	1320
accacagga	gcgcaaccct	gacgaacacc	atcgcaccag	acgtgatcac	gcagatccca	1380
ctgggtcaagg	ccttcaacct	ccactccggc	gcaacgggtg	tcagggggccc	aggcttcacc	1440
ggcggcgaca	tcctcaggag	gacgaacgtg	ggcagcttcg	gcgatatgcg	cgtcaacatc	1500
accgcaccac	tgtcccagag	gtacaggggtg	cgcattccgt	acgcgagcac	cacggacctc	1560
cagttccaca	cctccatcaa	cggcagggca	atcaaccagg	ccaacttccc	ggcgacgata	1620
gagaaccgcg	gcaacctgca	gttcggcgat	ttcaggaccg	tgggcttcac	cacgccattc	1680
acgttctccg	acgccaacag	cacgttcacc	atcggcgcg	ggagcttctc	cagcggcaac	1740
gaggtctaca	tcgatcgcat	cgagttcgtg	ccagcagagg	tcaccttcga	ggcagagtac	1800
aacctcaagc	aggcccagca	ggcgggtgaac	gccctgttca	cgtccagcaa	ccagatcggc	1860
ctcaagacgg	acgtcaccga	ttaccagata	gaccaggtga	gcaacctcgt	cgagtgcctg	1920
tccgacgagt	tctgcctgga	tgagaagcgc	gagctcagcg	agaagggtgaa	gcacgccaa	1980
aggctcagcg	acgagcgcaa	cctcctgcag	gattccaact	tcaggggaat	caacaggcag	2040
ccagacaggg	gctggagggg	ctccaccgac	atcatcatcc	agggcggcga	tggcgtgttc	2100
aaggagaact	acgtcacgct	gccgggcacc	ttcgacgagt	gctacccaac	ctacctctac	2160
cagaagatcg	atgagagcaa	gctgaaggcc	tacacgaggt	accagctccg	cggatacatc	2220
gaggactccc	aggatctcga	gatctacctg	atcaggtaca	acgcgaagca	cgagaccgtg	2280
aacgtcccag	gctccggcag	cctgaggcca	ctcagcgcac	agaacccaat	cggcaagtgc	2340

```

ggcgagccaa acaggtgcg accacacctg gagtgggaacc cagacctcga ttgctcctgc      2400
agggatggcg agaagtggcg acaccactcc caccacttca gcctggacat cgatgtgggc      2460
tgcaccgacc tcaacgagga tctgggcggtg tgggtcatct tcaaaatcaa gacgcaggac      2520
ggccacgcaa ggctgggcaa cctcgagttc ctggaggaga agccactcct gggcgaggca      2580
ctcgcaaggg tgaagagggc ggagaagaag tggcgcgata agagggagaa gctggagctc      2640
gagaccaaca tcgtctacaa ggaggccaag gagagcgtgg acgcgctgtt cgtcaactcc      2700
cagtacgaca ggctccaggc cgatacgaac atcgcgatga tccacgccgc ggataagcgc      2760
gtgcaccgca tcagggaggc ctacctcccg gagctgagcg tgatcccagg cgtcaacgca      2820
ggcatcttcg aggagctcga gggccgcctc ttcaccgcct acagccagta cgacgcgagg      2880
aacgtgatca agaacggcga tttcaacaac ggctgtcct gctggaacgt caagggccac      2940
gtggacgtcg aggagcagca caaccacagg agcgtgctcg tgggccaga gtgggaggca      3000
gaggtctccc aggaggtgag ggtctgccc aagcaggggat acatcctgag ggtgaccgca      3060
tacaaggagg gctacggcga gggcacggtc accatccacg agatcgagga ccacacggat      3120
gagctcaagt tcagcaactg cgtggaggag gaggtctacc cgaacaacac ggtgacctgc      3180
aacgattaca cggcaacca ggaggagtac ggcggcgcat acacctccag gaacaggggc      3240
tacgacgaga cgtacggctc caactccagc gtgagcgcg attacgcctc cgtctacgag      3300
gagaaggcat acaccgacgg caggagggag aacccatgcy agttcaacag gggctacggc      3360
gactacaccc cactgccagc aggctacgtg acgaaggagc tcgagtactt cccagagacc      3420
gataagctct ggatcgagat cggcgagacg gagggcacct tcctcgtgga ctccgctcag      3480
ctcctgctca tggaggagtg a                                          3501

```

```

<210> 16
<211> 781
<212> PRT
<213> Bacillus thuringiensis

```

```
<400> 16
```

```

Met Asn Pro Tyr Gln Asn Lys Asn Asp His Glu Ile Leu Asn Ala Thr
1           5           10           15

```

```

Gln Asn Asn Cys Arg Met Pro Asn Cys Tyr Ser Arg Tyr Pro Leu Ser
          20           25           30

```

Asn Asp Pro Gln Ile Ser Cys Gln Asn Asn His Tyr Lys Asp Trp Leu
 35 40 45

Asn Lys Tyr Glu Asn Ile Thr Thr Thr Ser Ser Leu Ser Gly Ile Leu
 50 55 60

Thr Ser Ile Gly Met Phe Asn Lys Val Leu Met Leu Thr Gly Val Leu
 65 70 75 80

Gly Val Thr Pro Ile Val Ile Asn Ile Ile Leu Asp Met Val Arg Leu
 85 90 95

Ile Asp Met Gly Thr Gly Asp Asp Leu Leu Ala His Thr Glu Gln Leu
 100 105 110

Ile Gln Gln Thr Leu Ala Ala Gln Tyr Arg Ser Ala Ala Thr Gly Ala
 115 120 125

Ile Tyr Gly Ile Ser Arg Ala Tyr Glu Asn Tyr Leu Met Phe Phe Lys
 130 135 140

Gln Trp Glu Arg Asn Arg Thr Pro Gln Asn Gly Gln Gln Val Glu Ser
 145 150 155 160

Ala Phe Thr Thr Val Asn Thr Leu Cys Ile Thr Ala Leu Ala Pro Gln
 165 170 175

Ala Leu Leu Ser Arg Arg Gly Phe Glu Thr Leu Leu Leu Pro Asn Tyr
 180 185 190

Ala Leu Ala Ala Asn Phe His Leu Leu Leu Leu Arg Asp Ala Val Ile
 195 200 205

Tyr Arg Asn Ser Trp Leu Pro Ser Asn Val Ser Ile Lys Ser Val Asn
 210 215 220

Leu Thr Ile Leu Arg Asn Ala Ile Ser Glu Tyr Arg Asn His Cys Asn
 225 230 235 240

His Trp Tyr Asn Asp Gly Leu Asn Arg Phe Ala Arg Ser Ser Phe Asn
 245 250 255

Asp Trp Val Arg Phe Asn Ala Tyr Arg Arg Asp Met Thr Leu Ser Val
 260 265 270

Leu Asp Phe Ile Thr Val Phe Pro Thr Tyr Asp Pro Ile Ile Tyr Pro
 275 280 285

Lys Ala Thr Asn Ile Glu Leu Thr Arg Ile Val Tyr Thr Asn Pro Ile
 290 295 300

Ile Pro Pro Arg Gly Phe Pro Arg Thr Asn Pro Pro Thr Phe Asn Gln
 305 310 315 320

Met Glu Asn Leu Ile Ile Ser Gly Cys Pro Ser Phe Leu Asn Gln Leu
 325 330 335

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

Phe Trp Ala Gly Asn Gln Asn Phe Leu Ser Asn Gly Thr Ile Arg Thr
 355 360 365

Ser Gly Ala Leu Ala Gly Arg Arg Ile Asp Phe Ser Met Gln Asn Ile
 370 375 380

Asp Val Phe Arg Val Asn Met Thr Thr His Asp Ile Asp Ser Leu Ser
 385 390 395 400

Arg Ser Tyr Gly Gly Val His Arg Ser Asp Phe Ile Gly Val Asn Ile
 405 410 415

Gln Asn Asn Gln Gly Thr Ser Ile Leu Tyr Asn Val Pro Leu Asp Thr
 420 425 430

Thr Arg Phe Met Thr Arg Asn Glu Thr Ala Phe Leu Pro Gly Asp Ser
 435 440 445

Gly Leu Glu Pro Asn Glu Arg Asn Tyr Thr His Arg Leu Phe Gln Val
 450 455 460

Met Thr Thr Tyr Arg Ser Asn Pro Asn Ala Arg Arg Ala Ala Phe Leu
 465 470 475 480

His Ala Trp Thr His Arg Ser Leu Arg Arg Arg Asn Gly Phe Arg Ala
485 490 495

Asp Gln Ile Ile Gln Ile Pro Ala Val Lys Ala Ile Asn Thr Ser Asp
500 505 510

Asp Ser Ala Val Val Ser Asn Val Gly Glu Asn Met Met Lys Leu Asp
515 520 525

Asn Leu Thr Thr Ser Leu Ser Tyr Lys Leu Thr Ala Glu Asp Ser Glu
530 535 540

Ala Ser Asn Thr Arg Phe Ile Val Arg Ile Arg Tyr Ala Ser Met Asn
545 550 555 560

Asn Asn Lys Leu Asn Leu Ile Leu Asn Gly Thr Gln Ile Ala Ser Leu
565 570 575

Asn Val Glu Gly Thr Met Gln Asn Gly Gly Ser Leu Thr Asp Leu Gln
580 585 590

Tyr Glu Asn Phe Lys Tyr Ala Ala Phe Ala Ser Asn Phe Lys Ile Gly
595 600 605

Ser Gln Ser Ile Ile Gly Ile Phe Lys Glu Thr Ser Asn Ala Asp Phe
610 615 620

Ile Leu Asp Lys Ile Glu Leu Ile Pro Val Asp Phe Ile Pro Leu Leu
625 630 635 640

Lys Gln Thr Gln Asp Cys Asn Asn Tyr Asn Gln Asp Thr Ile Tyr Thr
645 650 655

Cys Asn Gln Gly Tyr Asp Ala Tyr Glu Pro Asn Ala Ser Asp Ser Tyr
660 665 670

Asp Ser Val Tyr His Pro Asp Ala Asn Asp Ala Tyr Asp Glu Gly Asp
675 680 685

Asp Asn Thr Tyr Asn Gln Asn Tyr Asn Ser Met Tyr Asn Gln Asn Ala
690 695 700

Asp Asn Thr Tyr Asp Gln Ser Tyr Asn Asn Tyr Asp Pro Asn Met Glu
705 710 715 720

Asn Thr Tyr Gln Pro Ser Tyr Asp Asn Gly Tyr Glu Gln Asn Asn Tyr
725 730 735

Asp Ser Asp Asp Ser Asp Tyr Asn Asn Thr Tyr Asn Gln Asn Thr Asn
740 745 750

Asn Met Tyr Asn Gln Glu Tyr Asn Asn Asn Asp Asn Pro Asn Asn Gly
755 760 765

Cys Thr Cys Lys Gln Cys Tyr Asn Gln Asn Tyr Pro Lys
770 775 780

<210> 17
<211> 1169
<212> PRT
<213> *Bacillus thuringiensis*

<400> 17

Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Val Ile Asp Ala Pro His
1 5 10 15

Cys Gly Cys Pro Ser Glu Asp Ile Val Lys Tyr Pro Leu Thr Asp Asp
20 25 30

Pro Asn Ala Gly Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met
35 40 45

His Gly Gly Asp Tyr Thr Asp Pro Leu Ile Asn Pro Asn Leu Ser Val
50 55 60

Ser Gly Lys Asp Val Ile Gln Val Gly Ile Asn Ile Val Gly Arg Leu
65 70 75 80

Leu Ser Phe Phe Gly Phe Pro Phe Ser Ser Gln Trp Val Thr Val Tyr
85 90 95

Ser Tyr Leu Leu Asn Ser Leu Trp Pro Asp Asp Glu Asn Ser Val Trp
100 105 110

Asp Ala Phe Met Glu Arg Val Glu Glu Leu Ile Asp Gln Lys Ile Ser

115	120	125
Thr Leu Val Lys Gly Asn Ala Leu Asp Asn Leu Thr Gly Leu Gln Asn 130 135 140		
Asn Tyr Asn Leu Tyr Val Glu Ala Leu Asp Glu Trp Leu Asn Arg Pro 145 150 155 160		
Asn Gly Ala Arg Ala Ala Leu Val Ser Gln Arg Phe Tyr Ile Leu Asp 165 170 175		
Ser Leu Phe Thr Gln Phe Met Pro Ser Phe Gly Ser Gly Ser Gly Ser 180 185 190		
Gln Asn Tyr Ala Thr Ile Leu Leu Ala Val Tyr Ala Gln Ala Ala Asn 195 200 205		
Leu His Leu Leu Leu Leu Lys Asp Val Glu Ile Tyr Gly Ala Arg Trp 210 215 220		
Gly Leu Asn Gln Thr Gln Ile Asp Gln Phe His Ser Arg Gln Gln Ser 225 230 235 240		
Leu Thr Gln Thr Tyr Thr Asn His Cys Val Thr Ala Tyr Asn Asp Gly 245 250 255		
Leu Ala Glu Leu Arg Gly Thr Ser Val Glu Ser Trp Leu Lys Tyr His 260 265 270		
Gln Tyr Arg Arg Glu Met Thr Val Thr Ala Met Asp Leu Val Ala Leu 275 280 285		
Phe Pro Tyr Tyr Asn Val Arg Gln Tyr Pro Asn Gly Ala Asn Pro Gln 290 295 300		
Leu Thr Arg Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Leu Pro Glu 305 310 315 320		
Arg Pro Ser Gly Ala Phe Cys Glu Ser Phe Tyr Thr Ile Arg Ala Ala 325 330 335		
Arg Glu Arg Leu Thr Phe Ser Gln Leu Glu Asn Ala Ile Ile Arg Pro		

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

565 570 575
 Val Thr Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr Ala
 580 585 590
 Ser Thr Ile Asp Phe Asp Phe Phe Val Thr Arg Gly Gly Thr Thr Ile
 595 600 605
 Asn Asn Phe Arg Phe Thr Arg Thr Met Asn Ser Gly Gln Glu Ser Arg
 610 615 620
 Tyr Glu Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe Thr
 625 630 635 640
 Gln Ser Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn
 645 650 655
 Gly Glu Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro Thr
 660 665 670
 Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala Ser
 675 680 685
 Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp Tyr
 690 695 700
 Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu Gln
 705 710 715 720
 Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala Lys
 725 730 735
 Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn Ala
 740 745 750
 Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val Thr
 755 760 765
 Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala Leu Gln Leu Ala
 770 775 780
 Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val Asp Ala

785					790						795					800
Ser	Val	Leu	Lys	Pro	Tyr	Thr	Arg	Tyr	Arg	Leu	Asp	Gly	Phe	Val	Lys	
				805					810					815		
Ser	Ser	Gln	Asn	Leu	Glu	Ile	Asp	Leu	Ile	His	His	His	Lys	Val	His	
			820					825					830			
Leu	Val	Lys	Asn	Val	Pro	Asp	Asn	Leu	Val	Ser	Asp	Thr	Tyr	Ser	Asp	
		835					840					845				
Gly	Ser	Cys	Ser	Gly	Ile	Asn	Arg	Cys	Asp	Glu	Gln	His	Gln	Val	Asp	
	850					855					860					
Met	Gln	Leu	Asp	Ala	Glu	His	His	Pro	Met	Asp	Cys	Cys	Glu	Ala	Ala	
865					870					875					880	
Gln	Thr	His	Glu	Phe	Ser	Ser	Tyr	Ile	Asn	Thr	Gly	Asp	Leu	Asn	Ser	
				885					890					895		
Ser	Val	Asp	Gln	Gly	Ile	Trp	Val	Val	Leu	Lys	Val	Arg	Thr	Thr	Asp	
			900					905					910			
Gly	Tyr	Ala	Thr	Leu	Gly	Asn	Leu	Glu	Leu	Val	Glu	Val	Gly	Pro	Leu	
		915					920					925				
Ser	Gly	Glu	Ser	Leu	Glu	His	Glu	Lys	Arg	Lys	Asn	Ala	Glu	Trp	Asn	
	930					935					940					
Ala	Gly	Leu	Gly	Arg	Lys	Arg	Thr	Glu	Thr	Asp	Arg	Val	Tyr	Gln	Ala	
945					950					955					960	
Ala	Lys	Gln	Ala	Ile	Asn	His	Leu	Phe	Val	Asp	Tyr	Gln	Asp	Gln	Gln	
				965					970					975		
Leu	Asn	Pro	Glu	Ile	Gly	Leu	Thr	Glu	Ile	Asn	Glu	Val	Ser	Asn	Leu	
			980					985					990			
Val	Glu	Ser	Ile	Thr	Asp	Val	Tyr	Ser	Asp	Thr	Val	Leu	Gln	Ile	Pro	
		995					1000					1005				
Gly	Ile	Asn	Tyr	Glu	Ile	Tyr	Lys	Glu	Leu	Ser	Asp	Arg	Leu	Gln		

[illegible]

Thr Ile Gln Lys Glu Trp Thr Glu Trp Lys Arg Thr Asp His Ser Leu
35 40 45

Tyr Val Ser Pro Ile Val Gly Thr Val Ala Ser Leu Leu Leu Lys Lys
50 55 60

Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Trp Gly Leu
65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Ala
85 90 95

Thr Glu Gln Phe Ile Asn Gln Arg Leu Asn Ala Asp Thr Leu Ala Arg
100 105 110

Ile Asn Val Glu Leu Val Gly Leu Gln Ala Asn Val Arg Glu Phe Asn
115 120 125

Gln Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Ala Val Pro Leu
130 135 140

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

Leu Pro Gln Phe Gln Ile Gln Gly Tyr Glu Leu Leu Leu Leu Pro Leu
165 170 175

Phe Ala Gln Ala Ala Asn Met His Leu Ser Phe Ile Arg Asp Val Ile
180 185 190

Leu Asn Ala Asp Glu Trp Asp Ile Ser Pro Ala Met Leu Arg Thr Tyr
195 200 205

Arg Asp His Leu Lys Arg Tyr Thr Asn Glu Tyr Ser Asn Tyr Cys Ile
210 215 220

Asn Thr Tyr Arg Thr Glu Phe Glu Gly Leu Asn Val Pro Leu Lys Asp
225 230 235 240

Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
245 250 255

Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Thr Ser Gly
260 265 270

Ala Asn Leu Tyr Val Ser Gly Thr Thr Val Gly Ala Asn Gln Ser Phe
275 280 285

Thr Ala Arg Asn Trp Pro Phe Leu Asn Thr Leu Phe Gln Val Asn Ser
290 295 300

Asn Tyr Ile Leu Thr Gly Phe Ser Gly Ala Arg Phe Gln Ser Thr Leu
305 310 315 320

Leu Asn Ser Phe Val Ile Thr Leu Leu Ala Ala Arg Ile His Tyr Met
325 330 335

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

Ser Cys Asn Leu Ser Ser Pro Ile Gln Leu Pro Tyr Arg Thr Ser Tyr
355 360 365

Leu Asp Ser Phe Ile Leu Arg Asp Gly Asn Arg Thr Asn Glu Tyr Phe
370 375 380

Thr Gly Ala Phe Leu Ala Leu Ile Leu Pro Cys Gly Val Thr Gly Pro
385 390 395 400

Thr Leu Gly Ala Ala Asn Ser Phe Pro Asn Tyr Ser Ile Arg Asn Ile
405 410 415

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

His Phe Asn Glu Ile Arg Asn Ile Glu Gly Gln Pro Gly Val Gly Gly
435 440 445

Gly Ser Arg Ile Ser Ser Leu Val Ser Val His Asn Arg Lys Asn Asn
450 455 460

Ile Val Ala His Glu Asn Gly Thr Met Ile His Leu Ala Pro Asp Asp
465 470 475 480

Tyr Thr Gly Phe Thr Ile Ser Pro Ile His Ala Thr Gln Val Asn Asn
485 490 495

Gln Thr Arg Thr Ile Ile Ser Glu Lys Phe Gly Asn Gln Gly Asp Ser
500 505 510

Leu Arg Phe Glu Gln Ser Asn Thr Thr Ala Arg Tyr Thr Leu Arg Gly
515 520 525

Asn Gly Asn Ser Tyr Asn Leu Tyr Leu Arg Val Ser Ser Ile Gly Asn
530 535 540

Ser Thr Ile Arg Val Thr Ile Asn Gly Arg Val Tyr Thr Ala Thr Asn
545 550 555 560

Val Asn Thr Thr Thr Asn Asn Asp Gly Val Leu Asp Asn Gly Ala Arg
565 570 575

Phe Ser Asp Ile Asn Ile Gly Asn Val Val Ala Ser Ala Asn Thr Asn
580 585 590

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

Leu Met Asn Ile Met Phe Val Pro Thr Asn Leu Pro Pro Val Tyr
610 615 620

<210> 19

<211> 640

<212> PRT

<213> Bacillus thuringiensis

<400> 19

Met Asp Ser Val Leu Asn Ser Gly Arg Ile Thr Asn Arg Asp Ala Tyr
1 5 10 15

Asn Val Val Ala Leu Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp
20 25 30

Thr Ile Gln Lys Glu Trp Thr Glu Trp Lys Lys Asn Asn His Ser Leu
35 40 45

Tyr Val Asp Pro Ile Val Gly Thr Val Ala Ser Phe Leu Leu Lys Lys
 50 55 60

Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Arg Asn Leu
 65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Glu
 85 90 95

Thr Glu Lys Phe Leu Asn Gln Arg Leu Asn Thr Asp Thr Leu Ala Arg
 100 105 110

Val Asn Ala Glu Leu Thr Gly Met Gln Ala Asn Val Asp Glu Phe Asn
 115 120 125

Arg Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Ala Val Pro Leu
 130 135 140

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

Leu Pro Gln Phe Gln Met Gln Gly Tyr Gln Leu Leu Leu Leu Pro Leu
 165 170 175

Phe Ala Gln Ala Ala Asn Leu His Leu Ser Phe Ile Arg Asp Val Ile
 180 185 190

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

Arg Asp Tyr Leu Lys Asn Tyr Thr Arg Glu Tyr Ser Asn Tyr Cys Ile
 210 215 220

Asn Thr Tyr Gln Ser Ala Phe Arg Gly Leu Asn Thr Ser Leu His Asp
 225 230 235 240

Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
 245 250 255

Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly
 260 265 270

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

Thr Gln Thr Tyr Thr Ser Gln Asp Trp Pro Phe Leu Asn Ser Leu Phe
 290 295 300

Gln Val Asn Ser Asn Tyr Val Leu Ser Gly Tyr Ser Gly Ala Ser Leu
 305 310 315 320

Phe Ile Ser Leu Pro Asn Ile Gly Leu Pro Pro Gly Ser Thr Thr Thr
 325 330 335

Gln Ala Leu Leu Gly Ala Arg Val Asn Tyr Cys Gly Gly Ile Ser Ser
 340 345 350

Gly Phe Ile Gly Val Ser Asn Phe Asn Gln Asn Phe Asn Cys Ser Ala
 355 360 365

Gly Pro Ser Pro Leu Ser Thr Pro Phe Ile Arg Ser Trp Leu Asp Ala
 370 375 380

Gly Ser Asp Arg Gln Gly Ala Asn Thr Ser Thr Asn Trp Gln Thr Glu
 385 390 395 400

Ala Phe Gln Thr Thr Gln Gly Arg Ser Cys Gly Ala Phe Gly Ala Arg
 405 410 415

Gly Asn Ser Asn Phe Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly
 420 425 430

Val Pro Leu Val Val Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr
 435 440 445

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

Arg Ala Tyr Met Val Ser Val His Asn Arg Lys Asn Asn Ile Tyr Ala
 465 470 475 480

Val His Glu Asn Gly Thr Met Ile His Leu Ala Pro Glu Asp Tyr Thr
 485 490 495

Gly Phe Thr Ile Ser Pro Ile His Ala Thr Gln Val Asn Ser Gln Ala
500 505 510

Asn Gln Ala Arg Thr Phe Ile Ser Glu Lys Phe Gly Asn Gln Gly Asp
515 520 525

Ser Leu Arg Val Glu Gln Ser Thr Ser Thr Val Gln Tyr Thr Phe Arg
530 535 540

Gly Asn Gly His Ser Tyr Asn Leu Tyr Leu Arg Val Ser Ser Ile Gly
545 550 555 560

Ser Ser Thr Ile Arg Val Gly Leu Asn Gly Thr Val Tyr Thr Val Arg
565 570 575

Asp Val Asn Thr Thr Thr Asn Asn Asp Gly Val Leu Asp Asn Gly Ala
580 585 590

Arg Phe Ser Asp Ile Tyr Ile Gly Asn Val Val Ala Ser Ala Asn Thr
595 600 605

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

Glu Leu Met Asn Ile Met Phe Val Pro Thr Asn Leu Pro Pro Val Tyr
625 630 635 640

<210> 20

<211> 1166

<212> PRT

<213> Bacillus thuringiensis

<400> 20

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

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

Val Ser Gln Val Ser Val Gly Leu Thr Arg Phe Leu Leu Glu Ala Ala
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 Glu Gln Ile Glu Gln
 65 70 75 80

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

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

Arg Glu Trp Glu Asn Asp Pro Asp Asn Pro Ala Ser Gln 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 Ala Ile Pro Asn Tyr Glu Ile Ala Thr Leu Ser Val Tyr Val
 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 Pro Val Asn Ile Glu Asp Leu Tyr Thr
 180 185 190

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

Tyr Asn Gln Gly Leu Asn Ser Ile Gly Ala Thr Asn Thr Arg Tyr Leu
 210 215 220

Asp Phe Gln Arg Asp 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 Asp Gly Val Asn
 260 265 270

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

Phe Phe Asp Arg Ile Ile Ile Tyr Thr Asn Thr Val Arg Ser Thr Pro
 290 295 300

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

Gly Asn Glu Ile Arg Phe Pro Leu Tyr Gly Val Ala Ala Asn Thr Glu
 325 330 335

Pro Pro Val Thr Ile Arg Pro Thr Gly Phe Thr Asp Glu Gln Arg Gln
 340 345 350

Trp Tyr Arg Val Leu Ser Arg Ile Val Ser Phe Arg Ser Ser Gly Gln
 355 360 365

Asp Phe Ser Leu Val Asp Ala Val Gly Phe Leu Thr Ile Phe Ser Ala
 370 375 380

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

Glu Ile Pro Ile Glu Gly Thr Asp Pro Leu Thr Gly Tyr Ser His Arg
 405 410 415

Leu Cys His Val Gly Phe Val Ala Ser Ser Pro Phe Ile Ser Gln Tyr
 420 425 430

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

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

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

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

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

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

Arg Ala Ile Asn Gln Ala Asn Phe Pro Ala Thr Ile Glu Asn Arg Gly
 530 535 540

Asn Leu Gln Phe Gly Asp Phe Arg Thr Val Gly Phe Thr Thr Pro Phe
 545 550 555 560

Thr Phe Ser Asp Ala Asn Ser Thr Phe Thr Ile Gly Ala Trp Ser Phe
 565 570 575

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

Glu Val Thr Phe Glu Ala Glu Tyr Asn Leu Lys Gln Ala Gln Gln Ala
 595 600 605

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

Val Thr Asp Tyr Gln Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu
 625 630 635 640

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

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

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

Thr Asp Ile Ile Ile Gln Gly Gly Asp Gly Val Phe Lys Glu Asn Tyr
 690 695 700

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

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

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

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

Arg Pro Leu Ser Ala Gln Asn Pro Ile Gly Lys Cys Gly Glu Pro Asn
 770 775 780

Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys
 785 790 795 800

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

Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val
 820 825 830

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

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

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Leu
 865 870 875 880

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

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

Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr
 915 920 925

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

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

Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn
 965 970 975

Val Lys Gly His Val Asp Val Glu Glu Gln His Asn His Arg Ser Val
 980 985 990

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

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

Gly Tyr Gly Glu Gly Thr Val Thr Ile His Glu Ile Glu Asp His
 1025 1030 1035

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

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

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

Thr Tyr Gly Ser Asn Ser Ser Val Ser Ala Asp Tyr Ala Ser Val
 1085 1090 1095

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

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

Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Ile
 1130 1135 1140

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

Val Glu Leu Leu Leu Met Glu Glu
1160 1165

<210> 21

<211> 781

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Protein BT-0049 đột biến.

<400> 21

Met Asn Pro Tyr Gln Asn Lys Asn Asp His Glu Ile Leu Asn Ala Thr
1 5 10 15

Gln Asn Asn Cys Arg Met Pro Asn Cys Tyr Ser Arg Tyr Pro Leu Ser
20 25 30

Asn Asp Pro Gln Ile Ser Cys Gln Asn Asn His Tyr Lys Asp Trp Leu
35 40 45

Asn Lys Tyr Glu Asn Ile Thr Thr Thr Ser Ser Leu Ser Gly Ile Leu
50 55 60

Thr Ser Ile Gly Met Phe Asn Lys Val Leu Met Leu Thr Gly Val Leu
65 70 75 80

Gly Val Thr Pro Ile Val Ile Asn Ile Ile Leu Asp Met Val Arg Leu
85 90 95

Ile Asp Met Gly Thr Gly Asp Asp Leu Leu Ala His Thr Glu Gln Leu
100 105 110

Ile Gln Gln Thr Leu Ala Ala Gln Tyr Arg Ser Ala Ala Thr Gly Ala
115 120 125

Ile Tyr Gly Ile Ser Arg Ala Tyr Glu Asn Tyr Leu Met Phe Phe Lys
130 135 140

Gln Trp Glu Arg Asn Arg Thr Pro Gln Asn Gly Gln Gln Val Glu Ser
145 150 155 160

Ala Phe Thr Thr Val Asn Thr Leu Cys Ile Thr Ala Leu Ala Pro Gln
165 170 175

Ala Leu Leu Ser Arg Arg Gly Phe Glu Thr Leu Leu Leu Pro Asn Tyr
 180 185 190

Ala Leu Ala Ala Asn Phe His Leu Leu Leu Leu Arg Asp Ala Val Ile
 195 200 205

Tyr Arg Asn Ser Trp Leu Pro Ser Asn Val Ser Ile Lys Ser Val Asn
 210 215 220

Leu Thr Ile Leu Arg Asn Ala Ile Ser Glu Tyr Arg Asn His Cys Asn
 225 230 235 240

His Trp Tyr Asn Asp Gly Leu Asn Arg Phe Ala Arg Ser Ser Phe Asn
 245 250 255

Asp Trp Val Arg Phe Asn Ala Tyr Arg Arg Asp Met Thr Leu Ser Val
 260 265 270

Leu Asp Phe Ile Thr Val Phe Pro Thr Tyr Asp Pro Ile Ile Tyr Pro
 275 280 285

Lys Ala Thr Asn Ile Glu Leu Thr Arg Ile Val Tyr Thr Asn Pro Ile
 290 295 300

Ile Pro Pro Arg Gly Phe Pro Arg Thr Asn Pro Pro Thr Phe Asn Gln
 305 310 315 320

Met Glu Asn Leu Ile Ile Ser Gly Cys Pro Ser Phe Leu Asn Gln Leu
 325 330 335

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

Phe Trp Ala Gly Asn Gln Asn Phe Leu Ser Asn Gly Thr Ile Arg Thr
 355 360 365

Ser Gly Ala Leu Ala Gly Arg Arg Ile Asp Phe Ser Met Gln Asn Ile
 370 375 380

Asp Val Phe Arg Val Asn Met Thr Thr His Asp Ile Asp Ser Leu Ser
 385 390 395 400

Arg Ser Tyr Gly Gly Val His Arg Ser Asp Phe Ile Gly Val Asn Ile
 405 410 415

Gln Asn Asn Gln Gly Thr Ser Ile Leu Tyr Asn Val Pro Leu Asp Thr
 420 425 430

Thr Arg Phe Met Thr Arg Asn Glu Thr Ala Phe Leu Pro Gly Asp Ser
 435 440 445

Gly Leu Glu Pro Asn Glu Arg Asn Tyr Thr His Arg Leu Phe Gln Val
 450 455 460

Met Thr Thr Tyr Arg Ser Asn Pro Asn Ala Arg Arg Ala Ala Phe Leu
 465 470 475 480

His Ala Trp Thr His Arg Ser Leu Arg Arg Arg Asn Gly Phe Arg Ala
 485 490 495

Asp Gln Ile Ile Gln Ile Pro Ala Val Lys Ala Ile Asn Thr Ser Asp
 500 505 510

Asp Ser Ala Val Val Ser Asn Val Gly Glu Asn Met Met Lys Leu Asp
 515 520 525

Asn Leu Thr Thr Ser Leu Ser Tyr Lys Leu Thr Ala Glu Asp Ser Glu
 530 535 540

Ala Ser Asn Thr Arg Phe Ile Val Arg Ile Arg Tyr Ala Ser Met Asn
 545 550 555 560

Asn Asn Lys Leu Asn Leu Ile Leu Asn Gly Thr Gln Ile Ala Ser Leu
 565 570 575

Asn Val Glu Gly Thr Met Gln Asn Gly Gly Ser Leu Thr Asp Leu Gln
 580 585 590

Tyr Glu Asn Phe Lys Tyr Ala Ala Phe Ala Ser Asn Phe Lys Ile Gly
 595 600 605

Ser Gln Ser Ile Ile Gly Ile Phe Lys Glu Thr Ser Asn Ala Asp Phe
 610 615 620

Ile Leu Asp Lys Ile Glu Leu Ile Pro Val Asp Phe Leu Pro Leu Leu
625 630 635 640

Lys Gln Thr Gln Asp Cys Asn Asn Tyr Asn Gln Asp Thr Leu Tyr Thr
645 650 655

Cys Asn Gln Gly Tyr Asp Ala Tyr Glu Pro Asn Ala Ser Asp Ser Tyr
660 665 670

Asp Ser Val Tyr His Pro Asp Ala Asn Asp Ala Tyr Asp Glu Gly Asp
675 680 685

Asp Asn Thr Tyr Asn Gln Asn Tyr Asn Ser Met Tyr Asn Gln Asn Ala
690 695 700

Asp Asn Thr Tyr Asp Gln Ser Tyr Asn Asn Tyr Asp Pro Asn Met Glu
705 710 715 720

Asn Thr Tyr Gln Pro Ser Tyr Asp Asn Gly Tyr Glu Gln Asn Asn Tyr
725 730 735

Asp Ser Asp Asp Ser Asp Tyr Asn Asn Thr Tyr Asn Gln Asn Thr Asn
740 745 750

Asn Met Tyr Asn Gln Glu Tyr Asn Asn Asn Asp Asn Pro Asn Asn Gly
755 760 765

Cys Thr Cys Lys Gln Cys Tyr Asn Gln Asn Tyr Pro Lys
770 775 780

<210> 22

<211> 1169

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Protein BT-0063 đột biến.

<400> 22

Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Val Ile Asp Ala Pro His
1 5 10 15

Cys Gly Cys Pro Ser Glu Asp Ile Val Lys Tyr Pro Leu Thr Asp Asp
 20 25 30

Pro Asn Ala Gly Leu Gln Asn Met Asn Tyr Lys Glu Tyr Leu Gln Met
 35 40 45

His Gly Gly Asp Tyr Thr Asp Pro Leu Ile Asn Pro Asn Leu Ser Val
 50 55 60

Ser Gly Lys Asp Val Ile Gln Val Gly Ile Asn Ile Val Gly Arg Leu
 65 70 75 80

Leu Ser Phe Phe Gly Phe Pro Phe Ser Ser Gln Trp Val Thr Val Tyr
 85 90 95

Ser Tyr Leu Leu Asn Ser Leu Trp Pro Asp Asp Glu Asn Ser Val Trp
 100 105 110

Asp Ala Phe Met Glu Arg Val Glu Glu Leu Ile Asp Gln Lys Ile Ser
 115 120 125

Thr Leu Val Lys Gly Asn Ala Leu Asp Asn Leu Thr Gly Leu Gln Asn
 130 135 140

Asn Tyr Asn Leu Tyr Val Glu Ala Leu Asp Glu Trp Leu Asn Arg Pro
 145 150 155 160

Asn Gly Ala Arg Ala Ala Leu Val Ser Gln Arg Phe Tyr Ile Leu Asp
 165 170 175

Ser Leu Phe Thr Gln Phe Met Pro Ser Phe Gly Ser Gly Ser Gly Ser
 180 185 190

Gln Asn Tyr Ala Thr Ile Leu Leu Ala Val Tyr Ala Gln Ala Ala Asn
 195 200 205

Leu His Leu Leu Leu Leu Lys Asp Val Glu Ile Tyr Gly Ala Arg Trp
 210 215 220

Gly Leu Asn Gln Thr Gln Ile Asp Gln Phe His Ser Arg Gln Gln Ser
 225 230 235 240

Leu Thr Gln Thr Tyr Thr Asn His Cys Val Thr Ala Tyr Asn Asp Gly
 245 250 255

Leu Ala Glu Leu Arg Gly Thr Ser Val Glu Ser Trp Leu Lys Tyr His
 260 265 270

Gln Tyr Arg Arg Glu Met Thr Val Thr Ala Met Asp Leu Val Ala Leu
 275 280 285

Phe Pro Tyr Tyr Asn Val Arg Gln Tyr Pro Asn Gly Ala Asn Pro Gln
 290 295 300

Leu Thr Arg Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Leu Pro Glu
 305 310 315 320

Arg Pro Ser Gly Ala Phe Cys Glu Ser Phe Tyr Thr Ile Arg Ala Ala
 325 330 335

Arg Glu Arg Leu Thr Phe Ser Gln Leu Glu Asn Ala Ile Ile Arg Pro
 340 345 350

Pro Arg Leu Phe Glu Arg Phe Gln Ala Leu Gly Ile Tyr Thr Gly Glu
 355 360 365

Ala Arg Leu Asn Ala Asn Ser Ala Pro Met Asn Tyr Trp Ile Gly His
 370 375 380

Phe Ile Arg Asn Thr Arg Leu Gly Asp Ser Thr Thr Ile Thr Thr Asn
 385 390 395 400

Tyr Gly Thr Val Asn Asn Arg Leu Thr Asn Phe Ile Pro Pro Thr Thr
 405 410 415

Ser Asp Val Tyr Gln Ile Asn Ser Ile Ser Ser Asn Leu Ala Ala Ile
 420 425 430

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

Gly Ile Ile Trp Ser Tyr Val Gly Gln Asn Ser Val Leu Ala Gln Cys
 450 455 460

His Gln Asn Tyr Asn Ser Ile Glu Glu Leu Pro Asn Gln Ser Asp Glu
 465 470 475 480

Pro Thr Val Arg Ser Tyr Ser His Arg Leu Ser His Ile Thr Ser Phe
 485 490 495

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

Pro Val Tyr Val Trp Thr His Arg Ser Val Asp Leu Asn Asn Thr Ile
 515 520 525

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

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

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

Val Thr Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr Ala
 580 585 590

Ser Thr Ile Asp Phe Asp Phe Phe Val Thr Arg Gly Gly Thr Thr Ile
 595 600 605

Asn Asn Phe Arg Phe Thr Arg Thr Met Asn Ser Gly Gln Glu Ser Arg
 610 615 620

Tyr Glu Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe Thr
 625 630 635 640

Gln Ser Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn
 645 650 655

Gly Glu Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro Thr
 660 665 670

Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala Ser
 675 680 685

Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp Tyr
 690 695 700

Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu Gln
 705 710 715 720

Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala Lys
 725 730 735

Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn Ala
 740 745 750

Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val Thr
 755 760 765

Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala Leu Gln Leu Ala
 770 775 780

Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val Asp Ala
 785 790 795 800

Ser Val Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe Val Lys
 805 810 815

Ser Ser Gln Asn Leu Glu Ile Asp Leu Ile His His His Lys Val His
 820 825 830

Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser Asp Thr Tyr Ser Asp
 835 840 845

Gly Ser Cys Ser Gly Ile Asn Arg Cys Asp Glu Gln His Gln Val Asp
 850 855 860

Met Gln Leu Asp Ala Glu His His Pro Met Asp Cys Cys Glu Ala Ala
 865 870 875 880

Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp Leu Asn Ser
 885 890 895

Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg Thr Thr Asp
 900 905 910

Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val Gly Pro Leu
 915 920 925

Ser Gly Glu Ser Leu Glu His Glu Lys Arg Lys Asn Ala Glu Trp Asn
 930 935 940

Ala Gly Leu Gly Arg Lys Arg Thr Glu Thr Asp Arg Val Tyr Gln Ala
 945 950 955 960

Ala Lys Gln Ala Ile Asn His Leu Phe Val Asp Tyr Gln Asp Gln Gln
 965 970 975

Leu Asn Pro Glu Ile Gly Leu Thr Glu Ile Asn Glu Val Ser Asn Leu
 980 985 990

Val Glu Ser Ile Thr Asp Val Tyr Ser Asp Thr Val Leu Gln Ile Pro
 995 1000 1005

Gly Ile Asn Tyr Glu Ile Tyr Lys Glu Leu Ser Asp Arg Leu Gln
 1010 1015 1020

Gln Ala Ser Asn Leu Tyr Thr Ser Arg Asn Ala Val Gln Asn Gly
 1025 1030 1035

Asp Phe Asp Ser Gly Leu Asp Gly Trp Asn Ala Thr Thr Asp Thr
 1040 1045 1050

Ser Val Gln Gln Asp Gly Asn Met His Phe Leu Val Leu Ser His
 1055 1060 1065

Trp Asp Ala Gln Val Ser Gln Gln Leu Arg Val His Pro Asn Cys
 1070 1075 1080

Lys Tyr Val Leu Arg Val Thr Ala Arg Lys Val Gly Gly Gly Asp
 1085 1090 1095

Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Gln Lys Thr Leu
 1100 1105 1110

Thr Phe Asn Ala Cys Asp Tyr Asp Val Asn Gly Thr Tyr Val Asp
 1115 1120 1125

Asp Asn Thr Tyr Ile Thr Lys Asp Val Leu Phe Tyr Pro Glu Thr
 1130 1135 1140

Lys His Met Trp Val Glu Val Ser Glu Ser Glu Gly Ser Phe Tyr
 1145 1150 1155

Leu Asp Ser Ile Glu Phe Ile Glu Thr Gln Glu
 1160 1165

<210> 23

<211> 623

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Protein BT-0064 đột biến.

<400> 23

Met Asn Asn Ile Leu Asn Asn Glu Arg Thr Thr Ile Cys Asp Ala His
 1 5 10 15

Asn Val Val Ala His Asp Pro Phe Ser Phe Glu His Lys Ser Val Asp
 20 25 30

Thr Ile Gln Lys Glu Trp Thr Glu Trp Lys Arg Thr Asp His Ser Leu
 35 40 45

Tyr Val Ser Pro Ile Val Gly Thr Val Ala Ser Leu Leu Lys Lys
 50 55 60

Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Trp Gly Leu
 65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Ala
 85 90 95

Thr Glu Gln Phe Ile Asn Gln Arg Leu Asn Ala Asp Thr Leu Ala Arg
 100 105 110

Ile Asn Val Glu Leu Val Gly Leu Gln Ala Asn Val Arg Glu Phe Asn
 115 120 125

Gln Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Ala Val Pro Leu
 130 135 140

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

Leu Pro Gln Phe Gln Ile Gln Gly Tyr Glu Leu Leu Leu Leu Pro Leu
 165 170 175

Phe Ala Gln Ala Ala Asn Met His Leu Ser Phe Ile Arg Asp Val Ile
 180 185 190

Leu Asn Ala Asp Glu Trp Asp Ile Ser Pro Ala Met Leu Arg Thr Tyr
 195 200 205

Arg Asp His Leu Lys Arg Tyr Thr Asn Glu Tyr Ser Asn Tyr Cys Ile
 210 215 220

Asn Thr Tyr Arg Thr Glu Phe Glu Gly Leu Asn Val Pro Leu Lys Asp
 225 230 235 240

Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
 245 250 255

Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Thr Ser Gly
 260 265 270

Ala Asn Leu Tyr Val Ser Gly Thr Thr Val Gly Ala Asn Gln Ser Phe
 275 280 285

Thr Ala Arg Asn Trp Pro Phe Leu Asn Thr Leu Phe Gln Val Asn Ser
 290 295 300

Asn Tyr Ile Leu Thr Gly Phe Ser Gly Ala Arg Phe Gln Ser Thr Leu
 305 310 315 320

Leu Asn Ser Phe Val Ile Thr Leu Leu Ala Ala Arg Ile His Tyr Met
 325 330 335

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

Ser Cys Asn Leu Ser Ser Pro Ile Gln Leu Pro Tyr Arg Thr Ser Tyr
 355 360 365

Leu Asp Ser Phe Ile Leu Arg Asp Gly Asn Arg Thr Asn Glu Tyr Phe
 370 375 380

Thr Gly Ala Phe Leu Ala Leu Ile Leu Pro Cys Gly Val Thr Gly Pro
 385 390 395 400

Thr Leu Gly Ala Ala Asn Ser Phe Pro Asn Tyr Ser Ile Arg Asn Ile
 405 410 415

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

His Phe Asn Glu Ile Arg Asn Ile Glu Gly Gln Pro Gly Val Gly Gly
 435 440 445

Gly Ser Arg Ile Ser Ser Leu Val Ser Val His Asn Arg Lys Asn Asn
 450 455 460

Ile Val Ala His Glu Asn Gly Thr Met Ile His Leu Ala Pro Asp Asp
 465 470 475 480

Tyr Thr Gly Phe Thr Ile Ser Pro Ile His Ala Thr Gln Val Asn Asn
 485 490 495

Gln Thr Arg Thr Ile Ile Ser Glu Lys Phe Gly Asn Gln Gly Asp Ser
 500 505 510

Leu Arg Phe Glu Gln Ser Asn Thr Thr Ala Arg Tyr Thr Leu Arg Gly
 515 520 525

Asn Gly Asn Ser Tyr Asn Leu Tyr Leu Arg Val Ser Ser Ile Gly Asn
 530 535 540

Ser Thr Ile Arg Val Thr Ile Asn Gly Arg Val Tyr Thr Ala Thr Asn
 545 550 555 560

Val Asn Thr Thr Thr Asn Asn Asp Gly Val Leu Asp Asn Gly Ala Arg
 565 570 575

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

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

Leu Met Asn Ile Met Phe Val Pro Thr Asn Leu Pro Pro Val Tyr
610 615 620

<210> 24

<211> 640

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Protein BT-0086 đột biến.

<400> 24

Met Asp Ser Val Leu Asn Ser Gly Arg Ile Thr Asn Arg Asp Ala Tyr
1 5 10 15

Asn Val Val Ala Leu Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp
20 25 30

Thr Ile Gln Lys Glu Trp Thr Glu Trp Lys Lys Asn Asn His Ser Leu
35 40 45

Tyr Val Asp Pro Ile Val Gly Thr Val Ala Ser Phe Leu Leu Lys Lys
50 55 60

Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Arg Asn Leu
65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Glu
85 90 95

Thr Glu Lys Phe Leu Asn Gln Arg Leu Asn Thr Asp Thr Leu Ala Arg
100 105 110

Val Asn Ala Glu Leu Thr Gly Met Gln Ala Asn Val Asp Glu Phe Asn
115 120 125

Arg Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Ala Val Pro Leu
130 135 140

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

Leu Pro Gln Phe Gln Met Gln Gly Tyr Gln Leu Leu Leu Leu Pro Leu
 165 170 175

Phe Ala Gln Ala Ala Asn Leu His Leu Ser Phe Ile Arg Asp Val Ile
 180 185 190

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

Arg Asp Tyr Leu Lys Asn Tyr Thr Arg Glu Tyr Ser Asn Tyr Cys Ile
 210 215 220

Asn Thr Tyr Gln Ser Ala Phe Arg Gly Leu Asn Thr Ser Leu His Asp
 225 230 235 240

Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
 245 250 255

Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly
 260 265 270

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

Thr Gln Thr Tyr Thr Ser Gln Asp Trp Pro Phe Leu Asn Ser Leu Phe
 290 295 300

Gln Val Asn Ser Asn Tyr Val Leu Ser Gly Tyr Ser Gly Ala Ser Leu
 305 310 315 320

Phe Ile Ser Leu Pro Asn Ile Gly Leu Pro Pro Gly Ser Thr Thr Thr
 325 330 335

Gln Ala Leu Leu Gly Ala Arg Val Asn Tyr Cys Gly Gly Ile Ser Ser
 340 345 350

Gly Phe Ile Gly Val Ser Asn Phe Asn Gln Asn Phe Asn Cys Ser Ala
 355 360 365

Gly Pro Ser Pro Leu Ser Thr Pro Phe Ile Arg Ser Trp Leu Asp Ala
 370 375 380

Gly Ser Asp Arg Gln Gly Ala Asn Thr Ser Thr Asn Trp Gln Thr Glu
 385 390 395 400

Ala Phe Gln Thr Thr Gln Gly Arg Ser Cys Gly Ala Phe Gly Ala Arg
 405 410 415

Gly Asn Ser Asn Phe Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly
 420 425 430

Val Pro Leu Val Val Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr
 435 440 445

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

Arg Ala Tyr Met Val Ser Val His Asn Arg Lys Asn Asn Ile Tyr Ala
 465 470 475 480

Val His Glu Asn Gly Thr Met Ile His Leu Ala Pro Glu Asp Tyr Thr
 485 490 495

Gly Phe Thr Ile Ser Pro Ile His Ala Thr Gln Val Asn Ser Gln Ala
 500 505 510

Asn Gln Ala Arg Thr Phe Ile Ser Glu Lys Phe Gly Asn Gln Gly Asp
 515 520 525

Ser Leu Arg Val Glu Gln Ser Thr Ser Thr Val Gln Tyr Thr Phe Arg
 530 535 540

Gly Asn Gly His Ser Tyr Asn Leu Tyr Leu Arg Val Ser Ser Ile Gly
 545 550 555 560

Ser Ser Thr Ile Arg Val Gly Leu Asn Gly Thr Val Tyr Thr Val Arg
 565 570 575

Asp Val Asn Thr Thr Thr Asn Asn Asp Gly Val Leu Asp Asn Gly Ala
 580 585 590

Arg Phe Ser Asp Ile Tyr Ile Gly Asn Val Val Ala Ser Ala Asn Thr
595 600 605

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

Glu Leu Met Asn Leu Met Phe Val Pro Thr Asn Leu Pro Pro Val Tyr
625 630 635 640

<210> 25

<211> 1166

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Protein BT-0498 đột biến.

<400> 25

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

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

Val Ser Gln Val Ser Val Gly Leu Thr Arg Phe Leu Leu Glu Ala Ala
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 Glu Gln Ile Glu Gln
65 70 75 80

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

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

Arg Glu Trp Glu Asn Asp Pro Asp Asn Pro Ala Ser Gln 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 Ala Ile Pro Asn Tyr Glu Ile Ala Thr Leu Ser Val Tyr Val
 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 Pro Val Asn Ile Glu Asp Leu Tyr Thr
 180 185 190

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

Tyr Asn Gln Gly Leu Asn Ser Ile Gly Ala Thr Asn Thr Arg Tyr Leu
 210 215 220

Asp Phe Gln Arg Asp 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 Asp Gly Val Asn
 260 265 270

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

Phe Phe Asp Arg Ile Ile Ile Tyr Thr Asn Thr Val Arg Ser Thr Pro
 290 295 300

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

Gly Asn Glu Ile Arg Phe Pro Leu Tyr Gly Val Ala Ala Asn Thr Glu
 325 330 335

Pro Pro Val Thr Ile Arg Pro Thr Gly Phe Thr Asp Glu Gln Arg Gln
 340 345 350

Trp Tyr Arg Val Leu Ser Arg Ile Val Ser Phe Arg Ser Ser Gly Gln
 355 360 365

Asp Phe Ser Leu Val Asp Ala Val Gly Phe Leu Thr Ile Phe Ser Ala
 370 375 380
 Leu Ser Phe Tyr Arg Asn Gly Phe Gly Phe Asn Thr Asp Thr Ile Asn
 385 390 395 400
 Glu Ile Pro Ile Glu Gly Thr Asp Pro Leu Thr Gly Tyr Ser His Arg
 405 410 415
 Leu Cys His Val Gly Phe Val Ala Ser Ser Pro Phe Ile Ser Gln Tyr
 420 425 430
 Ala Arg Val Pro Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr
 435 440 445
 Asn Thr Ile Ala Pro Asp Val Ile Thr Gln Ile Pro Leu Val Lys Ala
 450 455 460
 Phe Asn Leu His Ser Gly Ala Thr Val Val Arg Gly Pro Gly Phe Thr
 465 470 475 480
 Gly Gly Asp Ile Leu Arg Arg Thr Asn Val Gly Ser Phe Gly Asp Met
 485 490 495
 Arg Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile
 500 505 510
 Arg Tyr Ala Ser Thr Thr Asp Leu Gln Phe His Thr Ser Ile Asn Gly
 515 520 525
 Arg Ala Ile Asn Gln Ala Asn Phe Pro Ala Thr Ile Glu Asn Arg Gly
 530 535 540
 Asn Leu Gln Phe Gly Asp Phe Arg Thr Val Gly Phe Thr Thr Pro Phe
 545 550 555 560
 Thr Phe Ser Asp Ala Asn Ser Thr Phe Thr Ile Gly Ala Trp Ser Phe
 565 570 575
 Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala
 580 585 590

Glu Val Thr Phe Glu Ala Glu Tyr Asn Leu Lys Gln Ala Gln Gln Ala
595 600 605

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

Val Thr Asp Tyr Gln Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu
625 630 635 640

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

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

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

Thr Asp Ile Ile Ile Gln Gly Gly Asp Gly Val Phe Lys Glu Asn Tyr
690 695 700

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

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

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

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

Arg Pro Leu Ser Ala Gln Asn Pro Ile Gly Lys Cys Gly Glu Pro Asn
770 775 780

Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys
785 790 795 800

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

Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val
820 825 830

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

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

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Leu
865 870 875 880

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

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

Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr
915 920 925

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

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

Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn
965 970 975

Val Lys Gly His Val Asp Val Glu Glu Gln His Asn His Arg Ser Val
980 985 990

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

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

Gly Tyr Gly Glu Gly Thr Val Thr Ile His Glu Ile Glu Asp His
1025 1030 1035

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

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

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

Thr Tyr Gly Ser Asn Ser Ser Val Ser Ala Asp Tyr Ala Ser Val
 1085 1090 1095

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

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

Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Leu
 1130 1135 1140

Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Leu Val Asp Ser
 1145 1150 1155

Val Glu Leu Leu Leu Met Glu Glu
 1160 1165

<210> 26
 <211> 722
 <212> PRT
 <213> Bacillus thuringiensis

<400> 26

Met Asn Ser Tyr Lys Asn Asn His Thr Met Val Asn Ser Pro Glu Asn
 1 5 10 15

Tyr Ser Asn Ile Val Asn Lys Tyr Pro Tyr Ala Ser Asn Pro Asn Ile
 20 25 30

Glu Met Gln Asn Met Asn Tyr Lys Asp Trp Met Ala Gly Tyr Glu Glu
 35 40 45

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

Asn Gln Val Ile Thr Leu Thr Gly Val Leu Gly Lys Thr Pro Glu Ile
 65 70 75 80

Ile Asn Ile Val Gln Glu Met Val Arg Leu Ile Arg Gly Asn Thr Gly
 85 90 95

Asn Asp Leu Leu Val His Val Glu Gln Leu Ile Gln Gln Thr Leu Ala
 100 105 110

Thr Gln Tyr Arg Ser Ala Ala Thr Gly Ala Ile Tyr Gly Ile Ser Arg
 115 120 125

Ala Tyr Asn Asp Tyr Leu Lys Phe Phe Arg Gln Trp Glu Arg Asn Arg
 130 135 140

Thr Pro Gln Asn Gly Gln Gln Val Glu Ser Ala Phe Thr Thr Val Asn
 145 150 155 160

Thr Leu Cys Ile Asn Ala Leu Ala Pro Gln Ala Ser Leu Ser Arg Arg
 165 170 175

Gly Phe Glu Thr Leu Leu Leu Pro Asn Tyr Ala Leu Ala Thr Asn Phe
 180 185 190

His Leu Leu Leu Leu Arg Asp Ala Val Leu Tyr Arg Thr Gln Trp Leu
 195 200 205

Ala Asn Ser Met Ser Thr Arg Asn Ile Asn Ile Gln Ile Leu Thr Arg
 210 215 220

Ala Ile Asp Glu Tyr Arg Asn His Cys Asn Tyr Trp Tyr Asn Ile Gly
 225 230 235 240

Leu Asn Arg Phe Thr Arg Thr Ser Phe Asn Asp Trp Val Arg Phe Asn
 245 250 255

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

Phe Pro Thr Tyr Asp Pro Ile Arg Tyr Pro Arg Pro Thr Asn Val Glu
 275 280 285

Leu Thr Arg Ile Val Tyr Thr Asp Pro Ile Ser Pro Pro Arg Gly Phe
 290 295 300

Pro Arg Thr Asn Pro Pro Ser Phe Asn Gln Met Glu Asn Leu Ile Ile
 305 310 315 320

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

Tyr His Asp Pro His His Val Tyr Arg Asp Tyr Trp Ala Gly Asn Arg
 340 345 350

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

Trp Ser Thr Asn Ile Pro Met Gln Asn Ile Asp Ile Phe Arg Val Asn
 370 375 380

Leu Thr Thr His Asp Ile Asp Ser Met Gln Arg Ser Tyr Gly Gly Val
 385 390 395 400

His Arg Ser Asp Phe Ile Gly Val Asn Thr Ile Asn Asn Gln Arg Thr
 405 410 415

Thr Leu Phe Tyr His Gln Asn Val Asn Thr Ser Leu Phe Leu Thr Arg
 420 425 430

Asn Glu Thr Val Phe Leu Pro Gly Asp Ser Gly Leu Glu Pro Asn Glu
 435 440 445

Gln Asn Tyr Thr His Arg Leu Phe Gln Val Met Thr Thr Tyr Arg Thr
 450 455 460

Asn Pro Asn Ala Arg Arg Ala Ala Phe Leu His Ala Trp Thr His Arg
 465 470 475 480

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

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

Asn Tyr Gly Glu Asn Ile Met Lys Leu Asp Asn Leu Thr Ser Gly Leu
515 520 525

Ser Tyr Lys Val Thr Ala Thr Asp Ser Ala Ala Ser Asn Thr Arg Phe
530 535 540

Ile Val Arg Val Arg Tyr Ala Ser Met Asp Asn Asn Lys Leu Asn Leu
545 550 555 560

Val Leu Asn Gly Ala Gln Ile Ala Ser Leu Asn Val Glu Arg Thr Val
565 570 575

Gln Asn Gly Gly Glu Ser Leu Thr Ala Leu Gln Cys Glu Asp Phe Lys
580 585 590

Tyr Ala Thr Phe Ala Gly Asp Phe Gln Met Gly Ser Gln Ser Ile Leu
595 600 605

Gly Ile Phe Lys Asp Ile Ser Asn Ala Asp Phe Ile Leu Asp Lys Ile
610 615 620

Glu Leu Ile Pro Ser His Phe Met Ser Ser Leu Glu Gln Thr Gln Asp
625 630 635 640

Asp Tyr Ser Tyr Asn Gln Asn Thr Ile Tyr Thr Cys Asn Gln Gly Tyr
645 650 655

Gly Thr Tyr Asp His Asn Ser Ser Asn Met Tyr Asp His Gln Asn Tyr
660 665 670

Lys Asn Tyr Thr Gln Asp Met Asp Thr Thr Tyr Gln Pro Asp Tyr Asp
675 680 685

Asn Tyr Asn Gln Asn Asn Thr Asp Met Tyr Asp Ser Gly Tyr Asn Asn
690 695 700

Ser Gln Asn Thr Gly Cys Thr Cys Asn Gln Gly Tyr Asn Asn Asn Phe
705 710 715 720

Pro Lys

<210> 27
 <211> 36
 <212> PRT
 <213> Bacillus thuringiensis

<400> 27

Ser Val Tyr His Pro Asp Ala Asn Asp Ala Tyr Asp Glu Gly Asp Asp
 1 5 10 15

Asn Thr Tyr Asn Gln Asn Tyr Asn Ser Met Tyr Asn Gln Asn Ala Asp
 20 25 30

Asn Thr Tyr Asp
 35

<210> 28
 <211> 17
 <212> PRT
 <213> Bacillus thuringiensis

<400> 28

Gly Tyr Glu Gln Asn Asn Tyr Asp Ser Asp Asp Ser Asp Tyr Asn Asn
 1 5 10 15

Thr