



(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)



1-0037357

(51)^{2020.01} C12N 15/82; A01N 63/10; C07K
14/325; A01N 63/00; A01N 63/23

(13) B

(21) 1-2017-02648

(22) 03/12/2015

(86) PCT/US2015/063620 03/12/2015

(87) WO 2016/094165 16/06/2016

(30) 62/090,900 12/12/2014 US

(45) 27/11/2023 428

(43) 27/11/2017 356A

(73) SYNGENTA PARTICIPATIONS AG (CH)

Schwarzwaldallee 215, CH-4058 Basel, Switzerland

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

(74) Công ty TNHH Ban Ca (BANCA)

(54) GEN KHẢM CHỨA VÙNG KHỞI ĐỘNG KHÁC LOẠI LIÊN KẾT HOẠT ĐỘNG
VỚI PHÂN TỬ AXIT NUCLEIC CHỨA TRÌNH TỰ NUCLEOTIT MÃ HÓA
PROTEIN DIỆT CÔN TRÙNG, CHẾ PHẨM VÀ PHƯƠNG PHÁP KIỂM SOÁT
CÔN TRÙNG

(57) Sáng chế đề cập đến gen khảm chứa vùng khởi động khác loại liên kết hoạt động với phân tử axit nucleic chứa trình tự nucleotit mã hóa protein độc đối với côn trùng. Sáng chế bộc lộ protein diệt côn trùng được phân lập từ *Bacillus thuringiensis* mà có hoạt tính kháng loài gây hại côn trùng bộ cánh vảy. ADN mã hóa protein diệt côn trùng có thể được dùng để biến nạp các sinh vật nhân sơ và sinh vật nhân thực khác nhau để biểu hiện protein diệt côn trùng. Các sinh vật tái tổ hợp này có thể được dùng để kiểm soát côn trùng bộ cánh vảy trong các môi trường khác nhau.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến protein diệt loài gây hại và phân tử axit nucleic mã hóa chúng, cũng như chế phẩm 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 trồng tạo thành bào tử gam dương khác biệt được đặc trưng bởi khả năng tạo ra thể vùi tinh thể của nó đặc biệt độc đối với một số bộ và loài nhất định của loài gây hại cây trồng, bao gồm côn trùng, nhưng vô hại đối với cây và các sinh vật không phải đích khác. Vì lý do này, chế phẩm chứa chủng *Bacillus thuringiensis* hoặc protein diệt côn trùng của chúng có thể được dùng làm chất diệt côn trùng chấp nhận được cho môi trường để kiểm soát các loài gây hại côn trùng nông nghiệp hoặc vectơ côn trùng của nhiều bệnh ở người hoặc động vật khác nhau.

Các protein Crystal (Cry) từ *Bacillus thuringiensis* có hoạt tính diệt côn trùng hiệu nghiệm kháng chủ yếu là ấu trùng bộ cánh vảy, bộ hai cánh, và bộ cánh cứng. Các protein này còn thể hiện hoạt tính kháng các loài gây hại trong các bộ gây hại cánh màng, cánh đều, chấy, ăn lông, và ve bét, cũng như các bộ không xương sống khác như giun tròn, sán dẹp, và trùng roi (Feitelson, J. 1993. The *Bacillus Thuringiensis* family tree. In *Advanced Engineered Pesticides*. Marcel Dekker, Inc., New York, N.Y.). Các protein này ban đầu được phân loại là CryI đến CryVI chủ yếu dựa trên hoạt tính diệt côn trùng của chúng. Các lớp chính là đặc hiệu bộ cánh vảy (I), đặc hiệu bộ cánh vảy và bộ hai cánh (II), đặc hiệu bộ cánh cứng (III), đặc hiệu bộ hai cánh (IV), và đặc hiệu giun tròn (V) và (VI). Các protein còn được phân loại thành các phân họ; các protein liên quan cao hơn trong mỗi họ được đặt các chữ cái phân chia như CryIA, CryIB, CryIC, v.v. Các protein liên quan gần hơn nữa trong mỗi nhóm được đặt tên như CryIC(a), CryIC(b), v.v. Các thuật ngữ “độc tố Cry” và “nội độc tố denta” được sử dụng hoán đổi với thuật ngữ “protein Cry.” Danh pháp hiện nay cho

protein Cry và gen là dựa trên sự tương đồng trình tự axit amin hơn là đặc tính nhắm 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.

Các protein Cry là các phân tử protein hình cầu mà tích tụ thành các tiền độc tố ở dạng tinh thể trong suốt giai đoạn hình thành bào tử của Bt. Sau khi loài gây hại ăn vào bụng, các tinh thể thường được hòa tan và giải phóng tiền độc tố, mà có thể có kích thước nằm trong khoảng, ví dụ, từ 130-140kDa đối với các protein Cry có hoạt tính với bộ cánh vảy và 60-80kDa đối với các protein Cry có hoạt tính với bộ cánh cứng. Các tiền độc tố được chuyển thành các đoạn độc tố chín muối (khoảng 60-70kDa vùng đầu N) bởi các proteaza trong ruột trong loài gây hại đích. Nhiều trong số các protein này là khá độc đối với côn trùng đích cụ thể, nhưng vô hại đối với cây trồng và các sinh vật không phải đích khác.

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

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

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

Do đó, có thể đạt được kiểm soát côn trùng tốt, nhưng một số chất hóa học nhất định đôi khi có thể còn ảnh hưởng đến côn trùng có ích không phải đích và một số chất sinh học nhất định có phổ hoạt tính rất hẹp. Ngoài ra, việc sử dụng tiếp một số phương pháp kiểm soát hóa học và sinh học nhất định làm tăng khả năng cho loài gây hại côn trùng phát triển khả năng kháng các phương pháp kiểm soát này. Điều này đã được giảm bớt một phần bằng nhiều thực tế quản lý sức kháng khác nhau, nhưng vẫn có nhu cầu đối với việc phát hiện các tác nhân kiểm soát loài gây hại mới và hiệu quả mà tạo ra lợi ích kinh tế cho người nông dân và chấp nhận được cho môi trường. Cụ thể là có nhu cầu đối với các tác nhân kiểm soát mà nhắm đến phổ rộng hơn của các loài gây hại côn trùng quan trọng về kinh tế và kiểm soát hiệu quả các loài côn trùng mà có thể trở nên kháng các tác nhân kiểm soát côn trùng hiện có.

Bản chất kỹ thuật của sáng chế

Trên cơ sở xem xét các nhu cầu này, mục đích của sáng chế là đề xuất các tác nhân kiểm soát loài gây hại mới bằng cách đề xuất các gen mới và protein diệt loài gây hại mà có thể dùng để kiểm soát nhiều loài gây hại cây trồng khác nhau.

Sáng chế đề xuất các chế phẩm và phương pháp tạo ra hoạt tính diệt loài gây hại lên vi khuẩn, cây, tế bào, mô và hạt của cây. Cụ thể là, sáng chế đề xuất gen khảm bao gồm các polynucleotit mới mà mã hóa protein Cry được phân lập từ *Bacillus thuringiensis* (Bt) và các trình tự cơ bản tương đồng với chúng, mà biểu hiện của chúng tạo ra trong protein độc tố đối với các loài gây hại côn trùng quan trọng về kinh tế, cụ thể là loài gây hại côn trùng mà phá hoại cây trồng. Sáng chế còn đề xuất protein Cry mới tạo ra từ biểu hiện của trình tự axit nucleic, và các chế phẩm và chế phẩm phối chế chứa protein Cry, mà độc với côn trùng bằng cách ức chế khả năng loài gây

hại côn trùng sống sót, phát triển và sinh sản, hoặc hạn chế thiệt hại hoặc thất thoát cho cây trồng liên quan đến côn trùng. Các protein Cry theo sáng chế bao gồm protein Cry tự nhiên và protein Cry đột biến mà có một hoặc nhiều việc thế, thêm hoặc loại axit amin. Các ví dụ của protein Cry đột biến bao gồm nhưng không giới hạn các loại mà được đột biến để có phổ hoạt tính rộng hơn bản sao protein Cry tự nhiên của chúng hoặc loại được đột biến để đưa vào một epitop để tạo ra các kháng thể mà nhận diện một cách khác protein được đột biến từ protein tự nhiên. Các protein Cry mới theo sáng chế có hoạt tính cao kháng loài gây hại côn trùng. Ví dụ, protein Cry theo sáng chế có thể được dùng để kiểm soát một hoặc nhiều loài gây hại côn trùng quan trọng về kinh tế như sâu xám (*Agrotisipsilon*), sâu đục thân ngô Châu Âu (*Ostrinianubilalis*), sâu xanh mùa thu (*Spodopterafrugiperda*), sâu đục trái trên cây ngô (*Helicoverpazea*), sâu đục thân mía (*Diatraeasaccharalis*), sâu ăn lá đậu tím (*Anticarsia gemmatalis*), sâu đo trên đậu tương (*Chrysodeixisincludes*), sâu đục thân ngô tây nam (*Diatraeagrandiosella*), sâu đục trái đậu phương Tây (*Richiaalbicosta*), sâu chồi thuốc lá (*Heliothisvirescens*), sâu đục thân ngô châu Á (*Ostriniafurancalis*), sâu đục quả bông (*Helicoverpaarmigera*), sâu đục thân sọc (*Chilosuppressalis*), sâu cú mèo (*Sesamiacalamistis*), sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocismedinalis*), và tương tự.

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

Sáng chế còn đề xuất caset biểu hiện và vector tái tổ hợp chứa polynucleotit mà mã hóa protein Cry theo sáng chế. Sáng chế còn đề xuất vi khuẩn, cây, tế bào, mô, và hạt của cây được biến nạp chứa gen khảm, hoặc caset biểu hiện hoặc vector tái tổ hợp mà chứa polynucleotit mã hóa 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 caset biểu hiện hoặc vector tái tổ hợp để biến nạp và biểu hiện trong sinh vật, bao gồm vi sinh vật và cây trồng. Nucleotit hoặc trình tự axit amin có thể là trình tự tổng hợp mà được thiết kế để biểu hiện trong sinh vật bao gồm, nhưng không giới hạn ở, vi sinh vật hoặc cây trồng hoặc tạo ra độc tố lai với hoạt tính diệt loài gây hại cải thiện. Sáng chế còn đề xuất phương pháp tạo ra protein Cry và đến phương pháp sử dụng trình tự axit nucleic, ví dụ trong vi sinh vật để kiểm soát côn

trùng hoặc trong cây chuyển gen để tạo ra sự bảo vệ khỏi thiệt hại do côn trùng, và phương pháp sử dụng protein Cry, và các chế phẩm và chế phẩm phối chế bao gồm protein Cry, ví dụ áp dụng protein Cry hoặc các chế phẩm hoặc chế phẩm phối chế lên vùng bị côn trùng phá hoại, hoặc để xử lý phòng ngừa vùng dễ bị côn trùng tấn công hoặc cây trồng để tạo ra sự bảo vệ chống lại loài gây hại côn trùng. Nucleotit hoặc trình tự axit amin có thể là trình tự tổng hợp mà được thiết kế để biểu hiện trong sinh vật bao gồm, nhưng không giới hạn ở, vi sinh vật hoặc cây trồng.

Các chế phẩm và phương pháp theo sáng chế hữu dụng để tạo ra sinh vật mà độc đối với côn trùng, đặc biệt là vi khuẩn và cây trồng. Các sinh vật này và các chế phẩm có nguồn gốc từ chúng được mong muốn cho các mục đích nông nghiệp. Các chế phẩm theo sáng chế còn hữu dụng để tạo ra protein Cry thay thế hoặc 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 hoặc vi sinh vật.

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

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

SEQ ID NO: 1 thể hiện trình tự nucleotit mã hóa protein BT-0009.

SEQ ID NO: 2 thể hiện trình tự nucleotit mã hóa protein BT-0012.

SEQ ID NO: 3 thể hiện trình tự nucleotit mã hóa protein BT-0013.

SEQ ID NO: 4 thể hiện trình tự nucleotit mã hóa protein BT-0023.

SEQ ID NO: 5 thể hiện trình tự nucleotit mã hóa protein BT-0067.

SEQ ID NO: 6 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0009.

SEQ ID NO: 7 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0012.

SEQ ID NO: 8 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0013.

SEQ ID NO: 9 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0023.

SEQ ID NO: 10 thể hiện trình tự tối ưu hóa bộ ba mã mã hóa protein BT-0067.

SEQ ID NO: 11 thể hiện trình tự nucleotit mã hóa protein BT-0009 đột biến.

SEQ ID NO: 12 thể hiện trình tự nucleotit mã hóa protein BT-0012 đột biến.

SEQ ID NO: 13 thể hiện trình tự nucleotit mã hóa protein BT-0013 đột biến.

SEQ ID NO: 14 thể hiện trình tự nucleotit mã hóa protein BT-0023 đột biến.

SEQ ID NO: 15 thể hiện trình tự nucleotit mã hóa protein BT-0067 đột biến.

SEQ ID NO: 16 thể hiện trình tự axit amin của protein BT-0009.

SEQ ID NO:17 thể hiện trình tự axit amin của protein BT-0012.
 SEQ ID NO:18 thể hiện trình tự axit amin của protein BT-0013.
 SEQ ID NO:19 thể hiện trình tự axit amin của protein BT-0023.
 SEQ ID NO:20 thể hiện trình tự axit amin của protein BT-0067.
 SEQ ID NO:21 thể hiện trình tự axit amin của protein BT-0009 đột biến.
 SEQ ID NO:22 thể hiện trình tự axit amin của protein BT-0012 đột biến.
 SEQ ID NO:23 thể hiện trình tự axit amin của protein BT-0013 đột biến.
 SEQ ID NO:24 thể hiện trình tự axit amin của protein BT-0023 đột biến.
 SEQ ID NO:25 thể hiện trình tự axit amin của protein BT-0067 đột biến.
 SEQ NOS:26-31 thể hiện đoạn môi hữu dụng trong sáng chế.

Mô tả chi tiết sáng chế

Phần mô tả này không nhằm làm danh mục cụ thể của tất cả cách khác nhau trong đó sáng chế có thể được thực hiện, hoặc tất cả các dấu hiệu có thể được thêm vào sáng chế. Ví dụ, các dấu hiệu được minh họa theo một phương án có thể được kết hợp với các phương án khác, và các dấu hiệu được minh họa theo phương án cụ thể có thể được xóa bỏ khỏi phương án này. Do đó, sáng chế dự tính rằng theo một số phương án theo sáng chế, dấu hiệu bất kỳ hoặc kết hợp các dấu hiệu được đề cập ở đây có thể được loại bỏ hoặc bỏ qua. Ngoài ra, nhiều biến thể và bổ sung cho các phương án khác nhau được khuyến khích ở đây sẽ được thực hiện bởi người có hiểu biết trung bình trong lĩnh vực theo sáng chế, mà không vượt ra khỏi phạm vi sáng chế. Vì vậy, phần mô tả sau đây nhằm minh họa một số phương án cụ thể của sáng chế, và không xác định 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 đối mà sáng chế thuộc về. Thuật ngữ được sử dụng trong phần mô tả của sáng chế ở đây là nhằm mục đích mô tả chỉ các phương án cụ thể và không nhằm giới hạn sáng chế. Còn được hiểu rằng thuật ngữ được sử dụng ở đây là nhằm mục đích chỉ mô tả các phương án cụ thể, và không nhằm giới hạn phạm vi của sáng chế.

Định nghĩa

Như được sử dụng ở đây và ở yêu cầu bảo hộ kèm theo, các dạng mạo từ số ít bao gồm viện dẫn số nhiều trừ khi ngữ cảnh quy định rõ ràng khác. Do đó, ví dụ, viện dẫn đến "cây trồng" là viện dẫn đến một hoặc nhiều cây trồng và bao gồm các tương đương của chúng đã biết đối với người có hiểu biết trung bình trong lĩnh vực, và v.v. Như được sử dụng ở đây, từ "hoặc" nghĩa là một thành phần bất kỳ của danh sách cụ thể và còn bao gồm kết hợp bất kỳ của các thành phần của danh sách đó (tức là, còn bao gồm "và").

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

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

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

Microbiology, Washington, D.C. (1993). Sản phẩm của khuếch đại được gọi là "amplicon."

Thuật ngữ "cấu trúc khảm" hoặc "gen khảm" hoặc "polynucleotit khảm" hoặc "axit nucleic khảm" (hoặc thuật ngữ tương tự) như được sử dụng ở đây đề cập đến cấu trúc hoặc phân tử bao gồm hai hoặc nhiều polynucleotit có nguồn gốc khác nhau được ghép thành phân tử axit nucleic đơn. Thuật ngữ "cấu trúc khảm", "gen khảm", "polynucleotit khảm" hoặc "axit nucleic khảm" đề cập đến bất kỳ cấu trúc hoặc phân tử mà bao gồm, không giới hạn, (1) polynucleotit (ví dụ, ADN), chứa polynucleotit điều khiển và mã hóa mà không thấy cùng nhau trong tự nhiên (tức là, ít nhất một trong số các polynucleotit trong cấu trúc là khác loại với ít nhất một trong số các polynucleotit khác của nó), hoặc (2) các phần mã hóa polynucleotit của protein không tiếp giáp trong tự nhiên, hoặc (3) các phần của vùng khởi động mà không tiếp giáp trong tự nhiên. Hơn nữa, cấu trúc khảm, gen khảm, polynucleotit khảm hoặc axit nucleic khảm có thể chứa polynucleotit điều khiển và polynucleotit mã hóa mà có nguồn gốc từ các nguồn khác nhau, hoặc chứa polynucleotit điều khiển 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 caset biểu hiện chứa polynucleotit theo sáng chế dưới sự kiểm soát của polynucleotit điều khiển, cụ thể là dưới sự kiểm soát của polynucleotit điều khiển chức năng trong cây trồng hoặc vi khuẩn.

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

Như được sử dụng ở đây, trình tự "tối ưu hóa bộ ba mã" nghĩa là trình tự nucleotit của polynucleotit tái tổ hợp, chuyển gen, hoặc tổng hợp trong đó bộ ba mã được chọn để phản ánh độ lệch bộ ba mã cụ thể mà tế bào chủ hoặc sinh vật có thể có. Điều này thường được thực hiện theo cách này để bảo quản trình tự axit amin của polypeptit được mã hóa bởi trình tự nucleotit tối ưu hóa bộ ba mã. Theo một số phương án nhất định, trình tự ADN của cấu trúc ADN tái tổ hợp bao gồm trình tự mà đã là bộ ba mã được tối ưu hóa cho tế bào (ví dụ, tế bào động vật, thực vật, hoặc nấm) trong đó cấu trúc là để được biểu hiện. Ví dụ, cấu trúc được biểu hiện trong tế bào của

cây có thể có tất cả hoặc các phần của trình tự của nó (ví dụ, yếu tố triệt gen thứ nhất hoặc yếu tố biểu hiện gen) bộ ba mã được tối ưu hóa để biểu hiện ở cây trồng. Xem, ví dụ, tài liệu bằng sáng chế Mỹ số 6,121,014, được kết hợp ở đây bằng cách viện dẫn.

"Kiểm soát" côn trùng nghĩa là ức chế, thông qua tác dụng độc, khả năng của loài gây hại côn trùng để sống, sinh trưởng, ăn, và/hoặc sinh sản, hoặc hạn chế thiệt hại hoặc mất mát liên quan đến côn trùng ở cây trồng hoặc bảo vệ sản lượng tiềm năng của cây trồng khi phát triển với sự có mặt của loài gây hại côn trùng. "Kiểm soát" côn trùng có thể có hoặc có thể không có nghĩa là giết côn trùng, mặc dù tốt hơn nên có nghĩa là giết côn trùng.

Các thuật ngữ "gồm có" và/hoặc "bao gồm," khi được sử dụng trong bản mô tả này, chỉ rõ sự có mặt của các dấu hiệu, toàn bộ, bước, hoạt động, yếu tố, và/hoặc thành phần được cho, nhưng không loại trừ sự có mặt hoặc thêm vào một hoặc nhiều dấu hiệu, toàn bộ, bước, hoạt động, yếu tố, thành phần, và/hoặc nhóm khác của chúng.

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

Trong ngữ cảnh của sáng chế, "ứng với" hoặc "tương ứng với" nghĩa là khi trình tự axit amin của protein Cry biến đổi được sắp thẳng hàng với nhau, các axit amin mà "tương ứng với" các vị trí được đánh số nhất định trong protein biến đổi hoặc tương đồng là những cái mà thẳng hàng với các vị trí này trong protein so sánh nhưng không cần thiết trong các vị trí số chính xác này so với trình tự axit amin so sánh cụ thể theo sáng chế. Ví dụ, nếu SEQ ID NO:16 là trình tự quy chiếu và được sắp thẳng hàng với SEQ ID NO:17, Ala150 nêu ở SEQ ID NO:17 "tương ứng với" Ala163 nêu ở SEQ ID NO:16.

"Phân phối" chế phẩm hoặc protein độc nghĩa là chế phẩm hoặc protein độc này tiếp xúc với côn trùng, tạo ra tác dụng độc và kiểm soát côn trùng. Chế phẩm hoặc protein độc được phân phối theo nhiều cách đã biết, ví dụ, qua đường miệng do được ăn vào bụng bởi côn trùng hoặc bằng cách tiếp xúc với côn trùng thông qua biểu hiện

của cây chuyển gen, (các) chế phẩm protein phối chế, (các) chế phẩm protein phun được, nền chất bả, hoặc bất kỳ hệ phân phối protein đã biết khác trong lĩnh vực.

Thuật ngữ "vùng" đề cập đến bộ của các axit amin được bảo tồn ở các vị trí cụ thể dọc theo sự sắp hàng của các trình tự của các protein liên quan đến tiến hóa. Trong khi các axit amin ở các vị trí khác có thể thay đổi giữa các tương đồng, các axit amin được bảo tồn cao ở các vị trí cụ thể cho thấy các axit amin mà có vẻ thiết yếu trong cấu trúc, sự ổn định hoặc chức năng của protein. Được xác định bằng mức bảo tồn cao của chúng trong các trình tự được sắp hàng của họ các tương đồng protein, chúng có thể được sử dụng làm phân định dạng để xác định nếu bất kỳ polypeptit cần biết thuộc họ polypeptit được xác định trước đó.

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

"Caset biểu hiện" như được sử dụng ở đây nghĩa là phân tử axit nucleic có khả năng điều khiển biểu hiện của ít nhất một polynucleotit được mong muốn trong tế bào chủ thích hợp, bao gồm vùng khởi động liên kết hoạt động với polynucleotit được mong muốn mà liên kết hoạt động với tín hiệu kết thúc. "Caset biểu hiện" thường còn chứa polynucleotit bổ sung cần để phiên mã đúng polynucleotit được mong muốn. Caset biểu hiện có thể còn chứa polynucleotit khác không cần trong biểu hiện trực tiếp của polynucleotit được mong muốn nhưng có mặt do vị trí giới hạn thích hợp để loại bỏ tổ hợp khởi vectơ biểu hiện. Caset biểu hiện bao gồm (các) polynucleotit được mong muốn có thể là khảm, có nghĩa là ít nhất một trong số các thành phần của nó là khác loại với ít nhất một trong số các thành phần khác của nó. Caset biểu hiện có thể còn là tổ hợp mà xuất hiện trong tự nhiên nhưng có thể thu được ở dạng tái tổ hợp hữu dụng để biểu hiện khác loài. Tuy nhiên, thông thường, caset biểu hiện là khác loại với vật chủ, tức là polynucleotit được mong muốn trong caset biểu hiện không xảy ra tự nhiên trong tế bào chủ và phải được đưa vào tế bào chủ hoặc tổ tiên của tế bào chủ bằng quá trình biến nạp hoặc quá trình gây giống. Biểu hiện của (các) polynucleotit được mong muốn trong caset biểu hiện thường là dưới sự kiểm soát của vùng khởi

động. Trong trường hợp sinh vật nhiều ngăn, như cây trồng, vùng khởi động có thể còn là đặc biệt hoặc ưu tiên cho mô, hoặc cơ quan, hoặc giai đoạn phát triển cụ thể. Caset biểu hiện, hoặc đoạn của chúng, còn có thể được gọi là "polynucleotit được chèn" hoặc "polynucleotit chèn vào" khi được biến nạp vào cây trồng.

"Gen" được xác định ở đây là đơn vị di truyền gồm có polynucleotit mà chiếm địa điểm cụ thể trên nhiễm sắc thể hoặc plasmit và mà chứa chỉ dấu di truyền cho các đặc trưng hoặc đặc điểm cụ thể trong sinh vật.

"Proteaza của ruột" là proteaza được tìm thấy tự nhiên trong đường tiêu hóa của côn trùng. Proteaza này thường tham gia vào việc tiêu hóa protein được ăn vào.

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

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

Trình tự axit nucleic "đăng mã hóa" với trình tự axit nucleic quy chiếu khi trình tự axit nucleic mã hóa polypeptit có trình tự axit amin giống như polypeptit được mã hóa bởi trình tự axit nucleic quy chiếu.

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

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

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

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

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

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

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

"Vật liệu cây" đề cập đến lá, thân, rễ, hoa hoặc các bộ phận của hoa, quả, phần hoa, tế bào trứng, hợp tử, hạt, cành giâm, mầm 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 trông như rễ, thân, lá, nụ hoa, hoặc phôi.

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

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

"Polynucleotit được mong muốn" đề cập đến bất kỳ polynucleotit mà, khi được chuyển vào sinh vật, ví dụ cây trồng, tạo ra cho sinh vật đặc trưng mong muốn như khả năng kháng kháng sinh, khả năng kháng virus, khả năng kháng côn trùng, khả năng kháng bệnh, hoặc khả năng kháng các loài gây hại khác, khả năng chịu thuốc diệt cỏ, cải thiện giá trị dinh dưỡng, cải thiện năng suất trong xử lý công nghiệp, tạo ra các enzym hoặc các chất chuyển hóa có giá trị thương mại hoặc thay đổi khả năng sinh sản.

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

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

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

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

Thuật ngữ "tương đồng" hoặc "cơ bản tương đồng" trong ngữ cảnh của hai trình tự axit nucleic hoặc protein, đề cập đến hai hoặc nhiều trình tự hoặc các trình tự phụ mà có ít nhất 60%, tốt hơn là 80%, tốt hơn nữa là 90%, còn tốt hơn nữa là 95%, và tốt nhất là ít nhất 99% tương đồng nucleotit hoặc gốc axit amin, khi so sánh và sắp thẳng hàng cho tương ứng tối đa, như được đo bằng cách sử dụng một trong các thuật toán so sánh trình tự sau hoặc bằng cách kiểm tra bằng mắt. Tốt hơn là, cơ bản tương đồng tồn tại trên vùng của các trình tự mà có chiều dài ít nhất khoảng 50 gốc, tốt hơn nữa là trên vùng của ít nhất khoảng 100 gốc, và tốt nhất là các trình tự là cơ bản tương đồng trên ít nhất khoảng 150 gốc. Theo phương án được ưu tiên đặc biệt, các trình tự là cơ

bản tương đồng trên toàn bộ chiều dài của vùng mã hóa. Ngoài ra, các trình tự axit nucleic hoặc protein cơ bản tương đồng cơ bản thực hiện cùng chức năng.

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

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

Một ví dụ của thuật toán mà thích hợp để xác định phần trăm tương đồng trình tự và tương tự 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 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 trình tự axit amin, chương trình BLASTP sử dụng mặc định độ dài từ (W) là 3, mức mong đợi (E) là 10, và ma trận tính điểm BLOSUM62 (tham khảo Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 (1989)).

Ngoài việc tính phần trăm tương đồng trình tự, thuật toán BLAST còn thực hiện phân tích thống kê sự 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 quy chiếu nếu tổng xác suất nhỏ nhất khi so sánh trình tự axit nucleic thử nghiệm với trình tự axit nucleic quy chiếu nhỏ hơn khoảng 0,1, tốt hơn nữa là nhỏ hơn khoảng 0,01, và tốt nhất là nhỏ hơn khoảng 0,001.

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

"Các điều kiện lai nghiêm ngặt" và "các điều kiện rửa lai nghiêm ngặt" trong ngữ cảnh các thử nghiệm lai axit nucleic như lai Southern và Northern là phụ thuộc trình tự, và khác nhau theo các thông số môi trường khác nhau. Các trình tự càng dài lai đặc biệt ở các nhiệt độ càng cao. Hướng dẫn khái quát về lai các axit nucleic được tìm thấy trong Kỹ thuật thí nghiệm Tijssen (1993) trong lai hóa sinh và sinh học phân

tử với trình tự dò axit nucleic phần I chương 2 "Overview of principles of hybridization and the strategy of axit nucleic probe assays" Elsevier, New York. Thông thường, các điều kiện lai và rửa nghiêm ngặt cao được chọn là thấp hơn khoảng 5°C so với điểm nóng chảy nhiệt (T_m) cho trình tự cụ thể ở cường độ ion và pH xác định. Thông thường, dưới "các điều kiện nghiêm ngặt" trình tự dò sẽ lai với trình tự phụ đích, mà không lai với các trình tự khác.

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

Sau đây là các ví dụ của bộ các điều kiện lai/rửa mà có thể được sử dụng để tách dòng trình tự nucleotit cùng loại mà cơ bản tương đồng với trình tự nucleotit quy chiếu theo sáng chế: trình tự nucleotit quy chiếu tốt hơn là lai với trình tự nucleotit quy

chiều trong 7% natri đodexyl 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 đodexyl 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 đodexyl 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 đodexyl 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 đodexyl 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" đề cập đến trình tự nucleotit bao gồm các đặc điểm cơ bản và/hoặc cấu trúc mà có mặt trong trình tự tự nhiên. Ví dụ, trình tự nhân tạo mã hóa protein Cry theo sáng chế mà giống gần hơn nội dung G+C và phân phối bộ ba mã thông thường của gen thực vật hai lá mầm và/hoặc một lá mầm được gọi là tổng hợp.

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

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

Nucleotit được biểu thị bằng các bazơ của chúng bằng các chữ viết tắt tiêu chuẩn sau: adenin (A), xytosin (C), thymin (T), và guanin (G). Tương tự các axit amin được biểu thị bằng các chữ viết tắt tiêu chuẩn sau: alanin (Ala; A), arginin (Arg; R),

asparagin (Asn; N), axit aspartic (Asp; D), xystein (Cys; C), glutamin (Gln; Q), axit glutamic (Glu; E), glyxin (Gly; G), histidin (His; H), 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 chế phẩm và phương pháp để kiểm soát loài gây hại cây trồng. Cụ thể là, sáng chế đề cập đến protein Cry mà độc đối với loài gây hại cây trồng và polynucleotit mà bao gồm trình tự nucleotit mà mã hóa protein Cry, và việc tạo ra và sử dụng polynucleotit và protein Cry để kiểm soát loài gây hại cây trồng.

Theo đó, theo một số phương án, gen khảm được đề xuất mà bao gồm vùng khởi động khác loại liên kết hoạt động với polynucleotit chứa trình tự nucleotit mà mã hóa độc tố protein đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự nucleotit (a) có ít nhất 80% (ví dụ 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) đến ít nhất 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOs:1-5; hoặc (b) mã hóa protein chứa trình tự axit amin mà có ít nhất 80% (ví dụ 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) với ít nhất 99% (99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%) tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOs: 16-20; hoặc (c) là trình tự tổng hợp của (a) hoặc (b) mà có bộ ba mã được tối ưu hóa để biểu hiện trong sinh vật chuyển gen.

Theo các phương án khác, vùng khởi động khác loại là vùng khởi động có thể biểu hiện thực vật. Ví dụ, không giới hạn, vùng khởi động có thể biểu hiện thực vật có thể được chọn từ nhóm bao gồm vùng khởi động ubiquitin, cmp, TrpA của ngô, 9 5' UTR gen T3 của vật ăn vi khuẩn, synthetaza 1 của đường ngô, dehydrogenaza 1 của rượu ngô, phức hệ thu sáng của ngô, protein sốc nhiệt của ngô, RuBP carboxylaza khối phụ nhỏ của đậu Hà Lan, Ti plasmit mannopin synthaza, Ti plasmit nopalin synthaza, chalcon isomeraza của dã yên thảo, protein 1 giàu glyxerin của đậu, patatin của khoai tây, lectin, CaMV 35S, và RuBP carboxylaza khối phụ nhỏ S-E9.

Theo các phương án bổ sung, protein được mã hóa bởi gen khảm là độc tố bổ sung cho một hoặc nhiều loài côn trùng được chọn từ nhóm bao gồm sâu xám (*Agrotis*

ipson), sâu xanh mùa thu (*Spodopterafrugiperda*), sâu đục trái trên cây ngô (*Helicoverpazea*), sâu đục thân mía (*Diatraeasaccharalis*), sâu ăn lá đậu tím (*Anticarsia gemmatalis*), sâu đo trên đậu tương (*Chrysodeixisincludes*), sâu đục thân ngô tây nam (*Diatraeagrandiosella*), sâu đục trái đậu phương Tây (*Richiaalbicosta*), sâu chồi thuốc lá (*Heliothisvirescens*), sâu đục thân ngô châu Á (*Ostriniafurancalis*), sâu đục quả bông (*Helicoverpaarmigera*), sâu đục thân sọc (*Chilosuppressalis*), sâu cú mèo (*Sesamiacalamistis*) và sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocismedinalis*).

Theo các phương án khác nữa, polynucleotit bao gồm trình tự nucleotit mà có ít nhất 80% đến ít nhất 99% tương đồng trình tự với SEQ ID NO:1, hoặc có ít nhất 80% đến ít nhất 99% tương đồng trình tự với SEQ ID NO:2, hoặc có ít nhất 80% đến ít nhất 99% tương đồng trình tự với SEQ ID NO:3, hoặc có ít nhất 80% đến ít nhất 99% tương đồng trình tự với SEQ ID NO:4, hoặc có ít nhất 80% đến ít nhất 99% tương đồng trình tự với SEQ ID NO:5.

Theo các phương án khác, polynucleotit bao gồm trình tự nucleotit mà mã hóa protein chứa trình tự axit amin mà có ít nhất 80% đến ít nhất 99% tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOS: 16-20.

Theo các phương án khác nữa, trình tự axit amin có ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với SEQ ID NO:16.

Theo các phương án khác nữa, trình tự axit amin có 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.

Theo các phương án khác nữa, trình tự axit amin có 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:18.

Theo các phương án khác, trình tự axit amin có ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất

96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với SEQ ID NO:19.

Theo các phương án khác nữa, trình tự axit amin 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 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với SEQ ID NO:20.

Theo một số phương án, gen khảm theo sáng chế bao gồm polynucleotit bao gồm trình tự tổng hợp của trình tự nucleotit mà có ít nhất 80%, hoặc ít nhất 81%, hoặc ít nhất 82%, hoặc ít nhất 83%, hoặc ít nhất 84%, hoặc ít nhất 85%, hoặc ít nhất 86%, hoặc ít nhất 87%, hoặc ít nhất 88%, hoặc ít nhất 89%, hoặc ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% với bất kỳ trong số các SEQ ID NOS:6-15, trong đó trình tự tổng hợp có bộ ba mã được tối ưu hóa để biểu hiện là sinh vật chuyên gen. Theo các phương án khác, gen khảm theo sáng chế bao gồm polynucleotit bao gồm trình tự tổng hợp của trình tự nucleotit mà mã hóa protein chứa 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 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với bất kỳ trong số các SEQ ID NOS:16-25, trong đó trình tự tổng hợp có bộ

ba mã được tối ưu hóa để biểu hiện là sinh vật chuyển gen. Theo các phương án khác nữa, sinh vật chuyển gen là vi khuẩn chuyển gen hoặc cây chuyển gen.

Theo một số phương án, sáng chế đề xuất polynucleotit tổng hợp bao gồm, chủ yếu gồm có hoặc gồm có trình tự nucleotit mà mã hóa protein có hoạt tính kháng ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự nucleotit 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 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOS:6-15.

Theo các phương án khác, sáng chế đề xuất polynucleotit tổng hợp bao gồm, chủ yếu gồm có hoặc gồm có trình tự nucleotit mà mã hóa protein có hoạt tính kháng ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự nucleotit mã hóa 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 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOS:16-25.

Theo một số phương án, sáng chế đề xuất protein phân lập mà có độc tố đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó protein bao gồm, chủ yếu gồm có hoặc gồm có (a) trình tự axit amin mà có ít nhất 80% tương đồng trình tự với ít nhất 99% tương đồng trình tự với trình tự axit amin được đại diện bởi bất kỳ một trong số các SEQ ID NOS:16-25 hoặc (b) trình tự axit amin mà được mã hóa bởi trình tự nucleotit mà có ít nhất 80% tương đồng trình tự với ít nhất 99% tương đồng trình tự với trình tự nucleotit được đại diện bởi bất kỳ một trong số các SEQ ID NOS:6-15.

Theo các phương án khác, protein phân lập bao gồm, chủ yếu gồm có hoặc gồm có trình tự axit amin mà có ít nhất 80% đến ít nhất 99% tương đồng trình tự với một trình tự bất kỳ nêu ở SEQ ID NOS:16-20. Theo các phương án khác nữa, trình tự axit amin có ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với SEQ ID NO:16.

Theo các phương án khác, trình tự axit amin có 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.

Theo các phương án khác, trình tự axit amin có 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:18.

Theo các phương án khác nữa, trình tự axit amin có ít nhất 90%, hoặc ít nhất 91%, hoặc ít nhất 92%, hoặc ít nhất 94%, hoặc ít nhất 94%, hoặc ít nhất 95%, hoặc ít nhất 96%, hoặc ít nhất 97%, hoặc ít nhất 98%, hoặc ít nhất 99%, hoặc ít nhất 99,1%, hoặc ít nhất 99,2%, hoặc ít nhất 99,3%, hoặc ít nhất 99,4%, hoặc ít nhất 99,5% hoặc ít nhất 99,6%, hoặc ít nhất 99,7%, hoặc ít nhất 99,8%, hoặc ít nhất 99,9% tương đồng trình tự với SEQ ID NO:19.

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

Theo một số phương án, trình tự axit amin bao gồm, chủ yếu gồm có hoặc gồm có một trình tự bất kỳ nêu ở SEQ ID NOs:16-25.

Các kháng thể được kích thích đáp ứng với thử thách miễn dịch bằng BT-0009, BT-0012, BT-0013, BT-0023 và BT-0067 tự nhiên hoặc đột biến và tương tự hoặc các protein liên quan theo sáng chế có thể được tạo ra bằng cách sử dụng các kỹ thuật miễn dịch tiêu chuẩn để tạo ra huyết thanh miễn dịch đa dòng và, nếu muốn, làm bất tử tế bào tạo ra kháng thể của vật chủ được gây miễn dịch đối với các nguồn sản sinh ra kháng thể đơn dòng. Các kỹ thuật để tạo ra các kháng thể cho bất kỳ chất được mong muốn là đã biết, ví dụ, như trong Harlow và Lane (1988) và như trong Goding (1986). Sáng chế bao hàm protein diệt côn trùng mà phản ứng chéo với các kháng thể được kích thích chống lại một hoặc nhiều trong số các protein Cry diệt côn trùng theo sáng chế.

Các kháng thể được tạo ra theo sáng chế còn hữu dụng trong các thử nghiệm miễn dịch để xác định lượng hoặc sự có mặt của BT-0044, BT-0051, BT-0068 và BT-0128 tự nhiên hoặc đột biến hoặc protein liên quan trong mẫu sinh học. Các thử nghiệm còn hữu dụng trong việc tạo ra kiểm soát được chất lượng các chế phẩm chứa một hoặc nhiều trong số các protein độc theo sáng chế hoặc protein độc liên quan. Ngoài ra, các kháng thể có thể được dùng để đánh giá hiệu quả của việc sản xuất một hoặc nhiều protein tái tổ hợp theo sáng chế hoặc protein liên quan, cũng như để sàng lọc các thư viện biểu hiện cho sự có mặt của trình tự nucleotit mã hóa một hoặc nhiều trong số các protein theo sáng chế hoặc các trình tự mã hóa protein liên quan. Các kháng thể còn hữu dụng làm phối tử ái lực để tinh chế và/hoặc phân lập bất kỳ một hoặc nhiều trong số các protein theo sáng chế và các protein liên quan. Các protein theo sáng chế và các protein chứa epitop kháng nguyên liên quan có thể thu được bằng cách biểu hiện quá toàn bộ hoặc một phần các chiều dài của trình tự mã hóa toàn bộ hoặc một phần của protein theo sáng chế hoặc protein liên quan trong tế bào chủ được ưu tiên.

Hiểu rằng ADN các trình tự 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:13-16, 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 thể, 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 còn có thể được thực hiện bởi một trong số vài dạng gây đột biến và/hoặc trong tiến triển được dẫn 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 phương án của sáng chế, trình tự nucleotit được đại diện bởi SEQ ID NOs: 1-5 được biến đổi để đưa vào sự thể axit amin trong protein được mã hóa. Theo một số phương án, protein đột biến thu được được mã hóa bởi polynucleotit đột biến tổng hợp bao gồm trình tự nucleotit được đại diện bởi bất kỳ một trong số các SEQ ID NOs: 11-15. Theo các phương án khác, protein đột biến bao gồm, chủ yếu gồm có hoặc gồm có trình tự axit amin được đại diện bởi bất kỳ một trong số các SEQ ID NOs: 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 chế phẩm theo sáng chế. Ví dụ, có thể biểu hiện protein Cry trong các tế bào chủ mà thể hiện tỷ lệ cao kết hợp sai bazơ trong khi sao chép ADN, như XL-1 Red (Stratagene, La Jolla, CA). Sau khi nhân giống trong các chủng này, có thể phân lập ADN (ví dụ bằng cách điều chế plasmid ADN, hoặc bằng cách khuếch đại bằng kỹ thuật PCR và tách dòng đoạn PCR thu được vào vector), nuôi cấy các đột biến protein Cry trong chủng không gây đột biến, và xác định gen được đột biến có hoạt tính diệt côn trùng, ví dụ bằng cách thực hiện thử nghiệm để thử nghiệm hoạt tính diệt côn trùng. Thông thường, protein được trộn và được sử dụng trong các thử nghiệm cho ăn. Xem, ví dụ Marrone và cộng sự (1985) *J. of Economic Entomology* 78:290-293. Các thử nghiệm có thể bao gồm cho cây trồng tiếp xúc một hoặc nhiều loài gây hại và xác định khả năng sống sót của cây và/hoặc gây ra chết loài gây hại. Các ví dụ của các đột biến mà làm tăng tính độc được tìm thấy trong Schnepf et al. (1998) *Microbiol. Mol. Biol. Rev.* 62:775-806.

Tùy ý, các thay đổi có thể được thực hiện đối với trình tự axit amin theo sáng chế tại các đầu amin hoặc carboxy mà không ảnh hưởng cơ bản đến hoạt tính. Thay đổi này có thể bao gồm chèn, loại bỏ, hoặc các thay đổi được đưa vào bằng các phương pháp phân tử hiện đại, như PCR, bao gồm khuếch đại PCR mà thay thế hoặc mở rộng trình tự mã hóa protein do đưa vào trình tự mã hóa axit amin trong oligonucleotit được sử dụng trong khuếch đại PCR. Tùy ý, các trình tự protein được thêm có thể bao gồm các trình tự mã hóa protein toàn bộ, như các trình tự thường được sử dụng trong lĩnh vực để tạo ra dung hợp protein. Các protein dung hợp này thường được sử dụng để (1) làm tăng biểu hiện của protein được mong muốn (2) đưa vào vùng liên kết, hoạt tính enzym, hoặc epitop để tạo điều kiện cho hoặc là tinh chế protein, phát hiện protein, hoặc các sử dụng thử nghiệm khác đã biết trong lĩnh vực (3) bài tiết 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 thế axit amin trong protein Cry tự nhiên tạo ra protein Cry đột biến có vùng kháng nguyên mà cho phép protein Cry đột biến phân biệt với protein Cry tự nhiên trong thử nghiệm phát hiện protein.

Theo một số phương án, sáng chế đề xuất phương pháp tạo ra kháng thể mà nhận diện một cách khác protein Cry được đột biến từ protein Cry tự nhiên mà protein Cry đột biến được dẫn xuất từ đó, phương pháp bao gồm các bước 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 vùng không bảo tồn bên ngoài vùng I của protein Cry tự nhiên. Theo phương án khác, cấu trúc vòng kháng nguyên không phải là cấu trúc vòng liên quan đến nhận diện thụ thể ruột côn trùng của protein Cry hoặc liên quan đến hoạt hóa proteaza của protein Cry.

Trình tự nucleotit và axit amin biến thể theo sáng chế còn bao hàm các trình tự có nguồn gốc từ các quy trình gây đột biến và sinh tái tổ hợp như kỹ thuật đổi chỗ ADN. Với quy trình này, một hoặc nhiều vùng mã hóa protein độc khác nhau có thể

được dùng để tạo ra protein độc mới sở hữu các đặc điểm mong muốn. Bằng cách này, các thư viện của polynucleotit tái tổ hợp được tạo ra từ quần thể của polynucleotit trình tự liên quan bao gồm các vùng trình tự mà có tương đồng trình tự cơ bản và có thể được tái tổ hợp cùng loại trong ống nghiệm hoặc trong cơ thể. Ví dụ, sử dụng nghiên cứu này, mô típ trình tự mã hóa vùng được mong muốn có thể được đổi chỗ giữa gen của diệt loài gây hại theo sáng chế và gen của diệt loài gây hại đã biết khác để thu được gen mới mã hóa protein với đặc tính được mong muốn được cải thiện, như tăng hoạt tính diệt côn trùng. Các chiến lược để đổi chỗ ADN này là đã biết trong lĩnh vực. Xem, ví dụ, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer 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 các bằng sáng chế Mỹ số 5,605,793 và 5,837,458.

Sự hoán đổi hoặc đổi chỗ vùng là cơ chế khác để tạo ra các protein Cry thay đổi theo sáng chế. Các vùng có thể được hoán đổi giữa các protein Cry, tạo ra các protein độc lai hoặc khảm với hoạt tính diệt loài gây hại hoặc phổ đích cải thiện. Phương pháp tạo ra các protein tái tổ hợp và thử nghiệm chúng về hoạt tính diệt loài gây hại là đã biết 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 vectơ tái tổ hợp chứa polynucleotit, phân tử axit nucleic, caset biểu hiện hoặc gen khảm theo sáng chế. Theo các phương án khác, vectơ còn được xác định là plasmid, cosmit, phagemit, nhiễm sắc thể nhân tạo, vectơ thể thực khuẩn hoặc virus. Một số vectơ dùng trong biến nạp thực vật và các sinh vật khác là đã biết trong lĩnh vực.

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

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

Ngoài vùng khởi động liên kết hoạt động với trình tự nucleotit theo sáng chế, caset biểu hiện theo sáng chế còn có thể bao gồm trình tự điều khiển khác. Như được sử dụng ở đây, "trình tự điều khiển" 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 khiển bao gồm, nhưng không giới hạn ở, gen tăng cường, intron, trình tự dẫn dịch mã, các tín hiệu kết thúc, và trình tự tín hiệu bổ sung poly.

Theo một số phương án, caset biểu hiện theo sáng chế còn có thể bao gồm trình tự nucleotit mã hóa các đặc điểm mong muốn khác. Trình tự nucleotit này có thể được chồng lên với bất kỳ kết hợp của trình tự nucleotit để tạo cho cây, bộ phận của cây hoặc tế bào của cây có kiểu hình mong muốn. Các kết hợp chồng lên nhau này có thể được tạo ra bằng phương pháp bất kỳ bao gồm, nhưng không giới hạn ở, lai giống cây trồng bằng phương pháp thông thường bất kỳ, hoặc bằng cách biến nạp gen (tức là chồng phân tử). Nếu được chồng lên bằng các cây biến nạp gen, trình tự nucleotit được mong muốn có thể kết hợp ở thời điểm và trật tự bất kỳ. Ví dụ, cây chuyển gen chứa một hoặc nhiều đặc điểm mong muốn có thể được sử dụng làm đích để đưa vào các đặc điểm khác bằng biến nạp liên tiếp. Trình tự nucleotit bổ sung có thể được đưa vào đồng thời trong yếu tố đồng biến nạp với trình tự nucleotit, phân tử axit nucleic, kết cấu axit nucleic, và/hoặc chế phẩm theo sáng chế, với điều kiện có kết hợp các caset biểu hiện. Ví dụ, nếu hai trình tự nucleotit sẽ được đưa vào, chúng có thể kết hợp trong các tổ hợp riêng rẽ (*trans*) hoặc có thể kết hợp trong cùng tổ hợp (*cis*). Biểu hiện của polynucleotit có thể dẫn đến nhờ vùng khởi động giống nhau hoặc vùng khởi động khác nhau. Còn thấy rằng polynucleotit có thể chồng lên ở vị trí bộ gen mong muốn sử

dụng hệ tái tổ hợp đặc hiệu tại chỗ. Xem, ví dụ, các tài liệu công bố đơn sáng chế quốc tế số WO 99/25821; WO 99/25854; WO 99/25840; WO 99/25855 và WO 99/25853.

Caset biểu hiện còn có thể bao gồm trình tự mã hóa cho một hoặc nhiều polypeptit đối với các đặc điểm nông học mà cơ bản là có lợi cho công ty hạt giống, người trồng cây hoặc người xử lý hạt. Polypeptit được mong muốn có thể là bất kỳ polypeptit mã hóa bởi trình tự nucleotit mong muốn. Các ví dụ không giới hạn của các polypeptit được mong muốn mà thích hợp để tạo ra trong cây trồng bao gồm các polypeptit tạo ra các đặc điểm quan trọng về nông học như khả năng kháng chất diệt cỏ (đôi khi được gọi là "khả năng chịu thuốc diệt cỏ"), khả năng kháng virus, khả năng kháng thể gây bệnh vi khuẩn, khả năng kháng côn trùng, khả năng kháng giun tròn, và/hoặc nấm. Xem, ví dụ, tài liệu 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 được mong muốn (ví dụ, chỉ dấu chọn lọc, màu bao hạt, v.v.). Các polypeptit được mong muốn khác nhau, cũng như phương pháp để đưa các polypeptit này vào cây trồng, được mô tả, ví dụ, trong các tài liệu 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 tài liệu 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 đề kháng/khả năng chịu chất diệt cỏ ức chế điểm phát triển hoặc mô phân sinh, như imidazalinon hoặc sulfonilurea có thể thích hợp theo một số phương án theo sáng chế. Polynucleotit ví dụ trong loại này mã hóa đối với enzym ALS và AHAS đột biến như mô tả, ví dụ, trong các tài liệu bằng sáng chế Mỹ số 5,767,366 và 5,928,937. Các tài liệu bằng sáng chế Mỹ số 4,761,373 và 5,013,659 nhắm đến cây chịu các chất diệt cỏ imidazalinon hoặc sulfonamid. Tài liệu bằng 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. Tài liệu bằng sáng chế Mỹ số 5,162,602 bộc lộ cây kháng sự ức chế bởi chất diệt cỏ xyclohexanedion và axit

aryloxyphenoxypropanoic. 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, các tài liệu bằng sáng chế Mỹ số 4,940,835 và 4,769,061. Tài liệu bằng sáng chế Mỹ số 5,554,798 bộc lộ cây ngô chuyển gen kháng glyphosat, mà sức đề kháng được tạo ra nhờ gen 5-enolpyruvyl-3-phosphoshikimat (EPSP) synthaza biến đổi.

Polynucleotit mã hóa đối với hợp chất kháng phosphono như glufosinat amoni hoặc phosphinothrixin, và pyridinoxy hoặc axit phenoxy propionic và xyclohexon cũng là thích hợp. Xem, tài liệu công bố đơn Châu Âu 0 242 246. Cũng xem, các tài liệu 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, tài liệu bằng sáng chế Mỹ số 4,810,648. Polynucleotit thích hợp khác mã hóa để kháng chất diệt cỏ bao gồm các polynucleotit mã hóa để kháng axit 2,2-dichloropropionic, sethoxydim, haloxyfop, chất diệt cỏ imidazolinon, chất diệt cỏ sulfonyleurea, 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, tài liệu công bố đơn sáng chế Mỹ số 2001/0016956 và tài liệu 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). Biết rằng lượng tạo ra polypeptit diệt loài gây hại ở cây cần để kiểm soát côn trùng hoặc loài gây hại khác có thể biến đổi phụ thuộc vào cây trồng, loại gây hại, yếu tố môi trường và tương tự. Polynucleotit hữu dụng để kháng côn trùng hoặc loài gây hại bổ sung bao gồm, ví dụ, polypeptit mã hóa độc tố được xác định trong sinh vật *Bacillus*. Polynucleotit bao gồm trình tự nucleotit mã hóa *Bacillus thuringiensis* (Bt)

protein Cry từ một vài loài phụ được tách dòng và việc tách dòng tái tổ hợp được thấy có độc tính đối với ấu trùng của côn trùng bộ cánh vảy, bộ hai cánh và bộ cánh cứng. Các ví dụ của protein diệt côn trùng Bt này bao gồm các protein Cry như Cry1Aa, Cry1Ab, Cry1Ac, Cry1B, Cry1C, Cry1D, Cry1Ea, Cry1Fa, Cry3A, Cry9A, Cry9B, Cry9C, và tương tự, cũng như protein diệt côn trùng sinh dưỡng như Vip1, Vip2, Vip3, và tương tự. Danh sách đầy đủ của các protein dẫn xuất Bt có thể tìm thấy trên trang web quốc tế tại Cơ sở dữ liệu danh phát chất độc *Bacillus thuringiensis* được duy trì bởi Đại học Sussex (cũng xem, Crickmore và cộng sự (1998) *Microbiol. Mol. Biol. Rev.* 62:807-813).

Các polypeptit thích hợp để tạo ra trong thực vật còn bao gồm các polypeptit tăng cường hoặc ngược lại tạo điều kiện chuyển hóa cây và/hoặc bộ phận của cây đã thu hoạch thành sản phẩm hữu dụng bán ra thị trường, bao gồm, ví dụ, hàm lượng carbohydrat cải thiện hoặc biến đổi và/hoặc phân phối, đặc tính lên men cải thiện, tăng hàm lượng dầu, tăng hàm lượng protein, tăng khả năng tiêu hóa, và tăng hàm lượng được dinh dưỡng, ví dụ, tăng hàm lượng phytosterol, tăng hàm lượng tocopherol, tăng hàm lượng stanol và/hoặc tăng hàm lượng vitamin. Các polypeptit được mong muốn còn bao gồm, ví dụ, polypeptit tạo ra hoặc góp phần làm giảm hàm lượng thành phần không mong muốn trong mùa vụ thu hoạch, ví dụ, axit phytic, hoặc enzym làm giảm đường. Thuật ngữ "tạo ra" hoặc "góp phần" là nhằm chỉ polypeptit được mong muốn có thể đóng góp trực tiếp hoặc gián tiếp vào sự tồn tại của đặc điểm được mong muốn (ví dụ, làm tăng thoái biến xenluloza bằng cách sử dụng enzyme xenlulaza khác loại).

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

Nhiều xylanaza từ vi sinh vật nấm và vi khuẩn đã được xác định và đặc trưng (xem, ví dụ, tài liệu 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 và Reinikainen, xuất bản (1993) *Foundation for Biotechnical and Industrial Fermentation Research* 8:125-135; các tài liệu công bố

đơn sáng chế Mỹ số 2005/0208178; và công bố đơn PCT số WO 03/16654). Cụ thể là, ba xylanaza cụ thể (XYL-I, XYL-II, và XYL-III) đã được xác định trong *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 cho sáng chế có thể là enzym làm giảm polysaccarit. 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, cellulaza, hemicellulaza, xylanaza, cyclodextrin 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 other exo-amylaza; 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 tài liệu bằng 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), glucose oxidaza, pectinaza, xylanaza, transglucosidaza, alpha 1,6 glucosidaza (ví dụ, E.C. 3.2.1.20); esteraza như ferulic axit esteraza (EC 3.1.1.73) và axetyl xylan esteraza (EC 3.1.1.72); và cutinaza (ví dụ E.C. 3.1.1.74).

Theo một số phương án, sáng chế đề xuất tế bào chủ chuyển gen không phải của người chứa polynucleotit, phân tử axit nucleic, gen khảm, caset biểu hiện hoặc vectơ tái tổ hợp theo sáng chế. Tế bào chủ chuyển gen không phải của người có thể bao gồm, nhưng không giới hạn ở, tế bào của cây, tế bào nấm men, tế bào vi khuẩn hoặc tế bào côn trùng. Theo đó, theo một số phương án, sáng chế đề xuất tế bào vi khuẩn được chọn từ 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, làm tác nhân kiểm soát côn trùng sinh học, các protein Cry theo sáng chế có thể được tạo ra bằng cách biểu hiện gen khảm mã hóa các protein Cry theo sáng chế trong tế bào vi khuẩn. Ví dụ, theo một phương án, sáng chế đề xuất tế bào *Bacillus thuringiensis* chứa gen khảm theo sáng chế.

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

Theo phương án khác của sáng chế, protein độc theo sáng chế được biểu hiện trong sinh vật cao hơn, ví dụ, thực vật. Trong trường hợp này, cây chuyển gen biểu hiện lượng có hiệu quả của protein độc bảo vệ chính chúng khỏi loài gây hại cây trồng như loài gây hại côn trùng. Khi côn trùng bắt đầu gây hại trên cây chuyển gen này, nó cũng ăn độc tố được biểu hiện. Điều này có thể ngăn cản côn trùng khỏi hút thêm vào mô thực vật hoặc có thể còn gây hại hoặc giết côn trùng. Polynucleotit theo sáng chế được chèn vào caset biểu hiện, mà sau đó được tích hợp ổn định trong bộ 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 thành loài thực vật cụ thể, có thể nhân giống loài này hoặc chuyển thành các dạng khác của loài tương tự, cụ thể là bao gồm các dạng kinh doanh được, sử dụng các kỹ thuật sinh sản truyền thống.

Polynucleotit theo sáng chế được biểu hiện trong cây chuyển gen, do đó tạo ra sinh tổng hợp protein Cry tương đương trong cây chuyển gen. Bằng cách này, cây chuyển gen có sự bảo vệ năng suất được tăng cường với sự có mặt của áp lực côn trùng được tạo ra. Đối với biểu hiện của chúng trong cây chuyển gen, trình tự nucleotit theo sáng chế có thể cần biến đổi và tối ưu hóa. Mặc dù trong nhiều trường hợp vi sinh vật có thể được biểu hiện ở cây ở các mức cao mà không cần biến đổi, biểu hiện thấp trong cây chuyển gen có thể tạo ra từ trình tự nucleotit của vi khuẩn có bộ ba mã mà không được ưu tiên trong cây trồng. Đã biết trong lĩnh vực rằng các sinh vật sống có các ưu tiên cụ thể đối với việc sử dụng bộ ba mã, và bộ ba mã của các trình tự

nucleotit được mô tả trong sáng chế có thể được thay đổi để thích hợp với ưu tiên của cây, trong khi vẫn giữ các axit amin được mã hóa nhờ đó. Hơn nữa, biểu hiện cao ở cây, ví dụ cây ngô, tốt nhất thu được từ các trình tự mã hóa mà có ít nhất khoảng 35% hàm lượng GC, hoặc ít nhất khoảng 45%, hoặc ít nhất khoảng 50%, hoặc ít nhất khoảng 60%. Trình tự nucleotit của vi khuẩn mà có hàm lượng GC thấp có thể biểu hiện nghèo nàn ở cây do sự tồn tại của mô típ ATTTA mà có thể làm bất ổn định các thông tin, và các mô típ AATAAA mà có thể tạo ra sự polyadenyl không thích hợp. Mặc dù một số các trình tự gen nhất định có thể được biểu hiện một cách thích hợp trong cả các loài cây một lá mầm và hai lá mầm, các trình tự có thể được biến đổi để tính cho các ưu tiên bộ ba mã cụ thể và ưu tiên hàm lượng GC của cây một lá mầm hoặc cây hai lá mầm như các ưu tiên này đã được thể hiện để phân biệt (Murray và cộng sự Nucl. Acids Res. 17:477-498 (1989)). Ngoài ra, trình tự nucleotit được sàng lọc đối với sự tồn tại của các vị trí nối không hợp lệ mà có thể gây ra cắt cụt thông tin. Tất cả các thay đổi cần được thực hiện trong các trình tự nucleotit như các trình tự được mô tả ở trên được thực hiện bằng cách sử dụng các kỹ thuật đã biết về gây đột biến hướng điểm, PCR, và xây dựng gen tổng hợp bằng cách sử dụng phương pháp được mô tả ví dụ trong các 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 gen tổng hợp được thực hiện 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 được ưu tiên của ngô, tức là, bộ ba mã đơn mà mã hóa thường xuyên nhất axit amin ở ngô, được sử dụng. Bộ ba mã được ưu tiên của ngô cho axit amin cụ thể có thể được dẫn xuất, ví dụ, từ các trình tự gen đã biết từ ngô. Ví dụ, việc sử dụng bộ ba mã của ngô cho 28 gen từ cây ngô được tìm thấy trong Murray và cộng sự, Axit nucleics Research 17:477-498 (1989), nội dung của nó được kết hợp ở đây bằng cách viện dẫn. Trình tự tổng hợp được lấy làm ví dụ đặc biệt theo sáng chế được thực hiện với bộ ba mã được tối ưu hóa của ngô được đại diện bởi bất kỳ một trong số các SEQ ID NOs: 16-25. 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 bao gồm trình tự nucleotit mà là trình tự tự nhiên một phần và trình tự được tối ưu hóa tổng hợp một phần.

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

Các trình tự mã hóa protein Cry mới theo sáng chế, hoặc là trình tự tự nhiên của chúng hoặc là các trình tự tổng hợp như được mô tả ở trên, có thể được dung hợp hoạt động với loại vùng khởi động để biểu hiện trong cây trồng bao gồm vùng khởi động cấu trúc, cảm ứng, được điều biến tạm thời, được điều biến phát triển, được điều biến hóa học, được ưu tiên mô và đặc hiệu mô để điều chế phân tử ADN tái tổ hợp, tức là, gen khảm. Việc lựa chọn vùng khởi động sẽ thay đổi tùy thuộc vào yêu cầu thời gian và không gian, và còn tùy thuộc vào loài đích. Do đó, biểu hiện của các trình tự nucleotit theo sáng chế trong lá, trong cuống hoặc thân, trong tai, trong cụm hoa (ví dụ bông, chùy, lõi, v.v.), trong rễ, và/hoặc cây con giống được ưu tiên. Trong nhiều trường hợp, tuy nhiên, đòi hỏi bảo vệ chống lại nhiều hơn một loại côn trùng gây hại, và do đó biểu hiện trong nhiều mô là cần thiết. Mặc dù nhiều vùng khởi động từ các cây hai lá mầm đã thể hiện hoạt động trong các cây một lá mầm và ngược lại, lý tưởng là các vùng khởi động của cây hai lá mầm được chọn để biểu hiện trong cây hai lá mầm và các vùng khởi động của cây một lá mầm để biểu hiện trong cây một lá mầm. Tuy nhiên, không có hạn chế đối với nguồn gốc của vùng khởi động được chọn; chỉ cần chúng hoạt động trong việc dẫn hướng biểu hiện của trình tự nucleotit trong tế bào mong muốn.

Các ví dụ của vùng khởi động chủ yếu hữu dụng theo sáng chế bao gồm vùng khởi động CaMV 35S và 19S (Fraley và cộng sự, tài liệu bằng sáng chế Mỹ số 5,352,605, được kết hợp ở đây bằng cách viện dẫn). Ngoài ra, vùng khởi động có nguồn gốc từ bất kỳ một trong số vài gen actin, mà được biểu hiện trong hầu hết các kiểu tế bào. Caset biểu hiện của vùng khởi động được mô tả bởi McElroy và cộng sự (Mol. Gen. Genet. 231: 150-160 (1991)) có thể dễ dàng biến đổi để biểu hiện gen độc

mới và đặc biệt phù hợp để dùng trong vật chủ của cây một lá mầm. Vùng khởi động chủ yếu khác nữa có nguồn gốc từ ubiquitin, mà là sản phẩm gen khác được biết để tích tụ trong nhiều kiểu tế bào. Vùng khởi động ubiquitin được tách dòng từ một vài loài để dùng trong cây chuyển gen, ví dụ, hướng dương (Binet et al., 1991. Plant Science 79: 87-94), ngô (Christensen et al., 1989. Plant Molec. Biol. 12: 619-632), và arabidopsis (Norris et al. 1993. Plant Molec. Biol. 21:895-906). Vùng khởi động ubiquitin của ngô được phát triển trong hệ một lá mầm chuyển gen và trình tự và vector của nó được cấu trúc cho biến nạp một lá mầm được bộc lộ trong công bố đơn sáng chế EP 0 342 926. Vùng khởi động ubiquitin thích hợp để biểu hiện gen độc mới trong cây chuyển gen, đặc biệt là cây một lá mầm.

Vùng khởi động đặc hiệu mô hoặc ưu tiên mô hữu dụng để biểu hiện trình tự mã hóa protein Cry mới theo sáng chế ở cây trồng, cụ thể là ngô, là các loại mà định hướng biểu hiện trong rễ, ruột cây, lá hoặc phần hoa. Vùng khởi động này được bộc lộ trong tài liệu đơn sáng chế Mỹ số 5625136, được kết hợp ở đây bằng cách viện dẫn toàn bộ. Vùng khởi động đặc hiệu mô khác hữu dụng theo sáng chế bao gồm vùng khởi động rubisco của bông được bộc lộ trong tài liệu bằng sáng chế Mỹ số 6,040,504; vùng khởi động đường synthaza của lúa được bộc lộ trong tài liệu bằng sáng chế Mỹ số 5,604,121; và vùng khởi động của virus cuốn lá vàng ở cây dứa hương được bộc lộ trong tài liệu bằng sáng chế Mỹ số 7,166,770, tất cả được kết hợp bằng cách viện dẫn toàn bộ. Vùng khởi động cảm ứng hóa học hữu dụng để dẫn hướng biểu hiện của gen độc mới trong cây được bộc lộ trong tài liệu bằng sáng chế Mỹ số 5,614,395 được kết hợp ở đây bằng cách viện dẫn toàn bộ.

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

Nhóm vùng khởi động khác hữu dụng theo sáng chế là loại mà cảm ứng cuộn. Nhiều vùng khởi động đã được mô tả mà được biểu hiện tại vị trí cuộn và cả ở các vị trí nhiễm bệnh thực vật. Lý tưởng là, vùng khởi động này chỉ có thể có hoạt tính cục bộ tại vị trí côn trùng xâm chiếm, và bằng cách này protein diệt côn trùng chỉ xếp

chồng trong các tế bào cần tổng hợp protein diệt côn trùng để diệt loài gây hại côn trùng xâm chiếm. Các ví dụ của vùng khởi động thuộc loại này bao gồm loại 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).

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

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

Nhiều trình tự khác có thể được kết hợp vào caset biểu hiện được 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 được mô tả là tạo ra sự nhắm đích của các sản phẩm gen đến các ngăn tế bào khác. Các trình tự đầu amin có thể chịu trách nhiệm nhắm đích protein được mong muốn đến ngăn tế bào bất kỳ, như, không bào, ty thể, vi thể peroxys, các thể protein, lưới nội bào, lục lạp, tinh bột hạt, hạt bột, vô bào hoặc thành tế bào của cây (ví dụ Unger et. al. Plant Molec. Biol. 13: 411-418 (1989); Rogers et. al. (1985) Proc. Natl. Acad. Sci. USA 82: 6512-651; tài liệu bằng sáng chế Mỹ số 7,102,057; công bố đơn số WO 2005/096704, tất cả chúng được kết hợp tại đây bằng cách viện dẫn. Tùy ý, trình tự tín hiệu có thể là trình tự tín hiệu đầu N từ sáp, trình tự tín hiệu đầu N từ gamma-đạm ngô, vùng liên kết tinh bột, vùng liên kết tinh bột đầu 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; tài liệu bằng sáng chế Mỹ số 5,639,949) hoặc trình tự tín hiệu bài tiết từ các tế bào aleuron (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Ngoài ra, các trình tự đầu amin cùng với các trình tự đầu carboxy chịu trách nhiệm 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 vùng khởi động cùng nguồn của chúng, mà còn cùng với vùng khởi động khác loại sao cho ảnh hưởng đến mục tiêu nhắm đích tế bào cụ thể dưới sự điều biến phiên mã của vùng khởi động mà có mẫu biểu hiện khác của vùng khởi động từ đó tạo ra tín hiệu nhắm đích.

Biến nạp cây

Quy trình để biến nạp cây là đã biết và quen thuộc trong lĩnh vực và được mô tả qua các tài liệu tham khảo. Các ví dụ không giới hạn của phương pháp để biến nạp cây bao gồm biến nạp thông qua phân phối axit nucleic trung gian vi khuẩn (ví dụ, thông qua *Agrobacterium*), phân phối axit nucleic trung gian virus, phân phối axit nucleic

trung gian silic cacbua hoặc sợi tinh thể axit nucleic, phân phối axit nucleic trung gian hạt mỡ, vi truyền, thả bom vi hạt, biến nạp trung gian canxi-phosphat, biến nạp trung gian xyclodextrin, điện di, biến nạp trung gian hạt nano, nghiền bằng sóng siêu âm, thẩm thấu, hấp thu axit nucleic trung gian PEG, cũng như bất kỳ cơ chế điện, hóa học, vật lý (cơ khí) và/hoặc sinh học khác mà tạo ra sự đưa axit nucleic vào tế bào của cây, bao gồm kết hợp bất kỳ của chúng. Các hướng dẫn chung về các phương pháp biến nạp cây khác nhau là đã biết trong lĩnh vực này bao gồm Miki và cộng sự. ("Procedures for Introducing Foreign DNA into Plants" trong *Method in Plant Molecular Biology and Biotechnology*, Glick, B. R. và Thompson, J. E., Eds. (CRC Press, Inc., Boca Raton, 1993), trang 67-88) và Rakowoczy-Trojanowska (*Cell. Mol. Biol. Lett.* 7:849-858 (2002)).

Đối với biến nạp trung gian *Agrobacterium*, các vector hai kỳ hoặc vector mang ít nhất một trình tự viền T-ADN thích hợp, trong khi đó đối với chuyển gen trực tiếp (ví dụ, thả bom hạt và tương tự) vector bất kỳ thích hợp và ADN thẳng chỉ chứa cấu trúc được mong muốn có thể được sử dụng. Trong trường hợp chuyển gen trực tiếp, biến nạp với các 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 đầu chọn lọc có thể là lựa chọn dương (Phosphomannoza Isomerase), tạo ra khả năng kháng kháng sinh (kanamycin, hygromycin hoặc methotrexat) hoặc chất diệt cỏ (glyphosat hoặc glufosinat). Tuy nhiên, việc lựa chọn đầu chọn lọc không thuộc phạm vi sáng chế.

Biến nạp trung gian *Agrobacterium* là phương pháp thường được dùng để biến nạp cây, cụ thể là, cây hai lá mầm, nhờ tính hiệu quả biến nạp cao của nó và nhờ tính hữu ích rộng với nhiều loài khác nhau. Biến nạp trung gian *Agrobacterium* thường liên quan đến việc chuyển vector hai kỳ mang ADN lạ được mong muốn cho chủng *Agrobacterium* thích hợp có thể phụ thuộc vào sự bổ sung của gen *vir* được mang bởi chủng *Agrobacterium* của vật chủ hoặc là trên Ti plasmit đồng nội trú hoặc 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.

Tùy ý, vector hai kỳ tái tổ hợp có thể được chuyển vào *Agrobacterium* bằng cách biến nạp axit nucleic (Höfgen & Willmitzer (1988) *Nucleic Acids Res.* 16:9877).

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 dấu 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, tài liệu 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, vector có thể được đưa vào tế bào bằng cách phủ các hạt bằng các vector chứa axit nucleic được mong muốn. Tùy ý, tế bào hoặc thể bào có thể bao quanh bởi vector mà vector mang vào tế bào bằng cách theo hạt. Các hạt hoạt tính sinh học (ví dụ, tế bào nấm men khô, vi khuẩn khô hoặc vật ăn vi khuẩn, mỗi hạt chứa một hoặc nhiều axit nucleic cần đưa vào) còn có thể được đẩy vào mô thực vật.

Theo phương án khác, polynucleotit theo sáng chế có thể được biến nạp trực tiếp vào bộ gen lặp thể. Lợi ích chủ yếu của biến nạp lặp thể là các lặp thể thường có khả năng biểu hiện gen vi khuẩn mà không làm biến đổi cơ bản, và các lặp thể có khả năng biểu hiện khung đọc mở phức dưới kiểm soát của vùng khởi động đơn. Kỹ thuật biến nạp lặp thể được mô tả chi tiết trong các tài liệu bằng sáng chế Mỹ số 5,451,513, 5,545,817, và 5,545,818, trong tài liệu 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 vùng ADN lặp thể đã tách dòng giới hạn hai đầu một chỉ dấu chọn lọc cùng với gen được mong muốn vào mô đích thích hợp, ví dụ, dùng súng bắn gen hoặc biến nạp tế bào trần (như canxi clorua hoặc biến nạp trung gian PEG). Các vùng giới hạn hai đầu 1-1,5kb, được gọi là các trình tự đích, tạo điều kiện cho tái tổ hợp tương đồng với bộ gen lặp thể và do đó cho phép thay thế hoặc sửa đổi các vùng cụ thể của hệ gen lặp thể. Đầu tiên, các đột biến điểm trong gen 16S rARN và rps12 lặp thể tạo ra khả năng kháng spectinomycin và/hoặc streptomycin có thể được sử dụng làm chỉ dấu lựa chọn để biến nạp (Svab, Z., Hajdukiewicz, P., và

Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., và Maliga, P. (1992) *Plant Cell* 4, 39-45). Sự có mặt của các điểm tách dòng giữa các dấu hiệu này cho phép tạo ra vector nhắm đích lap thể đưa vào các gen lạ (Staub, J.M., và Maliga, P. (1993) *EMBO J.* 12, 601-606). Sự tăng cơ bản về tần suất biến nạp có thể thu được bằng cách thay đổi vị trí của các gen lặn kháng kháng sinh rARN hoặc r-protein với chỉ dấu chọn lọc trội, gen aadA của vi khuẩn mã hóa enzym giải độc spectinomycin aminoglycosit- 3'- adenytransferaza (Svab, Z., và Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917). Trước đây, chỉ dấu này đã được dùng thành công để biến nạp tần suất cao bộ gen lap thể của tảo lục *Chlamydomonas reinhardtii* (Goldschmidt-Clermont, M. (1991) *Nucl. Acids Res.* 19:4083-4089). Các chỉ dấu chọn lọc khác hữu dụng cho biến nạp lap 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 lap thể, mà trong đó các gen được chèn bởi tái tổ hợp tương đồng vào toàn bộ vài nghìn bản sao của bộ gen lap thể vòng có trong mỗi tế bào của cây, lợi dụng số lượng rất lớn bản sao lợi thế hơn các gen được biểu hiện trong nhân để cho phép các mức biểu hiện có thể dễ dàng vượt quá 10% tổng protein thực vật hoà tan. Theo một phương án, polynucleotit theo sáng chế có thể được chèn vào vector định hướng lap thể và được biến nạp vào bộ gen lap thể của vật chủ thực vật mong muốn. Do đó, thực vật đồng nhất đối với các bộ gen lap thể chứa trình tự nucleotit của sáng chế có thể thu được, mà có khả năng biểu hiện cao polynucleotit.

Các phương pháp chọn lọc cho cây biến nạp, chuyển gen, tế bào của cây và/hoặc mô nuôi cấy mô thực vật là quen thuộc trong lĩnh vực và có thể được sử dụng trong các phương án theo sáng chế được đề xuất ở đây. Ví dụ, vector tái tổ hợp theo sáng chế còn có thể bao gồm caset biểu hiện bao gồm trình tự nucleotit cho chỉ dấu chọn lọc, mà có thể được dùng để lựa chọn cây, bộ phận của cây và/hoặc tế bào cây được biến nạp. Như được sử dụng ở đây, "chỉ dấu chọn lọc" nghĩa là trình tự nucleotit mà khi được biểu hiện thể hiện kiểu hình phân biệt với cây, bộ phận của cây và/hoặc tế bào của cây biểu hiện chỉ dấu và do đó cho phép cây, bộ phận của cây và/hoặc tế bào của cây được biến nạp này phân biệt với loại mà không có chỉ dấu. Trình tự nucleotit có thể mã hóa hoặc là chỉ dấu chọn lọc hoặc sàng lọc, phụ thuộc và chỉ dấu tạo ra đặc điểm có thể được chọn lọc bằng phương pháp hóa học, như bằng cách sử dụng chất

chọn lọc (ví dụ, chất kháng sinh, chất diệt cỏ, hoặc tương tự), hoặc liệu chỉ dấu đơn giản là đặc điểm mà có thể xác định thông qua việc quan sát hoặc thử, như bằng cách sàng lọc (ví dụ, đặc điểm nơi cư trú R). Đương nhiên, nhiều ví dụ về chỉ dấu chọn lọc thích hợp là đã biết trong lĩnh vực và có thể được sử dụng trong tổ hợp có thể biểu hiện mô tả ở đây.

Các ví dụ của chỉ dấu chọn lọc bao gồm, nhưng không giới hạn ở, trình tự nucleotit mã hóa *neo* hoặc *nptII*, mà tạo ra khả năng kháng kanamycin, G418, và tương tự (Potrykus *et al.* (1985) *Mol. Gen. Genet.* 199:183-188); trình tự nucleotit mã hóa *bar*, mà tạo ra khả năng kháng phosphinothrixin; trình tự nucleotit mã hóa 5-enolpyruvylshikimat-3-phosphat (EPSP) synthaza được thay đổi, mà tạo ra khả năng kháng glyphosat (Hinchee *et al.* (1988) *Biotech.* 6:915-922); trình tự nucleotit mã hóa nitrilaza như *bxn* từ *Klebsiella ozaenae* mà tạo ra khả năng kháng bromoxynil (Stalker *et al.* (1988) *Science* 242:419-423); trình tự nucleotit mã hóa axetolactat synthaza (ALS) được thay đổi mà tạo ra khả năng kháng imidazolinon, sulfonylurê hoặc các chất hóa học ức chế ALS khác (tài liệu đơn sáng chế châu Âu số 154204); trình tự nucleotit mã hóa dihydrofolat reductaza kháng methotrexat (DHFR) (Thillet *et al.* (1988) *J. Biol. Chem.* 263:12500-12508); trình tự nucleotit mã hóa dalapon dehalogenaza mà tạo ra khả năng kháng dalapon; trình tự nucleotit mã hóa mannoza-6-phosphat isomeraza (còn gọi là phosphomannoza isomeraza (PMI)) mà tạo ra khả năng chuyển hóa mannoza (các tài liệu bằng sáng chế Mỹ số 5,767,378 và 5,994,629); trình tự nucleotit mã hóa anthranilat synthaza được thay đổi mà tạo ra khả năng kháng 5-metyl tryptophan; và/hoặc trình tự nucleotit mã hóa *hph* mà tạo ra khả năng kháng hygromycin. Người có hiểu biết trung bình trong lĩnh vực có khả năng chọn chỉ dấu chọn lọc thích hợp để dùng trong caset biểu hiện của sáng chế.

Các chỉ dấu 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 thực vật (Dellaporta và cộng sự, "Molecular cloning of the maize R-nj allele by transposon-tagging with Ac" 263-282 trong: Cấu trúc và chức năng nhiễm sắc thể: Impact of New Concepts, 18th Stadler Genetics Symposium (Gustafson & Appels eds., Plenum Press 1988)); trình tự nucleotit mã hóa β -lactamaza, enzym mà các nền chất sinh màu

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

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

Theo đó, các đặc tính gen được thiết kế vào hạt chuyển gen và thực vật, bộ phận của cây, và/hoặc tế bào của cây theo sáng chế được mô tả ở trên có thể vượt qua tái tạo bằng giới tính hoặc nuôi sinh dưỡng và do đó có thể duy trì và nhân giống ở các cây thế hệ con cái. 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, bộ phận của cây và/hoặc tế bào của cây theo nhiều con đường bất kỳ đã biết trong lĩnh vực, như được mô tả ở trên. Do đó, không có phương pháp cụ thể để đưa một hoặc nhiều polynucleotit vào cây, mà có thể sử dụng phương pháp bất kỳ mà cho phép một hoặc nhiều polynucleotit được tích hợp ổn định trong bộ gen của cây. Khi hơn một polynucleotit được đưa vào,

polynucleotit tương ứng có thể được ghép thành một phần của phân tử axit nucleic đơn, hoặc là phân tử axit nucleic riêng, và có thể được định vị trên phân tử axit nucleic giống hoặc khác nhau. Theo đó, polynucleotit có thể được đưa vào tế bào được mong muốn trong sự kiện biến nạp đơn, trong sự kiện biến nạp riêng rẽ, hoặc, ví dụ, trong thực vật, là một phần của yếu tố nhân giống.

Các phương án bổ sung của sáng chế bao gồm sản phẩm thu hoạch được tạo ra từ cây chuyển gen và/hoặc các phần của chúng theo sáng chế, cũng như sản phẩm xử lý tạo ra từ sản phẩm thu hoạch. Các sản phẩm thu hoạch có thể là tổng thể cây hoặc bộ phận bất kỳ của cây, như được mô tả ở đây. Do đó, theo một số phương án, các ví dụ không giới hạn của sản phẩm thu hoạch bao gồm hạt, quả, hoa hoặc bộ phận của chúng (ví dụ, bao phấn, nhụy, và tương tự), lá, thân, và tương tự. Theo các phương án khác, sản phẩm xử lý bao gồm, nhưng không giới hạn ở, bột mịn, bột xay thô, dầu, tinh bột, ngũ cốc, và tương tự được tạo ra từ hạt được thu hoạch hoặc bộ phận khác của cây theo sáng chế, trong đó hạt này hoặc bộ phận khác của cây bao gồm trình tự phân tử axit nucleic/polynucleotit/nucleotit theo sáng chế.

Theo các phương án khác, sáng chế đề xuất phần chiết từ hạt chuyển gen và/hoặc cây chuyển gen theo sáng chế, trong đó phần chiết bao gồm phân tử axit nucleic, polynucleotit, trình tự nucleotit hoặc protein độc theo sáng chế. Các phần chiết từ cây trồng hoặc bộ phận của cây có thể được thực hiện 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)).

Chế phẩm diệt côn trùng

Theo một số phương án, sáng chế đề xuất chế phẩm diệt côn trùng bao gồm protein Cry theo sáng chế trong chất mang nông dụng. Như được sử dụng ở đây "chất mang nông dụng" có thể là tự nhiên hoặc tổng hợp, vật liệu hữu cơ hoặc vô cơ mà kết hợp với thành phần hoạt tính để tạo điều kiện cho việc áp dụng nó lên cây, hoặc bộ phận của cây. 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 chế phẩm nông nghiệp. Các chế phẩm này có thể được áp dụng theo cách

bất kỳ mà mang protein diệt loài gây hại hoặc tác nhân kiểm soát loài gây hại khác tiếp xúc với loài gây hại. Theo đó, chế phẩm có thể được áp dụng lên bề mặt của cây hoặc bộ phận của cây, bao gồm hạt, lá, hoa, thân, cuống, rễ, và tương tự. Chất mang nông dụng khác có thể là cây chuyển gen hoặc bộ phận của cây.

Theo các phương án khác nữa, chế phẩm diệt côn trùng bao gồm tế bào vi khuẩn chuyển gen theo sáng chế, trong đó tế bào vi khuẩn chứa gen khả năng theo sáng chế. Ví dụ, chế phẩm diệt côn trùng này có thể được điều chế bằng cách làm khô, làm khô lạnh, đông thể hóa, chiết, lọc, li tâm, đóng cặn, hoặc cô đặc mề nuôi cấy của các tế bào *Bacillus thuringiensis* chứa polynucleotit theo sáng chế. Theo các phương án bổ sung, chế phẩm bao gồm từ khoảng 1% đến khoảng 99% theo trọng lượng của protein Cry theo sáng chế.

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

Theo một số phương án, tác nhân kiểm soát loài gây hại thứ hai có thể là tác nhân được chọn từ nhóm bao gồm chất diệt loài gây hại hóa học, 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 dạng serin và cysteine), lectin, alpha - amylaza, peroxidaza và cholesterol oxidaza.

Theo các phương án khác, tác nhân kiểm soát loài gây hại thứ hai là chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm pyrethroid, carbamat, neonicotinoid, chất chặn kênh natri neuron, lacton vòng lớn diệt côn trùng, chất đối vận axit gamma

aminobutyric (GABA), urê diệt côn trùng và chất bắt chước juvenil hormon. Theo phương án khác, chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm abamectin, axephat, axetamiprit, amidoflomet (S-1955), avermectin, azadirachtin, azinphos-metyl, bifenthrin, binfenazat, buprofezin, carbofuran, chlorfenapyr, chlorfluazuron, chlorpyrifos, chlorpyrifos-metyl, chromafenozit, clothianidin, xyfluthrin, beta-xyfluthrin, xyhalothrin, lambda-xyhalothrin, xypermethrin, xyromazin, deltamethrin, diafenthiuron, diazinon, diflubenzuron, dimethoat, diofenolan, emamectin, endosulfan, esfenvalerat, ethiprol, fenothicarb, fenoxycarb, fenpropathrin, fenproximat, fenvalerat, fipronil, flonicamid, fluxythrinate, tau-fluvalinat, flufenerim (UR-50701), flufenoxuron, fonophos, halofenozit, hexaflumuron, imidacloprid, indoxacarb, isofenphos, lufenuron, malathion, metaldehyt, methamidophos, methidathion, methomyl, methopren, methoxychlor, monocrotophos, methoxyfenozit, nithiazin, novaluron, noviflumuron (XDE-007), oxamyl, parathion, parathion-metyl, permethrin, phorat, phosalon, phosmet, phosphamidon, pirimicarb, profenofos, pymetrozin, pyridalyl, pyriproxyfen, rotenon, spinosad, spiromesifin (BSN 2060), sulprofos, tebufenozit, teflubenzuron, tefluthrin, terbufos, tetrachlorvinphos, thiacloprid, thiamethoxam, thiodicarb, thiosultap-natri, tralomethrin, trichlorfon và triflumuron, aldicarb, oxamyl, fenamiphos, amitraz, chinomethionat, chlorobenzilat, xyhexatin, dicofol, dienochlor, etoxazole, fenazaquin, fenbutatin oxit, fenpropathrin, fenpyroximat, hexythiazox, propargit, pyridaben và tebufenpyrad. Theo phương án khác, chất diệt loài gây hại hóa học được chọn từ nhóm bao gồm 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, diofenolan, 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 các phương án khác, tác nhân kiểm soát loài gây hại thứ hai là protein Cry được chọn từ nhóm bao gồm Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ad, Cry1Ae, Cry1Af,

Cry1Ag, Cry1Ah, Cry1Ai, Cry1Aj, Cry1Ba, Cry1Bb, Cry1Bc, Cry1Bd, Cry1Be,
 Cry1Bf, Cry1Bg, Cry1Bh, Cry1Bi, Cry1Ca, Cry1Cb, Cry1Da, Cry1Db, Cry1Dc,
 Cry1Dd, Cry1Ea, Cry1Eb, Cry1Fa, Cry1Fb, Cry1Ga, Cry1Gb, Cry1Gc, Cry1Ha,
 Cry1Hb, Cry1Hc, Cry1Ia, Cry1Ib, Cry1Ic, Cry1Id, Cry1Ie, Cry1If, Cry1Ig, Cry1Ja,
 Cry1Jb, Cry1Jc, Cry1Jd, Cry1Ka, Cry1La, Cry1Ma, Cry1Na, Cry1Nb, Cry2Aa,
 Cry2Ab, Cry2Ac, Cry2Ad, Cry2Ae, Cry2Af, Cry2Ag, Cry2Ah, Cry2Ai, Cry2Aj,
 Cry2Ak, Cry2Al, Cry2Ba, Cry3Aa, Cry3Ba, Cry3Bb, Cry3Ca, Cry4Aa, Cry4Ba,
 Cry4Ca, Cry4Cb, Cry4Cc, Cry5Aa, Cry5Ab, Cry5Ac, Cry5Ad, Cry5Ba, Cry5Ca,
 Cry5Da, Cry5Ea, Cry6Aa, Cry6Ba, Cry7Aa, Cry7Ab, Cry7Ac, Cry7Ba, Cry7Bb,
 Cry7Ca, Cry7Cb, Cry7Da, Cry7Ea, Cry7Fa, Cry7Fb, Cry7Ga, Cry7Gb, Cry7Gc,
 Cry7Gd, Cry7Ha, Cry7Ia, Cry7Ja, Cry7Ka, Cry7Kb, Cry7La, Cry8Aa, Cry8Ab,
 Cry8Ac, Cry8Ad, Cry8Ba, Cry8Bb, Cry8Bc, Cry8Ca, Cry8Da, Cry8Db, Cry8Ea,
 Cry8Fa, Cry8Ga, Cry8Ha, Cry8Ia, Cry8Ib, Cry8Ja, Cry8Ka, Cry8Kb, Cry8La,
 Cry8Ma, Cry8Na, Cry8Pa, Cry8Qa, Cry8Ra, Cry8Sa, Cry8Ta, Cry9Aa, Cry9Ba,
 Cry9Bb, Cry9Ca, Cry9Da, Cry9Db, Cry9Dc, Cry9Ea, Cry9Eb, Cry9Ec, Cry9Ed,
 Cry9Ee, Cry9Fa, Cry9Ga, Cry10Aa, Cry11Aa, Cry11Ba, Cry11Bb,
 Cry12Aa, Cry13Aa, Cry14Aa, Cry14Ab, Cry15Aa, Cry16Aa, Cry17Aa, Cry18Aa,
 Cry18Ba, Cry18Ca, Cry19Aa, Cry19Ba, Cry19Ca, Cry20Aa, Cry20Ba, Cry21Aa,
 Cry21Ba, Cry21Ca, Cry21Da, Cry21Ea, Cry21Fa, Cry21Ga, Cry21Ha, Cry22Aa,
 Cry22Ab, Cry22Ba, Cry22Bb, Cry23Aa, Cry24Aa, Cry24Ba, Cry24Ca, Cry25Aa,
 Cry26Aa, Cry27Aa, Cry28Aa, Cry29Aa, Cry29Ba, Cry30Aa, Cry30Ba, Cry30Ca,
 Cry30Da, Cry30Db, Cry30Ea, Cry30Fa, Cry30Ga, Cry31Aa, Cry31Ab, Cry31Ac,
 Cry31Ad, Cry32Aa, Cry32Ab, Cry32Ba, Cry32Ca, Cry32Cb, Cry32Da, Cry32Ea,
 Cry32Eb, Cry32Fa, Cry32Ga, Cry32Ha, Cry32Hb, Cry32Ia, Cry32Ja, Cry32Ka,
 Cry32La, Cry32Ma, Cry32Mb, Cry32Na, Cry32Oa, Cry32Pa, Cry32Qa, Cry32Ra,
 Cry32Sa, Cry32Ta, Cry32Ua, Cry33Aa, Cry34Aa, Cry34Ab, Cry34Ac, Cry34Ba,
 Cry35Aa, Cry35Ab, Cry35Ac, Cry35Ba, Cry36Aa, Cry37Aa, Cry38Aa, Cry39Aa,
 Cry40Aa, Cry40Ba, Cry40Ca, Cry40Da, Cry41Aa, Cry41Ab, Cry41Ba, Cry42Aa,
 Cry43Aa, Cry43Ba, Cry43Ca, Cry43Cb, Cry43Cc, Cry44Aa, Cry45Aa, Cry46Aa,
 Cry46Ab, Cry47Aa, Cry48Aa, Cry48Ab, Cry49Aa, Cry49Ab, Cry50Aa, Cry50Ba,
 Cry51Aa, Cry52Aa, Cry52Ba, Cry53Aa, Cry53Ab, Cry54Aa, Cry54Ab, Cry54Ba,
 Cry55Aa, Cry56Aa, Cry57Aa, Cry57Ab, Cry58Aa, Cry59Aa, Cry59Ba, Cry60Aa,

Cry60Ba, Cry61Aa, Cry62Aa, Cry63Aa, Cry64Aa, Cry65Aa, Cry66Aa, Cry67Aa, Cry68Aa, Cry69Aa, Cry69Ab, Cry70Aa, Cry70Ba, Cry70Bb, Cry71Aa, Cry72Aa và Cry73Aa.

Theo các phương án khác nữa, tác nhân kiểm soát loài gây hại thứ hai là protein diệt côn trùng sinh dưỡng Vip3 được chọn từ nhóm bao gồm Vip3Aa1, Vip3Aa2, Vip3Aa3, Vip3Aa4, Vip3Aa5, Vip3Aa6, Vip3Aa7, Vip3Aa8, Vip3Aa9, Vip3Aa10, Vip3Aa11, Vip3Aa12, Vip3Aa13, Vip3Aa14, Vip3Aa15, Vip3Aa16, Vip3Aa17, Vip3Aa18, Vip3Aa19, Vip3Aa20, Vip3Aa21, Vip3Aa22, Vip3Aa23, Vip3Aa24, Vip3Aa25, Vip3Aa26, Vip3Aa27, Vip3Aa28, Vip3Aa29, Vip3Aa30, Vip3Aa31, Vip3Aa32, Vip3Aa33, Vip3Aa34, Vip3Aa35, Vip3Aa36, Vip3Aa37, Vip3Aa38, Vip3Aa39, Vip3Aa40, Vip3Aa41, Vip3Aa42, Vip3Aa43, Vip3Aa44, Vip3Ab1, Vip3Ab2, Vip3Ac1, Vip3Ad1, Vip3Ad2, Vip3Ae1, Vip3Af1, Vip3Af2, Vip3Af3, Vip3Ag1, Vip3Ag2, Vip3Ag3, HM117633, Vip3Ag4, Vip3Ag5, Vip3Ah1, Vip3Ba1, Vip3Ba2, Vip3Bb1, Vip3Bb2 và Vip3Bb3.

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

Theo các phương án bổ sung, phương pháp tạo ra độc tố protein đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*) được đề xuất, phương pháp bao gồm: nuôi cấy tế bào chủ chuyển gen không phải của người mà chứa polynucleotit hoặc gen khảm hoặc phân tử axit nucleic hoặc vector tái tổ hợp theo sáng chế dưới các điều kiện trong đó vật chủ tạo ra độc tố protein đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*). Theo một số phương án, tế bào chủ chuyển gen không phải của người là tế bào của cây. Theo một phương án, tế bào của cây là tế bào cây ngô. Theo các phương án khác, các điều kiện mà tế bào của cây hoặc tế bào cây ngô phát triển bao gồm ánh sáng mặt trời tự nhiên. Theo các phương án khác, tế bào chủ chuyển gen

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

Theo các phương án khác, protein được tạo ra có hoạt tính diệt côn trùng kháng ít nhất một côn trùng bổ sung, trong đó côn trùng bổ sung được chọn từ nhóm bao gồm sâu xám (*Agrotis ipsilon*), sâu xanh mùa thu (*Spodoptera frugiperda*), sâu đục trái trên cây ngô (*Helicoverpa zea*), sâu đục thân mía (*Diatraea saccharalis*), sâu ăn lá đậu tím (*Anticarsia gemmatilis*), sâu đo trên đậu tương (*Chrysodeixis includens*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục trái đậu phương Tây (*Richia albicosta*), sâu chồi thuốc lá (*Heliothis virescens*), sâu đục thân ngô châu Á (*Ostrinia furuncalis*), sâu đục quả bông (*Helicoverpa armigera*), sâu đục thân sọc (*Chilos suppressalis*), sâu cú mèo (*Sesamia calamistis*) hoặc sâu cuốn lá nhỏ hại lúa (*Cnaphalocrocis medinalis*), và kết hợp bất kỳ của chúng.

Theo các phương án khác, gen khảm bao gồm bất kỳ trong số các SEQ ID NOs: 1-5. 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: 16-20.

Theo một số phương án, gen khảm bao gồm trình tự nucleotit mà là bộ ba mã được tối ưu hóa để biểu hiện ở cây. Theo các phương án khác, gen khảm bao gồm bất kỳ trong số các SEQ ID NOs: 6-15. 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 các phương án khác nữa, sáng chế đề xuất phương pháp tạo ra cây chuyển gen kháng loài gây hại (ví dụ, kháng côn trùng), bao gồm: đưa vào cây polynucleotit, gen khảm, vectơ tái tổ hợp, caset biểu hiện hoặc phân tử axit nucleic theo sáng chế chứa trình tự nucleotit mà mã hóa protein Cry theo sáng chế, trong đó trình tự nucleotit được biểu hiện trong cây trồng, nhờ đó tạo cho cây khả năng kháng ít nhất sâu đục thân ngô Châu Âu, và tạo ra cây chuyển gen kháng loài gây hại (ví dụ, kháng côn trùng). Theo một số phương án, cây chuyển gen kháng loài gây hại là kháng ít nhất sâu xám (*Agrotis ipsilon*) so với cây kiểm soát không có polynucleotit, gen khảm, vectơ tái tổ hợp, caset biểu hiện hoặc phân tử axit nucleic theo sáng chế. Theo một số phương án, việc đưa vào đạt được bằng cách biến nạp cây. Theo các phương án khác, việc đưa vào đạt được bằng cách lai cây thứ nhất chứa gen khảm, vectơ tái tổ hợp, caset biểu hiện hoặc phân tử axit nucleic theo sáng chế với cây thứ hai khác.

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

Theo các phương án khác nữa, phương pháp kiểm soát ít nhất côn trùng sâu xám (*Agrotis ipsilon*) được đề xuất, phương pháp bao gồm việc phân phối lên côn trùng lượng có hiệu quả của protein Cry theo sáng chế. Để có hiệu quả, protein Cry đầu tiên được côn trùng ăn vào bụng qua đường miệng. Tuy nhiên, protein Cry được phân phối đến côn trùng theo nhiều cách đã biết. Các cách để phân phối protein theo đường miệng vào côn trùng bao gồm, nhưng không giới hạn ở, cung cấp protein (1) trong cây chuyển gen, trong đó côn trùng ăn (ăn vào bụng) một hoặc nhiều phần của cây chuyển gen, nhờ đó ăn vào bụng polypeptit mà được biểu hiện trong cây chuyển gen; (2) trong (các) chế phẩm protein được phối chế mà có thể được áp dụng hoặc kết hợp vào, ví dụ, môi trường phát triển của côn trùng; (3) trong (các) chế phẩm protein mà có thể được áp dụng lên bề mặt, ví dụ, được phun, lên bề mặt của bộ phận của cây, mà sau đó được côn trùng ăn vào bụng do côn trùng ăn một hoặc nhiều bộ phận của cây đã được phun; (4) nền chất bả; (5) thông qua việc tiêm vào côn trùng; hoặc (6) hệ phân phối protein 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 chuyển gen.

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

phẩm theo sáng chế lên bề mặt cây có thể được thực hiện bằng cách sử dụng phương pháp bất kỳ đã biết đối với người có hiểu biết trung bình trong lĩnh vực để áp dụng hợp chất, chế phẩm, công thức và tương tự lên bề mặt cây. Một số ví dụ không giới hạn phân phối hoặc cho tiếp xúc cây hoặc bộ phận của nó bao gồm phun, rải bụi, rải, tung rắc, sương mù, tán nhỏ, truyền phát, ngâm, bơm vào đất, kết hợp với đất, làm ướt sũng (ví dụ, 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.

Theo một số phương án, sáng chế bao hàm phương pháp cung cấp cho nông dân phương tiện kiểm soát loài gây hại côn trùng bộ cánh vảy, phương pháp bao gồm việc cung cấp hoặc bán cho nông dân vật liệu cây như hạt, vật liệu cây chứa polynucleotit, gen khả năng, caset biểu hiện hoặc vector tái tổ hợp có khả năng biểu hiện protein Cry theo sáng chế, như được mô tả ở trên.

Các phương án theo sáng chế có thể được hiểu rõ hơn bằng cách tham khảo đến các ví dụ sau đây. Phân mô tả trên và sau đây của các phương án theo sáng chế và các phương án khác nhau không nhằm giới hạn phạm vi bảo hộ, mà chỉ để minh họa cho yêu cầu bảo hộ. Do đó, cần hiểu rằng các yêu cầu bảo hộ không giới hạn ở các chi tiết cụ thể trong các ví dụ này. Người có hiểu biết trong lĩnh vực hiểu rằng các phương án khác của sáng chế có thể được thực hiện mà không vượt ra khỏi phạm vi và tinh thần của bản mô tả này, phạm vi của sáng chế được xác định trong yêu cầu bảo hộ kèm theo.

Ví dụ thực hiện sáng chế

Ví dụ 1. Xác định các chủng Bt hoạt tính

Các phân lập *Bacillus thuringiensis* được nuôi cấy từ các bào tử có mặt trong sự gom hiện tại và được giữ trên các tấm aga T3 + penixillin. Mỗi phân lập được nuôi lớn ưa khí trong các khối sâu 24 giếng trong khoảng 10 ngày ở 28°C đến khi hình thành bào tử, mà được kiểm tra bằng cách nhuộm màu bằng thuốc nhuộm Coomassie blue/axit axetic và quan sát bằng kính hiển vi. Sau khi hình thành bào tử cả các phần tan và không tan được thử hoạt tính kháng các loài bộ cánh vảy được mong muốn. Các đoạn được thử trong sinh thử nghiệm nhiễm xạ bề mặt, khi các đoạn được phủ lên

khẩu phần nhân tạo đa loài. Mỗi phân lập được sàng lọc kháng ít nhất bốn loài bộ cánh vảy, bao gồm *Helicoverpa zea* (sâu đục trái trên cây ngô), *Agrotis ipsilon* (sâu xám), *Ostrinia nubilalis* (sâu đục thân ngô Châu Âu), và *Spodoptera frugiperda* (sâu xanh mùa thu) với kích thước mẫu bằng 12 ấu trùng sơ sinh. Thời gian của mỗi thử nghiệm là khoảng 7 ngày ở nhiệt độ phòng; các tấm được chấm điểm đối với tỷ lệ chết cũng như ức chế tăng trưởng ấu trùng. Tỷ lệ chết quan sát được ở mức tăng bằng 30% so với đối chứng âm được coi là có hoạt tính. Dựa trên thử nghiệm côn trùng ban đầu, các chủng C0588, C0633, C0680 và M0262 được chọn để phân tích thêm.

Ví dụ 2: Phân lập và sắp trình tự các gen Bt

Cấu trúc thư viện gen Fosmit: Đối với một số chủng Bt mà được xác định trong ví dụ 1, các gen mã hóa các protein hoạt tính giả định được phân lập sử dụng phương pháp thư viện fosmit được mô tả trong Park và cộng sự. (FEMS Microbiol. Lett. 284:28-34 (2008). Thư viện fosmit được cấu trúc bằng cách sử dụng bộ kit sản xuất thư viện Fosmid CopyControl™ (Epicentre, Madison, WI) theo nguyên tắc của nhà sản xuất. Tóm lại, ADN tinh chế từ mỗi chủng Bt (khoảng 0,5µg) được xử lý bằng enzym để sửa đầu các đầu cụt, và sau đó được buộc vào vector fosmit pCC1FOS (Epicentre). Sau khi gói trong ống nghiệm vào thể thực khuẩn lambda và lây nhiễm *Escherichia coli* (*E. coli*) EPI1300-T1®, các tế bào vi khuẩn được trải lên Luria-Bertani (LB) mà chứa 12,5µg/ml chloramphenicol. Các tấm được ủ ở khoảng 37°C trong 24h trước khi chọn khuẩn lạc. Các khuẩn lạc *E. coli* được chuyển nhiễm được chuyển sang tấm 96 giếng mà chứa 150µl môi trường LB chứa chloramphenicol và được ủ ở 37°C trong 24h.

Sàng lọc lai khuẩn lạc: Thư viện fosmit được trải ra với mật độ bằng 300 cfu mỗi tấm 100 x15mm L-aga cộng thêm 15µg/ml chloramphenicol. Tổng số 3000 fosmit được trải ra. Lai màng lọc được thực hiện bằng cách sử dụng vòng lọc Immobilon-Ny+ 87 mm (EMD Millipore, Billerica, MA). Việc lấy khuẩn lạc ra được hoàn thành như sau: màng lọc được đặt lên các tấm trong khoảng 5 phút, sau đó sử dụng kẹp, màng lọc được lấy ra từ bề mặt aga và đặt khuẩn lạc úp lên giấy lọc Whatman được ngâm 0,5M NaOH trong 5 phút. Sau đó màng lọc khuẩn lạc được đặt lên giấy lọc Whatman được ngâm trong 2X SSC trong 5 phút. ADN được cố định lên màng với bộ UV Stratalinker® ở 2000 x100 µJ (Stratagene, Inc., La Jolla, CA). Sau đó màng lọc được phơi khô trong không khí trên giấy lọc Whatman. Màng lọc được lai trước và

được lai trong 250mM NaPO₄, pH 7,0, 7% SDS, 1% BSA ở 65°C như được mô tả bởi nhà cung cấp. Màng lọc lai được rửa trong 2X SSC, 0,5% SDS trong 30 phút ở 65°C, sau đó 0,2X SSC, 0,2% SDS trong 30 phút ở 65°C. Màng lọc được phơi nhiễm ra màng tia X (Kodak® BIOMAX® XAR, Fisher Scientific, Pittsburgh, PA) qua đêm với việc sàng mạnh ở -80°C. Khuẩn lạc dương được vá vào aga L cộng thêm 15 µg/ml chloramphenicol.

Mẫu dò lai : Đoạn mồi PCR được thiết kế để khuếch đại 720bp đoạn gen giống cry9B từ ADN bộ gen của chủng Bt được thiết kế C0633. Cặp đoạn mồi gồm đoạn mồi xuôi được thiết kế OAR2613a có trình tự AAACATGAACCGAAATAATCAAAATG (SEQ ID NO:26) và đoạn mồi ngược được thiết kế OAR2615a có trình tự ATCCGTCCCTTGTGCGTGTAATA (SEQ ID NO:27). Phản ứng PCR được chạy dưới các điều kiện chu kỳ sau: [94°C, 5 phút], 12x [94°C, 30 giây, 57°C đến 51°C, nhỏ giọt 0,5°C mỗi chu kỳ, 30 giây, 72°C 2,5 phút], và 35x[94°C, 30 giây, 52°C, 30 giây, 72°C, 2,5 phút]. Phản ứng chứa 1X đệm OneTaq® (New England Biolabs, Beverly, MA), 200µm dNTP, 80ng ADN, 2,5U OneTaq ADN polymeraza Taq®, 50ng mỗi đoạn mồi và nước cất vô trùng đến 50µl tổng phản ứng.

Amplicon thu được được tách trên 1% agarosa TAE gel chứa ethidium bromua. Amplicon được quan sát dưới ánh sáng UV và cắt gel. ADN được phân lập sử dụng bộ kit chiết gel như được mô tả bởi nhà cung cấp (Qiagen, Valencia, CA). Các mẫu dò được đánh dấu EasyTide (α-32P) dCTP 3000 Ci/mmol (Perkin Elmer, Waltham, MA) sử dụng hệ đánh dấu prime ngẫu nhiên Rediprime II (GE Healthcare, Pittsburgh, PA). Nucleotit không kết hợp được loại ra bằng cách sử dụng các cột sắc ký Micro Bio-Spin 30 (Biorad, Hercules, CA). Các mẫu dò được gia nhiệt ở 95°C trong 5 phút trước khi thêm dung dịch lai.

Sắp trình tự gen Bt: ADN điều chế cho 2-4 dòng độc lập được chuẩn bị theo chỉ dẫn của nhà sản xuất (Qiagen). Phản ứng sắp trình tự với đoạn mồi được thiết kế cho cả hai mạch của trình tự nucleotit dự báo được mong muốn được thực hiện bằng cách sử dụng bộ kit trình tự kết thúc BigDye™ (Applied Biosystem, Foster City, CA) theo hướng dẫn của nhà sản xuất. Sản phẩm phản ứng được điện di trên thiết bị sắp trình tự ABI373 hoặc ABI377. Tất cả dữ liệu sắp trình tự được phân tích sử dụng gói phần mềm Phred/Phrap/Consed (University of Washington) cho tỷ lệ sai bằng hoặc nhỏ hơn

10^{-4} ở mức trình tự liên ứng. Trình tự ghép bằng chương trình Sequencher™ (Phiên bản 4.7, Gene Codes Corp., Ann Arbor, MI). Mỗi gen được sắp trình tự đến 4X độ bao phủ.

Ví dụ 3. Tách dòng và tổng hợp gen Bt

Các cặp đoạn mồi đặc hiệu Cry9 được thiết kế để tạo điều kiện cho xác định và tách dòng các gen loại cry9. Các cặp đoạn mồi được thiết kế để lai với đầu 5' của gen loại cry9 bằng cách bổ sung vị trí giới hạn *PmeI* và với đầu 3' bằng cách bổ sung vị trí giới hạn *AscI*. Cặp đoạn mồi được sử dụng để khuếch đại đầu 5' bao gồm đoạn mồi xuôi có trình tự GTTTAAACATGAATCGAAATAATCAAAATG (SEQ ID NO:28) và đoạn mồi ngược có trình tự GGCGCGCCCTACTCTTGTGTTTCAATAAAA (SEQ ID NO:29). Cặp đoạn mồi được sử dụng để khuếch đại đầu 3' bao gồm đoạn mồi xuôi có trình tự GTTTAAACATGAATCAAAATAAACACGGA (SEQ ID NO:30) và đoạn mồi ngược có trình tự GGCGCGCCTTACTGTTGGGTTTCCATGAACT (SEQ ID NO:31). Các vị trí giới hạn được chèn vào được chèn trong các đoạn mồi tương ứng. Các phản ứng PCR được thực hiện sử dụng các điều kiện chu kỳ sau đây: [94°C, 5 phút] và 30x [94°C, 30 giây, 45°C, 30 giây, 72°C, 3,5 phút]. Phản ứng chứa 1X đệm OneTaq, 200µm dNTP, 80ng ADN, 2,5U OneTaq ADN polymeraza (New England Biolabs), 50ng mỗi đoạn mồi và nước cất vô trùng thành 50µl tổng phản ứng.

Amplicon thu được được tách dòng vào vector TOPO pCR 4.0 như được mô tả bởi nhà cung cấp (Life Technologies). Plasmid ADN được phân lập tự tiêu bằng *PmeI* và *AscI* như được mô tả bởi nhà cung cấp (New England Biolabs).

Đoạn *PmeI/AscI* được tách dòng vào vector hình thoi được thiết kế pCIB5634' được thiết kế để biểu hiện trong cả *E. coli* và *B. thuringiensis*. Vector pCIB5634' tự tiêu bằng *PmeI* và *AscI*. Vector tự tiêu và đoạn gen được tinh chế bằng cách chạy trên 1% gel trên cơ sở đệm agarosa Tris Axetat EDTA. Các đoạn bị cắt khỏi gel và làm sạch sử dụng bộ kit chiết gel QIAGEN như được mô tả bởi nhà cung cấp. Các đoạn được nối với nhau sử dụng bộ kit nối từ New England Biolabs như được mô tả bởi nhà cung cấp. Phản ứng nối được biến nạp vào các tế bào TOP10 (Life Technologies) như được mô tả bởi nhà cung cấp và được đặt trên L-agar chứa 100 mg/ml ampicillin. Plasmid ADN được phân lập từ khuẩn lạc đơn và dòng vô tính được xác định được sắp trình tự lại thành 2X độ bao phủ để xác nhận trình tự đúng.

Một số gen Bt mà được chọn để tạo ra tái tổ hợp nhưng không được tách dòng trực tiếp ngoài ADN bộ gen được bổ sung cho nhà cung cấp thứ ba để tổng hợp tổng thể gen. Các gen Bt được tổng hợp được tách dòng phụ vào vectơ hình thoi được mô tả để biểu hiện tiếp và thử thêm hoạt tính sinh học.

Ví dụ 4. Ghép và phân tích bộ gen

Một số gen Bt theo sáng chế được xác định bằng cách sử dụng nghiên cứu sắp trình tự tổng thể bộ gen. Tóm lại, ADN của *Bacillus* được cắt sử dụng thiết bị siêu âm Covaris S2 (Covaris, Inc., Woburn, MA) với chương trình ADN_400bp được cài đặt ở chu kỳ hoạt động: 10%; cường độ: 4; chu kỳ/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 Bioscience 1-57 (1-27 Brazil, 28-57 USA, UK và Switzerland) 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 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 75% hỗn hợp hạt (25% hạt AMPure cộng thêm 75% dung dịch NaCl/PEG TekNova cat # P4136). Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi thu được được chuyển vào giếng mới và 45µL 50% hỗn hợp hạt (50% hạt AMPure cộng thêm 50% dung dịch NaCl/PEG TekNova cat # P4136) được thêm. Hỗn hợp được khuấy và đặt trên giá từ. Dịch nổi thu được được loại ra và các hạt được rửa bằng 80% etanol. 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, 45s]; 12 x [98°C, 15s, 60°C, 30s, 72°C, 30s]; [72°C, 1 phút]. Mỗi phản ứng chứa: 5µL thư viện ADN library, 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 sắp trình tự các đầu bắt cặp (PE) (100 chu kỳ mỗi lần đọc; 12-24 thư viện mỗi lần) trên hệ sắp trình tự HiSeq 2500 sử dụng nguyên tắc sắp trình tự của nhà sản xuất (Illumina, Inc., San Diego, CA).

Dụng cụ phân tích máy tính *Bacillus* được phát triển để xác định và đặc trưng các gen có thể độc để ưu tiên dẫn đến thử nghiệm thêm trong phòng thí nghiệm.

Việc ghép và phân tích bộ gen cũng như phân tích thư viện bộ gen được mô tả ở trên dẫn đến xác định năm gen giống Cry9 trong chủng *Bacillus thuringiensis* có độc tố đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*). Các đặc điểm xác định của gen giống Cry9 và các protein được thể hiện trong bảng 1.

Bảng 1. Các gen giống Cry9 được xác định trong các chủng *Bacillus thuringiensis*.

Chủng	Tên Protein/Gen	Trọng lượng phân tử (kD)	Nucleotit SEQ ID NO:	Axit amin SEQ ID NO:
C0588	BT0009	132,2	1	19
C0633	BT0012	129,7	2	20
C0633	BT0013	129,6	3	21
C0680	BT0023	130,8	4	22
M0262	BT0067	130,3	7	25

Ví dụ 5. Tương đồng của BT0009, BT0012, BT0013, BT0023 và BT0067 với các protein Cry Bt đã biết

Nghiên cứu dữ liệu protein với trình tự axit amin của các protein theo sáng chế thấy rằng chúng tương đồng với protein diệt côn trùng. So sánh trình tự axit amin của các protein theo sáng chế với dữ liệu không dư (nr) được giữ bởi NCBI sử dụng thuật toán BLAST bộc lộ các protein sau đây là có khối lượng nhất của axit amin tương đồng với các trình tự theo sáng chế (bảng 2).

Bảng 2. Phần trăm tương đồng của protein Cry theo sáng chế với các protein Cry đã biết.

		Phần trăm tương đồng							
	Cry9Aa 1	Cry9Ba 1	Cry9B b1	Cry9Ca 1	Cry9Da 1	Cry9Db 1	Cry9Ea 1	CryFa 1	CryGa 1
0009	60	70	66	66	82	88	67	64	36
0012	97	57	53	56	56	56	56	52	62

0013	97	57	53	56	56	56	56	52	62
0023	59	70	65	70	73	71	86	67	36
0067	59	69	66	75	76	74	76	69	36

Ví dụ 6. Biểu hiện protein Bt trong tế bào chủ tái tổ hợp

Biểu hiện *Bacillus*. Gen được mong muốn được biểu hiện trong chủng *AcrySTALLIFEROUS Bacillus* không có hoạt tính với bộ cánh cứng hoặc bộ cánh vảy quan sát được qua vector biểu hiện pCIB5634' được mô tả ở trên, mà chứa vùng khởi động của protein Cry và chỉ báo kháng erythromycin thích hợp. Các cấu trúc được biến nạp vào chủng của vật chủ qua điện di và chọn lọc sau trên các tấm aga chứa erythromycin. Các chủng tái tổ hợp này được nuôi đến giai đoạn hình thành bào tử trong môi trường T3 ở 28°C trong 4-5 ngày. Các viên tế bào được thu và rửa đi rửa lại trước khi hòa tan trong đệm carbonat độ pH cao (50mM) chứa 2mM DTT.

Biểu hiện *E. coli*. Gen được mong muốn được biểu hiện trong nhiều chủng *E. coli* khác nhau sử dụng các vector pET28a hoặc pET29a (EMD Millipore). Các cấu trúc được biến nạp bằng điện di và chọn lọc sau trên các tấm aga chứa kanamycin. Các chủng tái tổ hợp này được nuôi và có thể biểu hiện cảm ứng sử dụng cảm ứng IPTG ở 28°C. Các tế bào được tạo huyền phù lại trong đệm carbonat độ pH cao (50mM) chứa 2mM DTT và sau đó phá vỡ bằng cách sử dụng thiết bị đồng nhất LV-1 vi dòng chảy.

Phân tích biểu hiện. Sau đó các hợp chất dung giải tế bào thu được (từ tế bào chủ bất kỳ) được phân loại thông qua li tâm và các mẫu được phân tích độ tinh khiết qua SDS-PAGE và điện di đồ (BioRad Experion). Tổng nồng độ protein được xác định qua thử nghiệm Bradford hoặc Thermo 660. Sau đó protein Cry tinh chế được thử trong sinh thử nghiệm.

Ví dụ 7. Hoạt tính của protein Cry trong sinh thử nghiệm

Các protein được tạo ra trong ví dụ 6 được thử kháng một hoặc nhiều trong số các loài gây hại côn trùng sau sử dụng phương pháp sinh thử nghiệm khẩu phần nhân tạo đã biết trong lĩnh vực: sâu xanh mùa thu (FAW; *Spodoptera frugiperda*), sâu đục trái trên cây ngô (CEW; *Helicoverpa zea*), sâu đục thân ngô Châu Âu (ECB; *Ostrinia nubilalis*), sâu xám (BCW; *Agrotis ipsilon*), sâu đục thân mía (SCB; *Diatraea saccharalis*), velvet đậu caterpillar (VBC; *Anticarsia gemmatilis*), sâu đo trên đậu tương (SBL; *Pseudoplusia includens*), sâu đục thân ngô Tây Nam (SWCB; *Diatraea grandiosella*), sâu đục trái đậu phương Tây (WBCW; *Striacosta albicosta*), sâu chồi thuốc lá (TBW; *Heliothis virescens*), sâu đục thân ngô châu Á (ACB;

Ostrinia furuncalis), sâu đục quả bông (CBW; *Helicoverpa armigera*), sâu đục thân sọc (SSB; *Chilosuppressalis*), sâu cú mèo (PSB; *Sesamia inferens*) và sâu cuốn lá lúa (RLF; *Cnaphalocrocis medinalis*).

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 tám 24 giếng. Sau khi bề mặt khẩu phần khô, ấu trùng của các loài côn trùng thử được thêm vào mỗi giếng. Các tấm được đóng kín và giữ ở điều kiện môi trường phòng thí nghiệm về nhiệt độ, ánh sáng và độ ẩm tương đối. Nhóm đối chứng dương bao gồm ấu trùng được phơi nhiễm ra chủng *Bacillus* kiểu dại có hoạt tính mạnh và phổ rộng. Nhóm đối chứng âm bao gồm ấu trùng được phơi nhiễm ra khẩu phần côn trùng được xử lý chỉ với dung dịch đệm và ấu trùng trên khẩu phần côn trùng không được xử lý; tức là chỉ khẩu phần. Tỷ lệ chết được đánh giá sau khoảng 120 giờ và cho điểm so với các đối chứng.

Các kết quả được thể hiện trong bảng 3, trong đó a "-" nghĩa là không có hoạt tính so với đối chứng, a "+/-" nghĩa là 0-10% hoạt tính so với đối chứng (loại này còn bao gồm 0% tỷ lệ chết với ức chế mạnh sự tăng trưởng ấu trùng), a "+" nghĩa là 10-25% hoạt tính so với đối chứng, a "++" nghĩa là 25-75% hoạt tính so với đối chứng, và a "+++" 75-100% hoạt tính so với đối chứng.

Bảng 3. Các kết quả sinh thử nghiệm với protein Cry.

Côn trùng	Các protein BT				
	0009	0012	0013	0023	0067
FAW	-	-	-	-	+/-
CEW	-	++	+	+	-
ECB	+++	+++	+++	+++	+++
BCW	-	-	-	-	-
SCB	++		-	++	+
VBC	+++	+++			-
SBL	-	+++			-
SWCB	++			+++	+/-
WBCW	-	-			
TBW		+++			++
ACB					
CBW					
SSB					
PSB					
RLF					

Ví dụ 8. Số phận của protein Cry trong thử nghiệm chất lưu dạ dày mô phỏng

Một số protein Cry nhất định được biểu hiện trong cây và hạt từ cây này được bán hàng năm cho nông dân để dùng trong kiểm soát nhiều loài gây hại côn trùng khác nhau. Các sản phẩm diệt loài gây hại tự bảo vệ này được kiểm tra và đăng ký bởi các tổ chức điều phối khác nhau bao gồm, ví dụ, Tổ chức bảo vệ môi trường Mỹ (EPA).

Phơi nhiễm khẩu phần là con đường chính mà con người có thể phơi nhiễm ra protein Cry được biểu hiện trong cây chuyển gen. Tính độc cấp đi qua đường miệng của động vật có vú và tiêu hóa protein là các điểm cuối để đánh giá nguy cơ sức khỏe con người của EPA. Bằng chứng khoa học khác về độ an toàn của protein Cry là chúng được thể hiện thoái biến nhanh trong ống nghiệm bằng cách sử dụng các chất lưu dạ dày mô phỏng. Kết quả của bảy thử nghiệm trong ống nghiệm được thực hiện với các protein Cry1, Cry2, và Cry3 đại diện cho thấy các protein bị thoái biến nhanh, thường trong vòng 30 giây. Các kết quả này hỗ trợ kết luận rộng hơn rằng các thành viên của nhóm protein Cry này (mà có chung sự tương đồng trình tự axit amin đáng kể) có khả năng thoái biến nhanh sau khi con người ăn phải. Vùng xem xét khác là protein Cry có thể gây phản ứng dị ứng hay không. Sự thoái biến nhanh trong ống nghiệm được chứng minh của protein Cry có thể tối thiểu hóa tiềm năng đối với sự kiện này. Bằng cách so sánh, chất gây dị ứng thức ăn thông thường vẫn còn trong mô hình dạ dày ruột trong ống nghiệm, trong khi đó protein thức ăn thông thường với lịch sử dị ứng suy thoái nhanh trong chất lưu dạ dày mô phỏng (Metcalf và cộng sự 1996).

Tìm hiểu bên trong tính gây dị ứng tiềm năng của protein có thể thu được qua việc phân tích khả năng tiêu hóa protein trong chất lưu dạ dày mô phỏng (SGF). Hầu như tất cả các protein Cry được biểu hiện trong cây chuyển gen mà được thử đến nay được tiêu hóa nhanh và do đó xác định là không gây dị ứng. Tuy nhiên, protein Cry9C được phát hiện trong sản phẩm ngô chuyển gen được gọi là Starlink được phát hiện ổn định một phần với SGF. Mặc dù Starlink Cry9C không độc đối với động vật, các đặc điểm về khả năng tiêu hóa một phần và tính ổn định xử lý một phần khiến nó khó cho EPA để loại bỏ hoàn toàn khả năng protein Cry9C Starlink có thể đóng vai trò là chất gây dị ứng thức ăn cuối cùng dẫn việc công ty mà phát triển Starlink phải thu hồi sản phẩm từ thị trường Mỹ.

Hiện nay, không có thử nghiệm cuối cùng để xác định tiềm năng gây dị ứng của các protein mới. Do đó, EPA sử dụng nghiên cứu bằng chứng khối lượng mà các yếu

tổ sau được xem xét: nguồn của đặc điểm; so sánh trình tự axit amin với các chất dị ứng đã biết; và đặc điểm sinh hóa của protein, bao gồm sự tiêu hóa trong ống nghiệm trong chất lưu dạ dày mô phỏng (SGF) và glycosyl hóa.

Thử nghiệm chất lưu dạ dày mô phỏng (SGF) đánh giá khả năng tiêu hóa trong ống nghiệm của protein thử ở các điều kiện được kiểm soát chặt chẽ đại diện bởi đường tiêu hóa của động vật có vú cấp cao. Tóm lại, protein Cry thử tạo ra vi khuẩn (ở nồng độ 0,5-5mg/ml) được phơi lộ ra enzym pepxin (từ niêm mạc dạ dày lợn, được hòa tan trong 2mg/ml NaCl, pH 1,2) ở tỷ lệ bằng 10 đơn vị hoạt tính pepxin / μ g protein thử qua khoảng thời gian bằng một giờ ở 37°C. Các mẫu được loại ở 1, 2, 5, 10, 30, và 60 phút và được dừng ngay bằng cách bổ sung đệm dừng được gia nhiệt trước (95°C – 2 phút) (65% 0,5M Natri bicarbonat pH 11, 35% đệm tải Trixin) hoàn lại ngay pepsin bất hoạt, và chuyển lại để gia nhiệt thêm 5 phút. Khi hoàn thành thử nghiệm, các mẫu và đối chứng điểm thời gian (chỉ protein thử, chỉ pepsin) được đánh giá bởi SDS-PAGE trên 10-20% Tris-Trixin gel (với các peptit nhìn thấy xuống còn 1 kDa) để theo dõi động học và mức tiêu hóa được thực hiện bởi pepsin.

Các kết quả của thử nghiệm SGF chứng minh rằng tất cả các protein Cry theo sáng chế đều thoái biến rất nhanh. Các kết quả này tạo ra bằng rằng các protein Cry theo sáng chế liên quan đến họ protein Cry 9, chúng khá khác nhau trong đáp ứng với thử nghiệm SGF so với các kết quả được công bố nhất định, ví dụ Cry9C trong Starlink, gợi ra sự khác nhau đáng kể về cấu trúc ở các vị trí phân cắt pepsin chính trong protein. Các kết quả này còn gợi ra rằng tiềm năng dị ứng của các protein Cry theo sáng chế là rất nhỏ.

Ví dụ 9. Tạo vector của gen để biểu hiện thực vật

Trước khi biểu hiện ở cây, polynucleotit tổng hợp bao gồm trình tự nucleotit mã hóa mỗi trong số các protein Cry Bt, BT-0009, BT-0012, BT-0013, BT-0023 và BT-0067 (SEQ ID NOs:6-10, tương ứng), và polynucleotit tổng hợp bao gồm trình tự nucleotit mã hóa mỗi trong số các protein Cry Bt đột biến, mBT-0009, mBT-0012, mBT-0013, mBT-0023 và mBT-0067 (SEQ ID NOs:21-25, tương ứng) được tổng hợp trên nền gen tự động hóa (Genscript, Inc., Piscataway, NJ). Ví dụ, caset biểu hiện thứ nhất được tạo bao gồm vùng khởi động ubiquitin của ngô (Ubi1) liên kết hoạt động với trình tự mã hóa protein Cry mà liên kết hoạt động với trình tự kết thúc Nos và caset biểu hiện thứ hai được tạo bao gồm vùng khởi động Ubi1 liên kết hoạt động với

trình tự mã hóa phosphomannosa isomerase (PMI) mà liên kết hoạt động với trình tự kết thúc Nos. Biểu hiện của PMI cho phép chọn lọc dương tính cây chuyển gen trên mannoza. Cả hai tổ hợp có thể biểu hiện tách dòng vào vector thích hợp cho biến nạp ngô trung gian *Agrobacterium*.

Ví dụ 11. Biểu hiện của protein Cry ở cây

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) bao gồm vector được mô tả trong ví dụ 12 phát triển trên môi trường rắn YEP (phần chiết nấm men (5 g/L), pepton (10g/L), NaCl (5g/L), 15g/l aga, pH 6,8) trong 2-4 ngày ở 28°C. Khoảng $0,8 \times 10^9$ *Agrobacterium* tế bào 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 chín từ dòng ngô lai cùng dòng được thực hiện từ các tai 8-12 ngày tuổi trong LS-inf lỏng + 100 µM As. Phôi được rửa một lần với môi trường gây nhiễm mới. Sau đó dung dịch *Agrobacterium* được thêm và phôi được tạo xoáy trong 30 giây và cho xử lý bằng vi khuẩn trong 5 phút. Sau đó chuyển các mầm trên mặt vảy nhỏ lên môi trường LSAs và nuôi cấy trong tối trong hai đến ba ngày. Sau đó, khoảng từ 20 đến 25 phôi mỗi tấm petri được chuyển vào môi trường LSDc được bổ sung xefotaxim (250mg/l) và bạc nitrat (1,6mg/l) và được nuôi cấy trong tối ở khoảng 28°C trong 10 ngày.

Các phôi chưa chín, tạo ra các thể sản phôi được chuyển vào môi trường LSD1M0.5S. Mề nuôi cấy được chọn trên môi trường này trong khoảng 6 tuần với bước nuôi cấy phụ ở khoảng 3 tuần. Thể sản sống sót được chuyển sang môi trường Reg1 được bổ sung mannoza. Sau khi nuôi cấy trong sáng (chế độ 16 giờ sáng/ 8 giờ tối), các mô xanh sau đó được chuyển sang môi trường Reg2 mà không có chất điều khiển tăng trưởng và được ủ trong khoảng 1-2 tuần. Cây non được chuyển sang hộp Magenta GA-7 (Magenta Corp, Chicago Ill.) chứa môi trường Reg3 và nuôi trong sáng. Sau khoảng 2-3 tuần, cây trồng được thử đối với sự có mặt của gen PMI và gen *cry* Bt bằng kỹ thuật 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 lá được cắt khỏi sự kiện sao chép đơn (giai đoạn V3-V4) và được cho tàn phá bằng ấu trùng sơ sinh, sau đó ủ ở nhiệt độ phòng trong 5 ngày. Kích thước mẫu cho sinh thử nghiệm đĩa lá thay đổi tùy theo loài côn trùng được thử (sâu đục thân ngô Châu Âu (ECB), n=10; sâu đục trái trên cây ngô (CEW), n=3, sâu xám (BCW), n=5). Thực hiện đọc để đánh giá tổn thương mô và tỷ lệ chết ở khoảng ngày 3 và ngày 5; các mẫu được đánh giá đối với tổn thương so với đối chứng âm sử dụng thước đo sau: "+" : <5% tổn thương mô, tất cả ấu trùng chết; "+/-": 5-20% tổn thương mô, tất cả ấu trùng chết; hoặc "-": >20% tổn thương mô, một số ấu trùng sống và/hoặc lũy tiến đến hình sao thứ 2.

Các kết quả của sinh thử nghiệm mô thực vật chuyển gen xác nhận rằng các protein Cry theo sáng chế khi được biểu hiện trong cây chuyển gen là độc đối với côn trùng đích. Ví dụ, mBT-0012 được biểu hiện trong ngô được biến nạp ổn định với gen khả năng theo sáng chế có hoạt tính kháng ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*).

YÊU CẦU BẢO HỘ

1. Gen khảm chứa vùng khởi động khác loại liên kết hoạt động với phân tử axit nucleic chứa trình tự nucleotit mã hóa protein độc đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự nucleotit chứa SEQ ID NO:2.
2. Gen khảm chứa vùng khởi động khác loại liên kết hoạt động với phân tử axit nucleic chứa trình tự nucleotit mã hóa protein độc đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự nucleotit mã hóa protein chứa SEQ ID NO:17 hoặc SEQ ID NO:22.
3. Gen khảm theo điểm 2, trong đó trình tự nucleotit mã hóa protein chứa SEQ ID NO:22.
4. Protein diệt côn trùng mà độc đối với ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), trong đó trình tự axit amin của protein này chứa SEQ ID NO:22.
5. Chế phẩm diệt côn trùng bao gồm protein theo điểm 4 và chất mang nông dụng.
6. Vectors tái tổ hợp chứa gen khảm theo điểm 2.
7. Cây chuyển gen chứa gen khảm theo điểm 2.
8. Hạt chuyển gen của cây chuyển gen theo điểm 7, trong đó hạt đã nêu chứa gen khảm.
9. Phương pháp tạo ra cây chuyển gen kháng côn trùng, phương pháp này bao gồm: đưa vào cây gen khảm theo điểm 2, trong đó protein diệt côn trùng được biểu hiện trong cây, nhờ đó tạo cho cây khả năng kháng ít nhất sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), và tạo ra cây chuyển gen kháng côn trùng.
10. Phương pháp kiểm soát ít nhất các côn trùng sâu đục thân ngô Châu Âu (*Ostrinia nubilalis*), phương pháp này bao gồm việc phân phối lên côn trùng lượng có hiệu quả của protein theo điểm 4.

DANH MỤC TRÌNH TỰ

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

<120> Gen khảm chứa vùng khởi động khác loại liên kết hoạt động với phân tử axit nucleic chứa trình tự nucleotit mã hóa protein diệt côn trùng, chế phẩm và phương pháp kiểm soát côn trùng

<130> 80670-VN-REG-ORG-P-1

<160> 31

<170> PatentIn phiên bản 3.5

<210> 1

<211> 3507

<212> ADN

<213> Bacillus thuringiensis

<400> 1

atgaatcgaa ataatcaaaa tgaatatgaa gttattgatg cccacattg tgggtgtccg
 60

gcagatgatg ttgtaaaata tcctttgaca gatgatccga atgctggatt gcaaaatatg
 120

aactataagg aatattttaca aacgtatggt ggagactata cagatcctct tattaatcct
 180

aacttatctg ttagtggaag agatgtaata caagttggaa ttaatatgt agggagatta
 240

ctaagctttt ttggattccc cttttctagt caatgggtta ctgtatatac ctatctttta
 300

aacagcttgt ggccggatga cgagaattct gtatgggacg cttttatgga gagagtagaa
 360

gaacttattg atcaaaaaat ctcagaagca gtaaagggtta gggcattgga tgacctaaact
 420

ggattacaat ataattataa tttatatgta gaagcattag atgagtggct gaatagacca
 480

aatggcgcaa gggcatcctt agtttctcag cgatttaaca ttttagatag cctatttaca
 540

caatttatgc caagcttttg ctctggtcct ggaagtcaaa attatgcaac tatattactt
 600

ccagtatatg cacaagcagc aaaccttcat ttgttattat taaaagatgc agacatttat
660

ggagctagat gggggctgaa tcaaactcaa atagatcaat tccattctcg tcaacaaagc
720

cttactcaga cttatacaaa tcattgtggt actgcgtata atgatggatt agcgggaatta
780

agaggcacia gcgttgcgag ttggctcaaa tatcatcaat accgtaggga aatgacagta
840

acggcaatgg atttagtggc attattccca tactataatg ttagacaata tccaaatggg
900

gcaaattcac aacttacacg tgaggtatat acagatccaa tcgtatttaa tccgcctgag
960

cgtccaagtg gcgctttctg cgaaagtfff tatactatcc gagcggctcg agaacgttta
1020

actttttcgc aacttgaaca tgcaataatt cgtccgccgc gcttgtttga aaggtttcaa
1080

gcttttagga tttatacagg cgaggcgaga ctgaatgcaa atagtgtcc aatgaactat
1140

tggattggac attttataag aaatacacgt ttaggtgact caacaacaat tactacaaat
1200

tatggaacaa ccaataatcg tttactaac ttcagtatgc cttctgatgt ttatcaaate
1260

aattcaacct caagtaattt agccgctatt ttaggcactt tatatggggg tactagagca
1320

caattccatt ttggatcagg aagtttttcg acgtatgtcg gacaaaatag cgttcttcca
1380

caatgtcatc aaaactataa ttcaatagaa gaattaccaa accaaagcga tgaacctaca
1440

gttagaagtt atagccatag attatctcat atcacctctt ttaatttcaa tgtacagctt
1500

aataatcctt taatttctgc gggcaatatg cctgtatatg tgtggacaca tcgcagtgtg
1560

gaccttacta acaggatttc ttcagataga attactcaaa taccagtggg aaaggcatat
1620

gagctaagta gtgggtgtac tgtcgtgaaa ggtccaggat tcacaggagg agatgtaate
1680

cgaagaacaa atactggtgg attcggagca ataaggggtgt cggtcactgg accgctaaca
1740

caacgatatc gcataagggtt ccgttatgct tcgacaatag attttgattt ctttgtaaca
1800

cgtggaggaa ctactataaa taatttttaga tttacacgta caatgaacag gggacaggaa
1860

tcaagatatg aatcctatcg tactgtagag tttacaactc cttttaactt tacacaaagt
1920

caagatataa ttcgaacatc tatccaggga cttagtggaa atggggaagt ataccttgat
1980

agaattgaaa tcatccctgt aaatccaaca cgagaagcgg aagaggattt agaagcggcg
2040

aagaaagcgg tggcgagctt gtttacacgc acaagggacg gattacaagt aaatgtgaaa
2100

gattatcaag tcgatcaagc ggcaaattta gtgtcatgct tatcagatga acaatatggg
2160

tatgacaaaa agatgttatt ggaagcggta cgcgcggcaa aacgcctcag ccgagaacgt
2220

aacttacttc aggacccaga ttttaataca atcaatagta cagaagaaaa tggatggaaa
2280

gcaagtaacg gcgttactat tagtgagggc ggtccattct ataaaggccg tgcacttcag
2340

ctagcaagtg cagcagaaaa ttatccaaca tacatttatc aaaaagtaga tgcacggag
2400

ttaaaacctt atacacgata tagactagat gggttcgtga agagtagtca agatttagaa
2460

attgatctca ttcaccatca taaagtccat cttgtgaaaa atgtaccaga taatttagta
2520

tctgatactt acccagatga ttcttgtagt ggaatcaatc gatgtcagga acaacagatg
2580

gtaaatgcgc aactggaaac agagcatcat catccgatgg attgctgtga agcagctcaa
2640

acacatgagt tttcttcta tattgatata ggggatttaa attcgactgt agaccaggga
2700

atctgggtga tctttaaagt tcgaacaacc gatggttatg cgacgttagg aaatcttgaa
2760

ttggtagagg tcggaccgtt attgggtgaa cctctagaac gtgaacaaag agaaaatgcg
2820

aatggaatg cagagttagg aagaaaacgt gcagaaacag atcgcggtga tcaagatgcc
2880

aaacaatcca tcaatcattt atttgtggat tatcaagatc aacaattaaa tccagaaata
2940

gggatggcag atattatgga cgctcaaaat cttgtcgcat caatttcaga tgtatatagc
3000

gatgccgtac tgcaaattccc tggaattaac tatgagattt acacagagct gtccaatcgc
3060

ttacaacaag catcgatatc gtatacgtct cgaaatgcgg tgcaaaatgg ggactttaac
3120

aacgggctag atagctggaa tgcaacagcg ggtgcatcgg tacaacagga tggcaatacg
3180

catttcttag ttcttttctca ttgggatgca caagtttctc aacaatttag agtgcagccg
3240

aattgtaa atgtattacg tgtaacagca gagaaagtag gcggcggaga cggatacgtg
3300

actatccggg atgatgctca tcatacagaa acgcttacat ttaatgcatg tgattatgat
3360

ataaatggca cgtacgtgac tgataatacg tatctaaca aagaagtgggt attctattca
3420

catacagatc acatgtgggt agaggtaagt gaaacagaag gtgcatttca tatagatagt
3480

attgaattcg ttgaaacaga aaagtaa
3507

<210> 2
<211> 3471
<212> ADN
<213> *Bacillus thuringiensis*

<400> 2
atgaatcaaa ataaacacgg aattattggc gcttccaatt gtggttgtgc atctgatgat
60

gttgcgaaat atccttttagc caacaatcca tattcatctg ctttaaattt aaattcttgt
120

caaaatagta gtattctcaa ctggattaac ataataggcg atgcagcaaa agaagcagta
180

tctattggga caaccatagt ctctcttatac acagcacctt ctcttactgg attaatttca
 240
 atagtatatg accttatagg taaagtacta ggaggtagta gtggacaatc catatcagat
 300
 ttgtctatat gtgacttatt atctattatt gattttacggg taagtcagag tgtcttaaatt
 360
 gatgggattg cagattttta tggttctgta ctcttataca ggaactatct agaggctctg
 420
 gatagctgga ataagaatcc taattctgct tctgctgaag aactccgtac tcgtttttaga
 480
 atcgccgact cagaatttga tagaatttta acccgagggt cttaacgaa tgggtggctcg
 540
 ttagctagac aaaatgcca aatattatta ttaccttctt ttgcgagcgc tgcatttttc
 600
 catttattac tactaaggga tgctactaga tatggcaacta attgggggct atacaatgct
 660
 acacctttta taaattatca atcaaaaacta gtagagctta ttgaactata tactgattat
 720
 tgcgtacatt ggtataatcg aggtttcaac gaactaagac aacgaggcac tagtgctaca
 780
 gcttggttag aatttcatag atatcgtaga gagatgacat tgatgggtatt agatatagta
 840
 gcatcatttt caagtcttga tattactaat tacccaatag aaacagattt tcagttgagt
 900
 agggtcattt atacagatcc aattggtttt gtacatcgta gtagtcttag gggagaaagt
 960
 tggttttagct ttgttaatag agctaatttc tcagatttag aaaatgcaat acctaattcct
 1020
 agaccgtctt ggttttttaa taatatgatt atatctactg gttcacttac attgcgggtt
 1080
 agcccaagta ctgatagagc gagggtatgg tatggaagtc gagatcgaat ttccccctgct
 1140
 aattcacaat ttattactga actaatctct ggacaacata cgactgctac acaaactatt
 1200
 ttagggcgaa atatatattag agtagattct caagcttgta atttaaatga taccacatat
 1260

ggagtgaata gggcgggtatt ttatcatgat gcgagtgaag gttctcaaag atccgtgtac
1320

gaggggtata ttcgaacaac tgggatagat aaccctagag ttcaaaatat taacacttat
1380

ttacctggag aaaattcaga tatcccaact ccagaagact atactcatat attaagcaca
1440

acaataaatt taacaggagg acttagacaa gtagcatcta atcgccgttc atctttagta
1500

atgtatggtt ggacacataa aagtctggct cgtaacaata ccattaatcc agatagaatt
1560

acacagatac cattgacgaa ggttgatacc cgaggcacag gtgtttctta tgtgaatgat
1620

ccaggattta taggaggagc tctacttcaa aggactgacc atggttcgct tggagtattg
1680

agggtccaat ttccacttca cttaagacaa caatatcgca ttagagtccg ttatgcttct
1740

acaacaaata ttcgattgag tgtgaatggc agtttcggta ctatttctca aaatctccct
1800

agtacaatga gattaggaga ggatttaaga tacggatctt ttgctataag agagtttaat
1860

acttctatta gacccactgc aagtccggac caaattcgat tgacaataga accatctttt
1920

attagacaag aggtctatgt agatagaatt gagttcattc cagttaatcc gacgcgagag
1980

gcgaaagagg atctagaagc agcaaaaaaa gcggtggcga gcttgtttac acgcacaagg
2040

gacggattac aagtaaatgt gaaagattat caagtcgatc aagcggcaaa tttagtgtca
2100

tgcttatcag atgaacaata tgggtatgac aaaaagatgt tattggaagc ggtacgtgcg
2160

gcaaaacgac ttagccgaga acgcaactta cttcaggatc cagattttta tacaattaat
2220

agtacagaag aaaatggatg gaaagcaagt aacggcggtta ctattagtga gggcgggcca
2280

ttctataaag gccgtgcaat tcagctagca agcgcacgag aaaattatcc aacatacatt
2340

tatcaaaaaag tagatgcatc ggagttaaag ccatatacac gctatagact ggatggatto
2400

gtgaagagta gtcaagattt agaaattgat ctcatcacc atcataaagt ccatcttggt
2460

aaaaatgtac cagataattt agtacttgat acttaccag atgattcctg caacggaatt
2520

aaccgttggt atgaacagaa gatggtaa atgcaactgg aaacagaaca tcatcatccg
2580

atggattgct gtgaagcggc tcaaacacat gagttttctt cctatattaa tacaggcgat
2640

ctaaatgcaa gtgtagatca aggcatttgg gttgtattga aagttcgaac aacagatggt
2700

tatgcgacgc taggaaatct tgaattggta gaggtcggac cgttatcggg tgaatctcta
2760

gaacgtgaac aaagggataa tgcgaaatgg agtgcagagc taggaagaaa gcgtgcagaa
2820

acagagagag tatattatgc tgccaaacaa tccatcaatc atttatttgt ggattatcga
2880

gatcaacaat taaatccaca aatagggatg gcagatatta tggacgctca aaatcttgct
2940

gcatcaattt cagatgtata tagcgatgcc gtactgcaaa tccctggaat taactatgag
3000

atttacacag agctatccaa tcgcttacaa caagcatcgt atctgcatat gtctcgaaat
3060

gcgatgcaaa atggggactt taacagcggg ctagatagtt ggaatgcaac agcgggtgct
3120

acggtacaac aggatggcaa tacgcatttc ttagttcttt ctcatggga tgcacaagtt
3180

tctcaacaat ttagagtga gccgaattgt aaatatgtat tacgtgtaac agcagagaaa
3240

gtaggcggcg gagacggata cgtgacaatc cgggatgggt ctcatcatac agaaacgctt
3300

acatttaatg catgtgatta tgatataaat ggcacgtacg tgactgataa tacgtatcta
3360

acaaaagaag tggattacca tccggagaca caacatatgt gggtagaggt aagtgaaca
3420

gaagggtgttt tccatataga cagtgttgag ttcattggaaa cccaacagta a
3471

<210> 3
<211> 3471
<212> ADN
<213> *Bacillus thuringiensis*

<400> 3
atgaatcaaa ataaacacgg aattattggc gcttccaatt gtggttggtc atctgatgat
60

gttgcgaaat atccttttagc caacaatcca tattcatctg ctttaaattt aaattcttgt
120

caaaatagta gtattctcaa ctggattaac ataataggcg atgcagcaaa agaagcagta
180

tctattggga caaccatagt ctctcttatac acagcacctt ctcttactgg attaatttca
240

atagtatatg accttatagg taaagtacta ggaggtagta gtggacaatc catatcagat
300

ttgtctatat gtgacttatt atctattatt gatttacggg taagtcagag tgtctttaat
360

gatgggattg cagattttta tggttctgta ctcttataca ggaactattt agaggctctg
420

gatagctgga ataagaatcc taattctgct tctgctgaag aactccgtac tctttttaga
480

atcgccgact cagaatttga tagaatttta acccgagggt cttaacgaa tgggtggctcg
540

ttagctagac aaaatgccca aatattatta ttaccttctt ttgcgagcgc tgcatttttc
600

catttattac tactaaggga tgctactaga tatggcacta attgggggct atacaatgct
660

acacctttta taaattatca atcaaaacta gtagagctta ttgaactata tactgattat
720

tgcgtagatt ggtataatcg aggtttcaac gaactaagac aacgaggcac tagtgctaca
780

gcttggttag aatttcatag atatcgtaga gagatgacat tgatggtatt agatatagta
840

gcatcatttt caagtcttga tattactaat tacccaatag aaacagattt tcagttgagt
 900
 agggtcattt atacagatcc aattggtttt gtacatcgta gtagtcttag gggagaaagt
 960
 tggtttagct ttgttaatag agctaatttc tcagatttag aaaatgcaat acctaadcct
 1020
 agaccgtctt ggtttttaaa taatatgatt atatctactg gttcacttac attgccggtt
 1080
 agcccaagta ctgatagagc gagggatatg tatggaagtc gagatcgaat tccccctgct
 1140
 aattcacaat ttattactga actaatctct ggacaacata cgactgctac acaaactatt
 1200
 ttagggcgaa atatatttag agtagattct caagcttgta atttaaata taccacatat
 1260
 ggagtgaata gggcgggtatt ttatcatgat gcgagtgaag gttctcaaag atccgtgtac
 1320
 gaggggtata ttcgaacaac tgggatagat aaccctagag ttcaaaatat taacacttat
 1380
 ttacctggag aaaattcaga tatcccaact ccagaagact atactcatat attaagcaca
 1440
 acaataaatt taacaggagg acttagacaa gtagcatcta atcgccgttc atcttttagta
 1500
 atgtatggtt ggacacataa aagtctggct cgtaacaata ccattaatcc agatagaatt
 1560
 acacagatac cattgacgaa ggttgatacc cgaggcacag gtgtttctta tgtgaatgat
 1620
 ccaggattta taggaggagc tctacttcaa aggactgacc atggttcgct tggagtattg
 1680
 agggccaat ttccacttca cttaagacaa caatatcgca ttagagtccg ttatgcttct
 1740
 acaacaata ttcgattgag tgtgaatggc agtttcggta ctatttctca aaatctccct
 1800
 agtacaatga gattaggaga ggatttaaga tacggatctt ttgctataag agagtttaat
 1860
 acttctatta gacccactgc aagtccggac caaatcgat tgacaataga accatctttt
 1920

attagacaag aggtctatgt agatagaatt gagttcattc cagttaatcc gacgcgagag
1980

gcgaaagagg atctagaagc agcaaaaaaa gcggtggcga gcttggtttac acgcacaagg
2040

gacggattac aagtaaatgt gaaagattat caagtcgac aagcggcaaa tttagtgtca
2100

tgcttatcag atgaacaata tgggtatgac aaaaagatgt tattggaagc ggtacgtgcg
2160

gcaaacgac ttagccgaga acgcaactta cttcaggatc cagattttaa tacaattaat
2220

agtacagaag aaaatggatg gaaagcaagt aacggcgta ctattagtga gggcggggcca
2280

ttctataaag gccgtgcaat tcagctagca agcgcacgag aaaattatcc aacatacatt
2340

tatcaaaaag tagatgcac ggagttaaag ccatatacac gctatagact ggatggatcc
2400

gtgaagagta gtcaagattt agaaattgat ctcattcacc atcataaagt ccatcttgtg
2460

aaaaatgtac cagataattt agtacttgat acttaccag atgattcctg caacggaatt
2520

aaccgttgtg atgaacagaa gatggtaa at gcgcaactgg aaacaggaca tcatcatccg
2580

atggattgct gtgaagcggc tcaaacacat gagttttctt cctatattaa tacaggcgat
2640

ctaaatgcaa gtgtagatca aggcatttgg gttgtattga aagttcgaac aacagatggg
2700

tatgcgacgc taggaaatct tgaattggta gaggtcggac cgttatcggg tgaatctcta
2760

gaacgtgaac aaagggataa tgcgaaatgg agtcgagagc taggaagaaa gcgtgcagaa
2820

acagagagag tatattatgc tgccaaacaa tccatcaatc atttatttgt ggattatcaa
2880

gatcaacaat taaatccaca aatagggatg gcagatatta tggacgctca aaatcttgtc
2940

gcatcaattt cagatgtata tagcgatgcc gtactgcaaa tccttggaat taactatgag
3000

atttacacag agctatccaa tcgcttacaa caagcatcgt atctgcatac gtctcgaaat
3060

gcgatgcaaa atggggactt taacagcggg ctagatagtt ggaatgcaac agcgggtgct
3120

acggtacaac aggatggcaa tacgcatttc ttagttcttt ctcatggga tgcacaagtt
3180

tctcaacaat ttagagtgcg gccgaattgt aaatatgtat tacgtgtaac agcagagaaa
3240

gtaggcggcg gagacggata cgtgacaatc cgggatgggtg ctcatcatac agaaacgctt
3300

acatttaatg catgtgatta tgatataaat ggcacgtacg tgactgataa tacgtatcta
3360

acaaaagaag tggatttcca tccggagaca caacatatgt gggtagaggt aagtgaacaa
3420

gaaggtgttt tccatataga cagtgttgag ttcattggaaa cccaacagta a
3471

<210> 4

<211> 3459

<212> ADN

<213> Bacillus thuringiensis

<400> 4

atgaatcgaa ataattccaaa tgaatatgaa attattgatg cccctattg tgggtgtccg
60

tcagatgatg atgtgaggta tcctttggca agtgacccaa atgcagcgtt ccaaaatatg
120

aactataaag agtattttaca aacgtatgat ggagactaca caggttctct tatcaatcct
180

aacttatcta ttaatcctag agatgtacta caaacaggta ttaatatgtt gggaagaata
240

ctagggtttt taggtgttcc atttgcggtt caactagtta ctttctatac ctttctctta
300

aatcagttgt ggccaactaa tgataatgca gtatgggaag cttttatggc gcaaatagaa
360

gagctaatac atcaaaaaat atcggcgcaa gtagtaagga atgcactcga tgacttaact
420

ggattacacg attattatga ggagtattta gcagcattag aggagtggct ggaaagaccg
480

aacggagcaa gagctaactt agttacacag aggtttgaaa acctgcatac tgcatttgta
540

actagaatgc caagctttgg tacgggtcct ggtagtcaaa gagatgcggt agcgttggtg
600

acggtatatg cacaagcagc gaatttgcac ttgttattat taaaagatgc agaaatctat
660

ggggcaagat ggggacttca acaagggcaa attaacttat attttaatgc tcaacaagaa
720

cgtactcgaa tttataccaa tcattgcgtg gaaacatata atagaggatt agaagatgta
780

agaggaacaa atacagaaag ttggttaaata taccatcgat tccgtagaga gatgacatta
840

atggcaatgg atttagtggc cttattccca tactataatg tgcgacaata tccaaatggg
900

gcaaatccac agcttacacg tgaaatatat acggacccaa tcgtatataa tccaccagct
960

aatcagggaa tttgccgacg ttgggggaat aatccgtata atacattttc tgaacttgaa
1020

aatgctttta ttgccccgcc acatcttttt gatagggtga atagattaac tatttctaga
1080

aaccgatata cagctccaac aactaatagc tacctagact attggtcagg tcataactta
1140

caaagccagt atgcaaataa cccgacgaca tatgaaacta gttacggtca gattacctct
1200

aacacacggt tattcaatac gactaatgga gccaatgcaa tagattcaag ggcaagaaat
1260

tttggttaact tatacgctaa tttgtatggc gttagcagct tgaacatttt cccaacagg
1320

gtgatgagt aaatcacctc agcccctaata acgtgttggc aagaccttac tacaactgag
1380

gaactaccac tagtgaataa taattttaat cttttatctc atgttacttt cttacgcttc
1440

aatactactc aggggtggccc ccttgcaact gtagggtttg taccacata tgtgtggaca
1500

cgtcaagatg tagattttta taatataatt actcccaata gaattactca aataccagtg
1560

gtaaaggcat atgagctaag tagtggtgct actgtcgtga aagggtccagg attcacagga
1620

ggagatgtaa tccgaagaac aaatactggt ggattcggag caataagggt gtcggtcact
1680

ggaccgctaa cacaacgata tcgcataagg ttccgttatg cttcgacaat agattttgat
1740

ttctttgtaa cacgtggagg aactactata aataatttta gatttacacg tacaatgaac
1800

aggggacagg aatcaagata tgaatcctat cgtactgtag agtttacaac tccttttaac
1860

tttacacaaa gtcaagatat aattcgaaca tctatccagg gacttagtgg aaatggggaa
1920

gtataccttg atagaattga aatcatccct gtaaatacaa cacgagaagc ggaagaggat
1980

ttagaagcgg cgaagaaagc ggtggcgagc ttgtttacac gcacaaggga cggattacaa
2040

gtaaagtga cagattatca agtcgatcaa gcggcaaatt tagtgtcatg cttatcagat
2100

gaacaatatg cgcattgataa aaagatgtta ttggaagcgg tacgcgcggc aaaacgcctc
2160

agccgagaac gtaacttact tcaggaccca gattttaata caatcaatag tacagaagaa
2220

aatggatgga aagcaagtaa cggcggttact attagtgagg gcggtccatt ctataaaggc
2280

cgtgcacttc agctagcaag tgcacgagaa aattacccaa catacatcta tcaaaaagta
2340

gatgcatcgg agttaaaacc ttatacacga tatagactgg atgggttcgt gaagagtagt
2400

caagatttag aaattgatct cattcaccat cataaagtcc atcttgatga aaatgtacta
2460

gataatttag tatctgatac ttaccagat gattcttgta gtggaatcaa tcgatgtgag
2520

gaacaacaga tggtaaatgc gcaactggaa acagaacatc atcatccgat ggattgctgt
2580

gaagcagctc aaacacatga gttttcttcc tatattgata caggggattt aaattcgact
2640

gtagaccagg gaatctgggt gatctttaaa gttcgaacaa cagatgggta tgcgacgcta
2700

ggaaatcttg aattggtaga ggtcggaccg ttattgggtg aacctctaga acgtgaacaa
2760

agagaaaatg cgaaatggaa tgcagagtta ggaagaaaac gtgcagaaac agatcgcggtg
2820

tatcaagatg ccaaacaatc catcaatcat ttatttgttg attatcaaga tcaacaatta
2880

aatccacaaa tagggatggc agatattatg gacgctcaaa atcttgtcgc atcaatttca
2940

gatgtatata gcgatgcagt actgcaaadc cctggaatta actatgagat ttacacagag
3000

ctgtccaatc gcttacaaca agcatcgtat ctgtatacgt ctcgaaatgc ggtgcaaaat
3060

ggggacttta acaacgggct agatagctgg aatgcaacag cgggtgcatc ggtacaacag
3120

gatggcaata cgcatttctt agttctttct cattgggatg cacaagtttc ccaacaattt
3180

agagtgcagc cgaattgtaa atatgtatta cgtgtaacag cagagaaagt aggcggcgga
3240

gacggatacg tgactatccg ggatgggtgct catcatacag aaacgcttac atttaatgca
3300

tgtgattatg atataaatgg cacgtacgtg actgataata cgtatctaac aaaagaagtg
3360

atattctatt cacatacaga acacatgtgg gtagaggtaa atgaaacaga aggtgcattt
3420

catatagata gtattgaatt cgttgaaaca gaaaagtaa
3459

<210> 5

<211> 3468

<212> ADN

<213> *Bacillus thuringiensis*

<400> 5

atgaatcgaa ataatacaga tgaatatgaa attattgacg cttccacttg tggttgttcg
60

tcagatgatg ttgttcaata ccctttggca agagatccga atgctgtatt ccaaaatatg
120

cattataaaag attatttgca aacgtatgat ggagactata caggttctct tataaatcct
180

aacttatcta ttaatcctag agatgtactg caaactggaa ttaatattgt gggaagatta
240

ctaggatttc taggtgttcc atttgctggt cagttagtta ctttctatac ctttctttta
300

aatcaactgt ggccaacaaa tgataatgca gtatgggaag cttttatggc acaaatagaa
360

gagcttatta atcaaagaat atccgaagca gtagtaggga cagcagcgga tcatttaacg
420

ggattacacg ataattatga gttatatgta gaggcattgg aggaatggct ggaaagaccg
480

aatgctgcta gaactaatct actttttaat agatttacca ccctagatag tctttttaca
540

caatttatgc caagctttgg tactggacct ggaagtcaaa actacgcagt tccattactt
600

acagtatacg cacaagcagc gaaccttcat ttgttattat taaaggatgc tgaaatatat
660

ggagcaagat ggggactgaa ccaaaatcag attaactcat tccatacgcg ccaacaagaa
720

cgtactcaat attatacaaa tcattgcgta acgacgtata ataccggttt agatagatta
780

agaggcacia atactgaaag ttggttaa at tatcatcggt tccgtagaga gatgacatta
840

atggcaatgg atttagtagc gctattccca tattataatg tacgacaata tccaaatgga
900

gcaaactctc agcttacacg tgaaatatat acagatccaa ttgtatttaa tccaccagct
960

aatgtgggat tatgtagacg ttggggcaat aacccatata atagattttc tgaactagaa
1020

aacgctttta tccgcccgcg acatcttttt gatagattga ataccttaac aattagtaga
1080

aatagatttg acgttgggtc aaactttata gaaccgtggt ctggacatac gttacgccgt
1140

agttattcga ataattcgac agtatatgaa gatagttatg gccaaattac agccacaaga
1200

acaacaatta atctgccggc taatggaact ggccgagtag aatcaacagc agtagatttt
1260

cgtagcgcgc ttgtggggat atacggtggt aatagagctt cttttattcc aggtggtgtg
1320

tttagtggca cgactcagcc ttctactgga ggatgtagag atttgtatga ttcaagtgat
1380

gaattaccac cagacgaaag tactggaagt ttgcccata gactatctca tgttaccttt
1440

ttaagtttta caactaatca ggccggttcc atagccaatt caggacgct ccctacttat
1500

gtctggaccc atcgcgatgt ggactttaat aatacaatta accccaatag aattactcaa
1560

ataccagtgg taaaggcata tgagctaagt agtggtgcta ctgtcgtgaa aggtccagga
1620

ttcacaggag gagatgtaat ccgaagaaca aatattggtg ggttcggagc aataagggtg
1680

tcggtcactg gaccgctaac acaacgatat cgcataaggt tccgttatgc ttcgacaata
1740

gattttgatt tctttgtaac acgtggagga actactataa ataattttag atttacacgt
1800

acaatgaaca ggggacagga atcaagatat gaatcctatc gtactgtaga gtttacaact
1860

ccttttaact ttacacaaag tcaagatata attcgaacat ctatccaggg acttagtgga
1920

aatggggaag tataccttga tagaattgaa atcatccctg taaatccaac acgagaagcg
1980

gaagaggatc tagaagcagc gaagaaagcg gtggcgagct tgtttacacg tacaagggac
2040

ggattacaag taaatgtgac agattatcaa gtcgatcaag cggcaaattt agtgtcatgc
2100

ttatcagatg aacaatatgg gcatgacaaa aagatgttat tggaagcggg aagagcggca
2160

aaacgectca gccgagaacg caacttactt caagatccag attttaatac aatcaatagt
2220

acagaagaaa atggatggaa agcaagtaac ggcgttacta ttagcgaggg cgtccattc
2280

tataaaggcc gtgcgcttca gctagcaagc gcaagagaaa attacccaac atacatttat
2340

caaaaagtaa atgcatcaga gttaaagccg tatacacgtt atagactgga tgggttcgtg
2400

aagagtagtc aagattttaga aattgatctc attcaccatc ataaagtcca tctcgtgaaa
2460

aatgtaccag ataatttagt atccgatact tactcggatg gttcttgagc tggaatgaat
2520

cgatgtgagg aacaacagat ggtaaagcgc caactggaaa cagaacatca tcatccgatg
2580

gattgctgtg aagcggctca aacacatgag ttttcttcct atattaatac aggcgatcta
2640

aattcaagtg tagatcaagg catttggtt gtattgaaag ttcgaacaac cgatgggttat
2700

gcgacgctag gaaatcttga attggtagag gtcggaccgt tatcgggtga atctctagaa
2760

cgtgaacaaa gggataatgc gaaatggagt gcagagctag gaagaaagcg tgcagaaaca
2820

gatcgcgtgt atcaagatgc caacaatcc atcaatcatt tatttggtga ttatcaagat
2880

caacaattaa atccagaaat agggatggca gatattattg acgctcaaaa tcttgctgca
2940

tcaatttcag atgtgtatag cgatgcagta ctgcaaacc ctggaattaa ctatgagatt
3000

tacacagagc tatccaatcg cttacaacaa gcatcgtatc tgtatacgtc tcgaaatgcg
3060

gtgcaaaatg gggactttta cagcggctca gatagttgga atgcaacagg gggggctacg
3120

gtacaacagg atggcaatac gcatttctta gttctttctc attgggatgc acaagtttct
3180

caacaattta gagtgcagcc gaattgtaaa tatgtattac gtgtaacagc agagaaagta
3240

ggcggcggag acggatacgt gacaatccgg gatgggtgctc atcatacaga aaagcttaca
3300

tttaatgcat gtgattatga tataaatggc acgtacgtga ctgataatac gtatctaaca
3360

aaagaagtgg tattctattc acatacagaa cacatgtggg tagaggtaag tgaaacagaa
3420

ggtgcatttc atatagatag tattgaattc gttgaaacag aaaagtaa
3468

<210> 6

<211> 3471

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 6

atgaatcaga acaagcacgg catcattggg gcctcgaact gcggctgcgc ctccgacgat
60

gtggctaagt accctctcgc taacaaccct tactccagcg ccctgaacct caattcctgc
120

cagaactcgt ctatcctcaa ttggattaac atcattgggg acgccgcgaa ggaggctgtc
180

agcatcggca ccacgattgt ttactgatc acggccccgt ccctgacagg cctcatcage
240

attgtgtacg acctcatcgg gaaggctcta ggcggtcat ccggccagag catttcggac
300

ctgtcgatct gcgatctcct gtctatcatt gacctcaggg tgtctcagtc agtcctgaac
360

gacgggatcg ccgatttcaa tggctcagtg ctctgtacc gcaactacct ggaggcgctc
420

gactcctgga acaagaatcc caactccgct agcgctgagg agctgaggac caggttcagg
480

attgccgact ccgagttcga tcgcattcctc actcgcgggt cactgaccaa tggcggctcc
540

ctcgcccgcc agaacgctca gatcctcctg ctccctagct tcgcgtcggc tgccttcttc
600

cacctgctcc tgctcaggga cgctacgcgc tacggaacca actggggcct ctacaatgct
660

actcogttca taaactacca gtccaagctg gtcgagctga tcgagctgta caccgactac
720

tgcgttcaact ggtacaatag ggggttcaac gagctgaggc agaggggcac gagcgctaca
780

gcttggtctg agttccatcg ctaccgcagg gagatgacgc tcatggttct ggacatcgtg
840

gccagcttca gctcgctcga tattactaac taccctatcg agaccgactt ccagctgtcg
900

cgctgattt acaccgaccc gatcgggttc gtccaccgct cttcactgag gggcgagtct
960

tggttctcat tcgtcaatag ggcgaaacttc tcggacctcg agaattgctat cccgaacccc
1020

cggccatctt ggttcttgaa caatatgac atttccactg gcagcctgac cctccccgtt
1080

tcgccatcta cggatcgcg gaggggtgtgg tacggctcac gggaccgcat ctcaccggct
1140

aactcccagt tcattacaga gctgatctct ggccagcaca caactgctac acagactatt
1200

ctgggccgga atatcttccg cgtggacagc caggcctgca atctcaacga taccacgtac
1260

ggcgtcaaca gggctgtttt ctaccatgac gcctcggagg gctctcagcg ctcagtctac
1320

gagggataca tcaggacaac tggcatcgat aatcctcggg tgcagaatat taacacgtac
1380

ctcccgggcg agaacagcga catccctacg ccggaggatt acacacacat tctgtcgacc
1440

acaatcaacc tcaccggcgg gctgaggcag gtcgcttcta acaggcgctc cagcctcgtt
1500

atgtacggct ggactcataa gtcactggcg cgcaacaaca caatcaaccc tgataggatc
1560

actcagattc cgctcaccaa ggtggacact agggggaccg gcgtgtcgta cgtcaacgat
1620

cccggcttca tcggcggggc cctgctccag cgcaccgacc acggctccct cggggttctg
1680

cgggtgcagt tcccactgca tctccgccag cagtacagga ttcgggtcgg ctacgcgtcc
1740

acaactaaca tccgcctcag cgtgaatggg tcgttcggca cgatctccca gaacctgccc
1800

agcacaatga ggctcgggga ggacctgcgc tacggctcct tcgccattcg ggagttcaac
1860

acgagcatcc gccccacagc gtcgccagat cagattaggc tcactatcga gccaagcttc
1920

atcaggcagg aggtctacgt tgaccggatc gagttcattc ctgtcaaccc gacgagggag
1980

gctaaggagg atctggaggc tgctaagaag gctgtggcta gcctcttcac caggacgagg
2040

gacggcctgc aagtgaatgt caaggactac caggttgatc aggccgcgaa cctcgtgtcc
2100

tgctgagcg acgagcagta cggctacgat aagaagatgc tgctcgaggc ggtcagggct
2160

gctaagaggc tctccaggga gaggaacctg ctccaggacc ctgatttcaa tacgatcaac
2220

agcacagagg agaacgggtg gaaggcgtct aatggcgtga ccatttcaga gggcggccca
2280

ttctacaagg gcagggctat ccagctggct tcggctcggg agaactaccc cacgtacatc
2340

taccagaagg tcgacgcctc tgagctgaag ccatacacac gctaccgcct ggatggcttc
2400

gtgaagtcgt ctcaggacct cgagatcgat ctgattcacc atcacaaggt ccacctggtt
2460

aagaatgtgc ccgacaacct ggtcctcgac acctaccag acgattcctg caatggcatc
2520

aacaggtgcg acgagcagaa gatggtgaac gccagctcg agaccgagca tcaccatccg
2580

atggactgct gcgaggcggc tcagacgcac gagttctcat cctacattaa cacaggggac
2640

ctgaatgcc a gcgtggatca gggcatctgg gtggctcctca aagtcaggac cacggacggg
2700

tacgctacgc tgggcaacct ggagctgggtt gaagtggggc cactctcggg cgagtctctg
2760

gagagggagc agagggacaa cgccaagtgg agcgtgagc tgggcaggaa gagggctgag
2820

accgagaggg tctactacgc cgcgaagcag tcgatcaatc acctcttcgt ggactacagg
2880

gatcagcagc tgaacccccca gattggcatg gctgacatca tggatgccca gaacctcgtc
2940

gcgtcaatct ccgacgtgta ctccgatgcg gtccctgcaga tcccaggcat caactacgag
3000

atctacacgg agctgagcaa ccggctgcag caggcctcgt acctccacat gtctcgcaat
3060

gcgatgcaga acggggactt caattctggc ctggattcat ggaatgcgac tgctggggcc
3120

accgttcagc aggacggcaa caccacttc ctgggtgctct cccattggga tgcccaagtt
3180

agccagcagt tccgcgtgca gccgaactgc aagtacgtcc tgagggttac tgctgagaag
3240

gtcggcgggg gcgacggcta cgttaccatc cgcgatggcg ctcaccatac agagactctc
3300

accttcaacg cctgcgacta cgatatcaat ggcacgtacg tgacagacaa cacttacctg
3360

accaaggagg ttgtgttcca cccggagacc cagcatatgt gggtcgaggt tagcgagacg
3420

gagggcgtgt tccacataga cagcgttgag ttcattggaga cccagcagta g
3471

<210> 7

<211> 3471

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 7

atgaatcaga acaagcacgg catcattggg gcctcgaact gcggctgcgc ctccgacgat
60

gtggctaagt accctctcgc taacaaccct tactccagcg ccctgaacct caattcctgc
120

cagaactcgt ctatcctcaa ttggattaac atcattgggg acgccgcgaa ggaggctgtc
180

agcatcggca ccacgattgt ttcactgatc acggccccgt ccctgacagg cctcatcagc
240

attgtgtacg acctcatcgg gaaggctcta ggcggtcat ccggccagag catttcggac
300

ctgtcgatct gcgatctcct gtctatcatt gacctcaggg tgtctcagtc agtcctgaac
360

gacgggatcg ccgatttcaa tggctcagtg ctctgtacc gcaactacct ggaggcgctc
420

gactcctgga acaagaatcc caactccgct agcgctgagg agctgaggac cagggttcagg
480

attgccgact ccgagttcga tcgcatactc actcgcggt cactgaccaa tggcggctcc
540

ctgccccgcc agaacgctca gatcctcctg ctccctagct tcgctcggc tgccttcttc
600

cacctgctcc tgctcagga cgctacgcg tacggaacca actggggcct ctacaatgct
660

actccgttca taaactacca gtccaagctg gtcgagctga tcgagctgta caccgactac
720

tgcgttcact ggtacaatag ggggttcaac gagctgaggc agaggggcac gagcgctaca
780

gcttggctgg agttccatcg ctaccgcagg gagatgacgc tcatggttct ggacatcgtg
840

gccagtttca gctcgctcga tattactaac taccctatcg agaccgactt ccagctgtcg
900

cgcgatgattt acaccgaccc gatcgggttc gtccaccgct cttcactgag gggcgagtct
960

tggttctcat tcgtcaatag ggcgaacttc tcggacctcg agaatgctat cccgaacccc
1020

cggccatctt ggttcctgaa caatatgac atttccactg gcagcctgac cctccccgtt
1080

tcgccaatcta cggatcgcg gaggggtgtg tacggctcac gggaccgcat ctcaccggct
1140

aactcccagt tcattacaga gctgatctct ggccagcaca caactgctac acagactatt
1200

ctgggccgga atatcttccg cgtggacagc caggcctgca atctcaacga taccacgtac
1260

ggcgtcaaca gggctgtttt ctaccatgac gcctcggagg gctctcagcg ctcagtctac
1320

gagggataca tcaggacaac tggcatcgat aatcctcggg tgcagaatat taacacgtac
1380

ctcccgggcg agaacagcga catccctacg ccggaggatt acacacacat tctgtcgacc
1440

acaatcaacc tcaccggcgg gctgaggcag gtgcgttcta acaggcgctc cagcctcgtt
1500

atgtacggct ggactcataa gtcactggcg cgcaacaaca caatcaaccc tgataggatc
1560

actcagattc cgctcaccaa ggtggacact agggggaccg gcgtgtcgta cgtcaacgat
1620

cccggcttca tcggcggggc cctgctccag cgcaccgacc acggctccct cggggttctg
1680

cgggtgcagt tcccactgca tctccgccag cagtacagga ttcgggtccg ctacgcgtcc
1740

acaactaaca tccgcctcag cgtgaatggg tcgttcggca cgatctccca gaacctgccc
1800

agcacaatga ggctcgggga ggacctgcgc tacggctcct tcgccattcg ggagttcaac
1860

acgagcatcc gcccacagc gtcgccagat cagattaggc tcaactatcga gccaagcttc
1920

atcaggcagg aggtctacgt tgaccggatc gagttcattc ctgtcaaccc gacgagggag
1980

gctaaggagg atctggagge tgctaagaag gctgtggcta gcctcttcac caggacgagg
2040

gacggcctgc aagtgaatgt caaggactac caggttgatc aggccgcgaa cctcgtgtcc
2100

tgctgagcg acgagcagta cggctacgat aagaagatgc tgctcgaggc ggtcagggct
2160

gctaagaggc tctccaggga gaggaacctg ctccaggacc ctgatttcaa tacgatcaac
2220

agcacagagg agaacgggtg gaaggcgtct aatggcgtga ccatttcaga gggcggccca
2280

ttctacaagg gcagggctat ccagctggct tcggctcggg agaactaccc cacgtacatc
2340

taccagaagg tcgacgcctc tgagctgaag ccatacacac gctaccgcct ggatggcttc
2400

gtgaagtcgt ctcaggacct cgagatcgat ctgattcacc atcacaaggt ccacctggtt
2460

aagaatgtgc ccgacaacct ggtcctcgac acctaccag acgattcctg caatggcatc
2520

aacaggtgcg acgagcagaa gatggtgaac gccagctcg agaccgagca tcaccatccg
2580

atggactgct gcgaggcggc tcagacgcac gagttctcat cctacattaa cacaggggac
2640

ctgaatgcc a gcgtggatca gggcatctgg gtggtcctca aagtcaggac cacggacggg
2700

tacgtacgc tgggcaacct ggagctggtt gaagtggggc cactctcggg cgagtctctg
2760

gagagggagc agagggacaa cgccaagtgg agcgctgagc tgggcaggaa gagggctgag
2820

accgagagg tctactacgc cgcaagcag tcgatcaatc acctcttcgt ggactacagg
2880

gatcagcagc tgaaccccc a gattggcatg gctgacatca tggatgccca gaacctcgtc
2940

gcgtcaatct ccgacgtgta ctccgatgcg gtccctgcaga tcccaggcat caactacgag
3000

atctacacgg agctgagcaa ccggctgcag caggcctcgt acctccacat gtctcgcaat
3060

gcgatgcaga acggggactt caattctggc ctggattcat ggaatgcgac tgctggggcc
3120

accgttcagc aggacggcaa caccacttc ctggtgctct ccattggga tgcccaagtt
3180

agccagcagt tccgcgtgca gccgaactgc aagtacgtcc tgagggttac tgctgagaag
3240

gtcggcgggg gcgacggcta cgttaccatc cgcgatggcg ctcaccatac agagactctc
3300

accttcaacg cctgcgacta cgatatcaat ggcacgtacg tgacagacaa cacttacctg
3360

accaaggagg ttgtgttcca cccggagacc cagcatatgt gggtcgaggt tagcgagacg
3420

gagggcgtgt tccacataga cagcgttgag ttcattggaga cccagcagta g
3471

<210> 8
 <211> 3471
 <212> ADN
 <213> Trình tự Nhân tạo

<220>
 <223> Gen Cry tổng hợp

<400> 8
 atgaatcaga acaagcacgg catcattggg gcctcgaact gcggctgcgc ctccgacgat
 60
 gtggctaagt accctctcgc taacaacct tactccagcg ccctgaacct caattcctgc
 120
 cagaactcgt ctatcctcaa ttggattaac atcattgggg acgccgcgaa ggaggctgtc
 180
 agcatcggca ccacgattgt ttactgatc acggccccgt ccctgacagg cctcatcagc
 240
 attgtgtacg acctcatcgg gaaggctcta ggcgggtcat ccggccagag catttcggac
 300
 ctgtcgatct gcgatctcct gtctatcatt gacctcaggg tgtctcagtc agtcctgaac
 360
 gacgggatcg ccgatttcaa tggctcagtg ctccgtgacc gcaactacct ggaggcgctc
 420
 gactcctgga acaagaatcc caactccgct agcgctgagg agctgaggac caggttcagg
 480
 attgccgact ccgagttcga tcgcatcctc actcgcggt cactgaccaa tggcggtcc
 540
 ctgccccgcc agaacgctca gatcctcctg ctccctagct tcgctcggc tgccctcttc
 600
 cacctgctcc tgctcaggga cgctacgcgc tacggaacca actggggcct ctacaatgct
 660
 actccgttca taaactacca gtccaagctg gtcgagctga tcgagctgta caccgactac
 720
 tgcgttcact ggtacaatag ggggttcaac gagctgaggc agaggggcac gagcgctaca
 780
 gcttggtctg agttccatcg ctaccgcagg gagatgacgc tcatggttct ggacatcgtg
 840
 gccagcttca gctcgctcga tattactaac taccctatcg agaccgactt ccagctgtcg
 900

cgcgatgattt acaccgaccc gatcgggttc gtccaccgct cttcactgag gggcgagtct
960

tggtttcat tctcaatag ggcgaacttc tcggacctcg agaattgctat cccgaacccc
1020

cggccatctt ggttcctgaa caatatgac atttccactg gcagcctgac cctccccgtt
1080

tcgccatcta cggatcgcg gaggggtgtg tacgggtcac gggaccgcat ctcaccggct
1140

aactcccagt tcattacaga gctgatctct ggccagcaca caactgctac acagactatt
1200

ctgggccgga atatcttccg cgtggacagc caggcctgca atctcaacga taccacgtac
1260

ggcgtcaaca gggctgtttt ctacccatgac gcctcggagg gctctcagcg ctcagtctac
1320

gagggataca tcaggacaac tggcatcgat aatcctcggg tgcagaatat taacacgtac
1380

ctcccgggcg agaacagcga catccctacg ccggaggatt acacacacat tctgtcgacc
1440

acaatcaacc tcaccggcgg gctgaggcag gtcgcttcta acaggcgctc cagcctcggt
1500

atgtacggct ggactcataa gtcactggcg cgcaacaaca caatcaaccc tgataggatc
1560

actcagattc cgctcaccaa ggtggacact agggggaccg gcgtgtcgta cgtcaacgat
1620

cccggcttca tcggcggggc cctgctccag cgcaccgacc acggctccct cggggttctg
1680

cgggtgcagt tcccactgca tctccgccag cagtacagga ttccgggtccg ctacgcgtcc
1740

acaactaaca tccgcctcag cgtgaatggg tcgttcggca cgatctccca gaacctgccc
1800

agcacaatga ggctcgggga ggacctgcgc tacgggtcct tcgccattcg ggagttcaac
1860

acgagcatcc gccccacagc gtcgccagat cagattaggc tcactatcga gccaagcttc
1920

atcaggcagg aggtctacgt tgaccggatc gagttcattc ctgtcaaccc gacgagggag
1980

gctaaggagg atctggaggc tgctaagaag gctgtggcta gcctcttcac caggacgagg
2040

gacggcctgc aagtgaatgt caaggactac cagggttgatc aggccgcgaa cctcgtgtcc
2100

tgctgagcg acgagcagta cggctacgat aagaagatgc tgctcgaggc ggtcagggct
2160

gctaagaggc tctccaggga gaggaacctg ctccaggacc ctgatttcaa tacgatcaac
2220

agcacagagg agaacgggtg gaaggcgtct aatggcgtga ccatttcaga gggcggccca
2280

ttctacaagg gcagggctat ccagctggct tcggctcggg agaactaccc cacgtacatc
2340

taccagaagg tcgacgcctc tgagctgaag ccatacacac gctaccgcct ggatggcttc
2400

gtgaagtcgt ctcaggacct cgagatcgat ctgattcacc atcacaaggt ccacctggtt
2460

aagaatgtgc ccgacaacct ggtcctcgac acctaccag acgattcctg caatggcatc
2520

aacaggtgcg acgagcagaa gatggtgaac gccagctcg agaccggaca tcaccatccg
2580

atggactgct gcgaggcggc tcagacgcac gagttctcat cctacattaa cacaggggac
2640

ctgaatgcca gcgtggatca gggcatctgg gtggtcctca aagtcaggac cacggacggg
2700

tacgctacgc tgggcaacct ggagctggtt gaagtggggc cactctcggg cgagtctctg
2760

gagagggagc agagggacaa cgccaagtgg agcgctgagc tgggcaggaa gagggctgag
2820

accgagaggg tctactacgc cgcaagcag tcgatcaatc acctcttcgt ggactaccaa
2880

gatcagcagc tgaacccccca gattggcatg gctgacatca tggatgcca gaacctcgtc
2940

gcgtcaatct ccgacgtgta ctccgatgcg gtctgcaga tcccaggcat caactacgag
3000

atctacacgg agctgagcaa ccggtgcag caggcctcgt acctccacac gtctcgcaat
3060

gcgatgcaga acggggactt caattctggc ctggattcat ggaatgcgac tgctggggcc
3120

accgttcagc aggacggcaa caccacttc ctggtgctct cccattggga tgcccaagtt
3180

agccagcagt tccgctgca gccgaactgc aagtacgtcc tgagggttac tgctgagaag
3240

gtcggcgggg gcgacggcta cgttaccatc cgcgatggcg ctcaccatac agagactctc
3300

accttcaacg cctgcgacta cgatatcaat ggcacgtacg tgacagacaa cacttacctg
3360

accaaggagg ttgtgttcca cccggagacc cagcatatgt gggtcgaggt tagcgagacg
3420

gagggcgtgt tccacataga cagcgttgag ttcattggaga cccagcagta g
3471

<210> 9

<211> 3459

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 9

atgaacagga acaacccaaa cgagtacgag attattgatg ctccatactg cggctgcccc
60

tccgatgacg atgtgcgcta ccccctcgcc agcgaccga acgccgcgtt ccagaatatg
120

aactacaagg agtacctgca gacctacgac ggcgattaca cggggtcact gattaatcca
180

aacctctcca tcaatcctcg cgacgtcctc cagaccggaa tcaacattgt tggccgcac
240

ctcggcttcc tgggcgtgcc gttcgctggc cagctggta ccttctacac gttcctcctg
300

aaccagctct ggcctacgaa tgacaacgcg gtgtgggagg ccttcatggc gcagatcgag
360

gagctgattg atcagaagat ctccgctcag gtggtcagga acgccctcga cgatctgacc
420

ggcctccacg actactacga ggagtagctg gctgctctcg aggagtggct cgagaggcca
480

aacggcgctc gcgctaattt ggtcacgcag cgcttcgaga acctccatac cgcttcgtg
540

acgaggatgc cgagcttcgg gacaggcccc gggtcgcaga gggacgcggt tgctctcctg
600

actgtgtacg cgcaggcggc taacctgcac ctctgtctcc tgaaggatgc tgagatctac
660

ggcgctcggg gggggctgca gcagggccag atcaacctct acttcaatgc ccagcaggag
720

cgacaagga tctacactaa ccaactgcgtt gagacctaca atagggggct cgaggacgtg
780

cggggcacga acacagagag ctgggtgaat taccatcgct tccgcaggga gatgacgctc
840

atggcgatgg acctcgtggc gctcttcccc tactacaacg tccgccagta cccaatggc
900

gccaaccctc agctgacaag ggagatctac actgacccga tcgtgtacaa cccaccagct
960

aatcagggga tctgcaggcg ctggggcaac aatccctaca acaccttctc tgagctggag
1020

aatgcgttca ttaggccacc tcacctcttc gaccgcctga acaggctcac catctcccgg
1080

aatcgtaca cggctccaac cacgaactcg tacctggatt actggtctgg ccataccctc
1140

cagtcacagt acgcgaacaa tcctacaact tacgagacgt cctacggcca gattacaagc
1200

aacactcgcc tcttcaatac cacgaacggc gctaattgta tcgacagcag ggcccggaac
1260

ttcgggaatc tgtacgcgaa tctctacggc gtctccagcc tgaacatttt cccaaccggc
1320

gttatgtcag agatcacctc cgcccctaac acgtgctggc aggacctcac aactaccgag
1380

gagctgccac tgggtgaacaa taacttcaat ctctgtccc acgtcacatt cctgagggtc
1440

aacacgacac agggcgggcc actcgcgact gttggcttcg tgcccaccta cgtgtggacg
1500

cggcaggacg tcgatttcaa taacatcatt acaccaaacc gcatcactca gattcctgtt
 1560
 gtgaaggcgt acgagctgtc gtctggcgct acagtcgtta agggcccggg gttcactggc
 1620
 ggggacgtca ttaggaggac taacaccggc gggttcgggg ctatcagggt gagcgtcaca
 1680
 ggccccctca ctcagcgcta caggattcgg ttccgctacg cgtcgaccat cgacttcgat
 1740
 ttcttcgtca cgcgcggcgg gactaccatc aataacttcc ggttcacgcg caccatgaac
 1800
 aggggccagg agagccggta cgagtcgtac cgcaccgtgg agttcacgac accgttcaac
 1860
 ttcacacaga gccaggacat cattcgcact tctattcagg gcctgtcagg caacggggag
 1920
 gtttacctcg accggatcga gatcattcca gtgaacccaa ccaggagggc tgaggaggat
 1980
 ctcgaggctg cgaagaaggc tgtggcctcg ctgttcaacta ggaccggga cggcctccag
 2040
 gttaatgtga cggactacca ggtggatcag gctgccaacc tggctctgtg cctctctgac
 2100
 gagcagtacg ctcacgataa gaagatgctc ctggaggccg tccgcgctgc taagaggctg
 2160
 tcgagggaga ggaacctcct gcaggacca gatttcaaca ccattaattc tacggaggag
 2220
 aacgggtgga aggcgtcgaa tggcgtgacc atctctgagg gcgggccttt ctacaagggc
 2280
 cgcgctctcc agctggcttc agctaggag aactaccga catacatcta ccagaaggtc
 2340
 gagcctccg agctgaagcc ctacacgcgc taccgcctcg atggcttcgt gaagtcatcc
 2400
 caggacctgg agatcgatct cattcaccat cacaaggctc acctgggtta gaacgtgctg
 2460
 gacaatctcg tcagcgatac ctaccggac gattcttgct caggcatcaa tcgctgcgag
 2520
 gagcagcaga tgggtgaacgc ccagctcgag accgagcatc accatccgat ggactgctgc
 2580

gaggctgctc agacgcatga gttcagctcg tacatcgaca caggggatct gaactccact
 2640
 gtggaccagg gcatctgggt catcttcaag gtgcgcacta ccgatgggta cgcgaccctc
 2700
 ggcaacctcg agctggtcga ggttgggccc ctcctgggcg agccactgga gagggagcag
 2760
 agggagaacg ctaagtggaa tgctgagctg ggcaggaagc gggctgagac cgaccgcgtc
 2820
 taccaggatg ctaagcagtc catcaatcac ctgttcgtgg actaccagga tcagcagctc
 2880
 aaccacaga ttggcatggc tgacatcatg gatgcccaga acctggtcgc gtccatcagc
 2940
 gacgtctact ctgatgccgt tctccagatc cctggcatta attacgagat ctacacagag
 3000
 ctgtcaaacc ggctccagca ggcttcctac ctgtacacta gccgcaacgc ggtgcagaat
 3060
 ggggacttca ataacggcct cgattcctgg aatgctacag ctggggctag cgtccagcag
 3120
 gacggcaaca ctacttcct cgttctgtct cattgggatg cccaggtttc acagcagttc
 3180
 aggggtgcagc cgaactgcaa gtatgtgctg agggttaccg ctgagaaggt tggcgggggc
 3240
 gacggctacg tgacgatcag ggatggcgct caccatacgg agacactgac tttcaacgcc
 3300
 tgcgactacg atatcaatgg cacctacgtc acggacaaca catacctcac taaggaggtc
 3360
 atattctaca gccacaccga gcatatgtgg gtggagggtc acgagacgga gggcgccctc
 3420
 cacatagact cgattgagtt cgtggagacc gagaagtga
 3459

<210> 10
 <211> 3468
 <212> ADN
 <213> Trình tự Nhân tạo

 <220>
 <223> Gen Cry tổng hợp

<400> 10

atgaatagga ataaccagga cgagtaggag atcattgacg cttcgacttg cgggtgcagc
60

tcagacgatg ttgttcagta cccctcgct cgcgatccca atgccgtgtt ccagaacatg
120

cactacaagg actacctcca gacatacgac ggcgattaca ctgggtcgct cattaacca
180

aatctgtcta tcaatcctcg cgacgtcctg cagacgggaa tcaacattgt tgggaggctc
240

ctgggcttcc tcgggggttcc gttcgctggc cagctgggtga cattctacac tttcctcctg
300

aaccagctct ggcccaccaa cgacaatgcg gtctgggagg cttcatggc gcagatcgag
360

gagctgatta accagcgcat ctccgaggct gtggtcggca ccgctgctga tcacctcagc
420

gggctgcatg acaactacga gctttacgtg gaggtctctg aggagtggct ggagaggcca
480

aacgtgcta ggacgaacct cctgttcaat aggttcacca cgctcgactc cctgttcaca
540

cagttcatgc catcattcgg gactggcccc ggctcccaga actacgccgt ccctctcctg
600

accgtttacg cgcaggcggc taacctccac ctctgctcc tgaaggacgc tgagatctac
660

ggcgctcgct gggggctgaa ccagaatcag atcaacagct tccacacaag gcagcaggag
720

cggacccagt actacagaa tcattgcgtg acaacttaca acaccggcct cgataggctg
780

agggggacca atacggagtc ctggctcaac taccataggt tccgcaggga gatgacctg
840

atggcgatgg acctcgtggc gctgttcccg tactacaacg tccggcagta cccaaacggc
900

gctaattcac agctcaccgc cgagatctac acggacccaa tcgtcttcaa cccaccagct
960

aatgttggcc tgtgcaggcg ctgggggaac aatccttaca atcgcttcag cgagctggag
1020

aacgccttca tcaggccacc tcacctgttc gatcggctca atacactgac tatttcgagg
1080

aaccggttcg acgtcggcag caatttcacg gagccgtggt cgggccatac gctgaggagg
1140

tcctacagca acaattcaac agtgtacgag gattcctacg gccagattac cgctacgagg
1200

accacgatca acctccctgc caatggcaca gggcgggtgg agtctactgc tgtcgacttc
1260

cgctcagccc tggttggcat ctacggggtg aacagggcga gtttcatccc gggcggggtc
1320

ttcagcggca caactcagcc atcgaccggc gggcgcaggg acctctacga ttccagcgac
1380

gagctgccac cagacgagag cacgggctcg ttcgctcaca ggctctctca tgtcacattc
1440

ctgtcattca ccacgaacca ggctgggctcg atcgccaatt ctgggagggt tccgacatac
1500

gtgtggactc accgggacgt ggatttcaac aataccatca acccaaatcg catcacgcag
1560

attcctgttg tgaaggcgta cgagctgtcc tccggcgcta ccgtcgtaa gggcccgggg
1620

ttcacgggcg gggacgtgat tcgcaggaca aacatcggcg gggtcggcgc tatcagggtc
1680

tccgttaccg gccccctgac gcagcgctac aggattcggg tccgctacgc ctccactatc
1740

gacttcgatt tcttcgtcac ccggggcggg acaactatca acaatttccg cttcacaagg
1800

actatgaaca ggggccagga gtcacggtag gagtcctacc gcaccgtgga gttcaccacg
1860

cccttcaact tcacgcagtc ccaggacatc attaggacat ccattcaggg cctcagcggc
1920

aacggggagg tttacctgga tcgcatcgag atcattccag tgaaccctac gaggagggt
1980

gaggaggacc tcgaggctgc gaagaaggct gtggctagcc tcttcaccag gacgagggat
2040

ggcctgcaag tgaatgtcac cgattaccag gtggaccagg ctgccaacct cgtctcttgc
2100

ctgtcagatg agcagtacg ccacgacaag aagatgctcc tggaggccgt ccgcgctgct
2160

aagaggctgt ccagggagag gaacctcctg caggacccag atttcaacac aatcaatagc
2220

actgaggaga atggctggaa ggcgtccaac ggggtgacca tcagcgaggg cgggcctttc
2280

tacaagggcc gcgctctcca gctggcttcg gctagggaga attacccaac ttacatctac
2340

cagaaggtca acgcctctga gctgaagcct tacacgcgt accgcctgga tggcttcgtg
2400

aagtcatccc aggatctcga gatcgacctg attcaccatc acaaggtgca cctcgtcaag
2460

aacgttccgg acaatctggc cagcgatacg tactcggacg gctcgtgctc tgggatgaat
2520

cgctgcgagg agcagcagat ggtgaacgcg cagctcgaga cagagcatca ccatccgatg
2580

gactgctgcg aggccgcgca gactcatgag ttcagctcgt acattaatac cggcgatctg
2640

aactcttcag tcgaccaggg catctgggtg gtcctcaagg ttcgcacaac tgatggctac
2700

gccacgctgg ggaacctcga gctggttgaa gtgggcccac tctccgggga gagcctggag
2760

agggagcaga gggacaacgc gaagtgggtcc gctgagctgg gcaggaagcg ggctgagacc
2820

gatagggttt accaggacgc taagcagtcc atcaatcacc tcttcgtgga ttaccaggac
2880

cagcagctga acccgagat tggcatggct gatatcattg acgccagaa cctcgtcgcg
2940

tcaatctccg atgtctacag cgacgccgtt ctgcagatcc ccggcattaa ttacgagatc
3000

tacacggagc tgtcgaacag gctgcagcag gcctcgtacc tctacacatc tcggaacgcg
3060

gtgcagaatg gcgatttcaa ctctgggctg gactcatgga atgctaccgg cggggctacg
3120

gtgcagcagg atggcaacac ccacttcctc gtctgtccc attgggacgc tcaggtgagc
3180

cagcagttcc gggccagcc gaactgcaag tatgtgctga gggtcactgc tgagaaggtt
3240

ggcgggggcg atggctacgt gaccatcagg gacggggctc accatacaga gaagctcact
3300

ttcaacgcct gcgactacga tatcaatggc acctacgtca cggacaacac atacctaact
3360

aaggaggttg tgttctactc gcacactgag catatgtggg tcgaggtttc tgagaccgag
3420

ggcgcttcc acatagactc aattgagttc gtggagaccg agaagtga
3468

<210> 11

<211> 3507

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 11

atgaatagga ataaccagaa cgagtacgag gtcattgatg ctccccattg cggctgcccg
60

gcggacgatg ttgtgaagta cccactgacc gacgatccga acgccggcct gcagaacatg
120

aattacaagg agtacctcca gacctacggc ggggactaca ccgacccgct gatcaaccct
180

aatctctcgg tttctggcaa ggacgtgatt caggtcggaa tcaacattgt ggggcgcctc
240

ctgtccttct tcgggttccc attctccagc cagtgggtga cagtctacac ttacctctg
300

aattcactgt ggcctgacga tgagaactcc gtttgggacg cttcatgga gcgggtggag
360

gagctgatcg atcagaagat tagcgaggct gtcaagggcc gcgcgctcga cgatctgacc
420

gggctccagt acaactacaa cctctacgtg gaggcgctgg acgagtggct caacaggcca
480

aatggcgctc gggcctctct cgtgtcacag aggttcaaca tcctggatto cctcttcacc
540

cagttcatgc catccttcgg gagcggcccc gggagccaga attacgcgac gatcctcctg
600

cctgtgtacg ctcaggccgc gaacctgcac ctctgtctcc tgaaggacgc tgatatctac
660

ggcgccaggt gggggctcaa ccagaccag attgaccagt tccactctag gcagcagtca
720

ctcaccaga cgtacacaaa ccattgcgtg acagcctaca atgacggcct ggctgagctg
780

agggggactt cggtcgctc ttggctgaag taccaccagt accgcaggga gatgactgtc
840

accgcgatgg acctggttgc tctcttccca tactacaacg tgcgccagta cccgaacggc
900

gcgaatcccc agctcaccag ggaggtgtac acggacccaa tcgtcttcaa cccaccagag
960

aggccttctg gcgctttctg cgagtcattc tacaccatca gggctgctag ggagaggctg
1020

acgttctccc agctcgagca cgcgatcatt cgcccaccta ggctgttcga gcggttccag
1080

gctctcggca tctacacagg ggaggccgc ctcaacgcta attccgcccc gatgaattac
1140

tggatcggcc atttcattag gaacacgagg ctgggggaca gcaccacgat cacaactaat
1200

tacggcacca cgaacaatag gctcaccaac ttctcaatgc cctccgacgt ctaccagatc
1260

aactccacgt cgtctaattc cgcggtatt ctgggcacac tctacggggt gactcgcgcc
1320

cagttccact tcggctcagg gtccttcagc acctacgtgg gccagaacag cgtcctgccg
1380

cagtgccatc agaactacaa tagcatcgag gagctgcaa accagtcgga cgagcctacc
1440

gttcgctcgt actctcacag gctgtcccat atcacgagct tcaacttcaa tgtgcagctg
1500

aacaatccgc tcatttctgc gggcaacatg cccgtttacg tgtggacca caggtcagtc
1560

gacctgacga accggatctc atccgatcg atcaccaga ttccggtggt caaggcttac
1620

gagctgtcct ccggcgctac agttgtgaag ggccccgggt tcaactggcg ggacgtgatc
1680

aggaggacga acacaggcgg gttcggcgcg attagggctt cggttaccgg gccgctcacg
1740

cagcgctacc gcatccgctt ccgctacgct tctaccattg acttcgattt cttcgtcacg
1800

cggggcgggg caactatcaa caatttcagg ttcactcgga ccatgaacag gggccaggag
1860

tcgcggtacg agtcttaccg caccgtggag ttcaccacgc ctttcaactt cacacagtcc
1920

caggacatca ttcgcacttc aatccagggc ctgtccggca atggggaggt ctacctcgac
1980

cgcattgaga tcattccggt taacccccacg agggaggctg aggaggatct ggaggctgcg
2040

aagaaggcgg tggctagcct gttcacgagg acacgggacg gcctccaagt gaatgtcaag
2100

gactaccagg tcgatcaggc tgccaacctg gttagctgcc tctcggacga gcagtacggc
2160

tacgataaga agatgctcct ggaggccgtg cgcgctgcta agaggctctc gaggggagagg
2220

aacctcctgc aggaccgga tttcaacaca atcaattcta ctgaggagaa cggctggaag
2280

gcttccaatg gggtgaccat tagcgagggc ggccatttct acaagggcag ggccctccag
2340

ctggcttcag ctcgggagaa ctacccaaca tacatctacc agaaggctcg cgcagcgag
2400

ctgaagcctt acactcgcta ccgcctcgat ggcttcgtga agtcttcaca ggacctggag
2460

atcgatctca ttcaccatca caagggtcac ctggtgaaga acgtcccga caatctcgtc
2520

tccgatacgt accccgacga ttcattgctc ggcatcaata ggtgccagga gcagcagatg
2580

gtgaacgccc agctggagac agagcatcac catccgatgg actgctgcga ggccgcgcag
2640

actcatgagt tctccagcta catcgacaca ggcgatctga acagcactgt ggaccagggg
2700

atctgggtca ttttcaagggt taggacaact gatggctacg cgaccctcgg gaacctcgag
2760

ctgggttgaag tgggcccact cctgggggag cctctggaga gggagcagag ggagaacgcg
2820

aagtggaatg ctgagctggg caggaagcgg gctgagaccg accgcgtcta ccaggatgcc
2880

aagcagagca tcaatcacct gttcggttgac taccaggatc agcagctcaa cccagagatc
2940

ggcatggccg acattatgga tgcgcagaaac ctgggtcgcta gcatctcgga cgtttactcg
3000

gatgccgtgc tccagatccc tggcattaat tacgagatct acaccgagct gagcaacagg
3060

ctccagcagg cgagctacct gtacacgtcg cggaacgctg tgcagaatgg cgacttcaac
3120

aatgggctcg attcttgga tgctaccgct ggcgcttcag tgcagcagga cgggaacacg
3180

cacttcctcg tcctgtcgca ttgggatgcc caggtttctc agcagttccg gytgcagccg
3240

aactgcaagt atgtgctcag ggtgacagct gagaaagtgg gcggggggcga cggctacgtc
3300

actatccgcg acgatgccc ccatactgag accctgacgt tcaacgcgtg cgactacgat
3360

atcaatggca cctacgtcac ggacaacaca tacatcacta aggaggtcgt tttctactcc
3420

cacacagatc atatgtgggt cgagggttagc gagactgagg gcgccttcca catcgactcg
3480

cttgagttcg tggagaccga gaagtga
3507

<210> 12
<211> 3471
<212> ADN
<213> Trình tự Nhân tạo

<220>
<223> Gen Cry tổng hợp

<400> 12
atgaatcaga acaagcacgg catcattggg gcctcgaact gcggctgcgc ctccgacgat
60

gtggctaagt accctctcgc taacaaccct tactccagcg ccctgaacct caattcctgc
120

cagaactcgt ctatcctcaa ttggattaac atcattgggg acgccgcgaa ggaggctgtc
180
agcatcggca ccacgattgt ttactgatc acggccccgt ccctgacagg cctcatcagc
240
attgtgtacg acctcatcgg gaagggtcat ggcgggtcat ccggccagag catttcggac
300
ctgtcgatct gcgatctcct gtctatcatt gacctcaggg tgtctcagtc agtcctgaac
360
gacgggatcg ccgatttcaa tggctcagtg ctctgtacc gcaactacct ggaggcgctc
420
gactcctgga acaagaatcc caactccgct agcgctgagg agctgaggac caggttcagg
480
attgccgact ccgagttcga tcgcatcctc actcgcggtc cactgaccaa tggcggctcc
540
ctcgcccgcc agaacgctca gatcctcctg ctccctagct tcgctcggc tgctttcttc
600
cacctgctcc tgctcaggga cgctacgcgc tacggaacca actggggcct ctacaatgct
660
actccgttca taaactacca gtccaagctg gtcgagctga tcgagctgta caccgactac
720
tgcgttcact ggtacaatag ggggttcaac gagctgaggc agaggggcac gagcgctaca
780
gcttggtcgg agttccatcg ctaccgcagg gagatgacgc tcattggttct ggacatcgtg
840
gccagcttca gctcgctcga tattactaac taccctatcg agaccgactt ccagctgtcg
900
cgcgtagattt acaccgaccg gatcgggttc gtccaccgct cttcactgag gggcgagtct
960
tggttctcat tcgtcaatag ggcgaacttc tcggacctcg agaattgctat cccgaacccc
1020
cggccatctt ggttcctgaa caatatgatc atttcactg gcagcctgac cctccccgtt
1080
tcgccatcta cggatcgcg gcgggtgtgg tacggctcac gggaccgat ctcaccggct
1140
aactcccagt tcattacaga gctgatctct ggccagcaca caactgctac acagactatt
1200

ctgggccgga atatcttccg cgtggacagc caggcctgca atctcaacga taccacgtac
1260

ggcgtcaaca gggctgtttt ctaccatgac gcctcggagg gctctcagcg ctcagtctac
1320

gagggataca tcaggacaac tggcatcgat aatcctcggg tgcagaatat taacacgtac
1380

ctccccggcg agaacagcga catccctacg ccggaggatt acacacacat tctgtcgacc
1440

acaatcaacc tcaccggcgg gctgaggcag gtcgcttcta acaggcgctc cagcctcggt
1500

atgtacggct ggactcataa gtcactggcg cgcaacaaca caatcaaccc tgataggatc
1560

actcagattc cgctcaccaa ggtggacact agggggaccg gcgtgtcgta cgtcaacgat
1620

cccggcttca tcggcggggc cctgctccag cgcaccgacc acggctccct cggggttctg
1680

cgggtgcagt tcccactgca tctccgccag cagtacagga ttcgggtccg ctacgcgtcc
1740

acaactaaca tccgcctcag cgtgaatggg tcgttcggca cgatctccca gaacctgcc
1800

agcacaatga ggctcgggga ggacctgcgc tacggctcct tcgccattcg ggagttcaac
1860

acgagcatcc gccccacagc gtcgccagat cagattaggc tcactatcga gccaaagcttc
1920

atcaggcagg aggtctacgt tgaccggatc gagttcattc ctgtcaaccc gacgagggag
1980

gctaaggagg atctggaggc tgctaagaag gctgtggcta gcctcttcac caggacgagg
2040

gacggcctgc aagtgaatgt caaggactac caggttgatc aggccgcgaa cctcgtgtcc
2100

tgcttgagcg acgagcagta cggctacgat aagaagatgc tgctcgaggc ggtcagggct
2160

gctaagaggc tctccaggga gaggaacctg ctccaggacc ctgatttcaa tacgatcaac
2220

agcacagagg agaacgggtg gaaggcgtct aatggcgtga ccatttcaga gggcggccca
2280

ttctacaagg gcagggctat ccagctggct tcggctcggg agaactaccc cacgtacatc
2340

taccagaagg tcgacgcctc tgagctgaag ccatacacac gctaccgcct ggatggcttc
2400

gtgaagtcgt ctcaggacct cgagatcgat ctgattcacc atcacaaggt ccacctggtt
2460

aagaatgtgc ccgacaacct ggtcctcgac acctaccag acgattcctg caatggcatc
2520

aacaggtgcg acgagcagaa gatggtgaac gccagctcg agaccgagca tcaccatccg
2580

atggactgct gcgaggcggc tcagacgcac gagttctcat cctacattaa cacaggggac
2640

ctgaatgcca gcgtggatca gggcatctgg gtggtcctca aagtcaggac cacggacggg
2700

tacgctacgc tgggcaacct ggagctgggt gaagtggggc cactctcggg cgagtctctg
2760

gagagggagc agagggacaa cgccaagtgg agcgctgagc tgggcaggaa gagggctgag
2820

accgagaggg tctactacgc cgccaagcag tcgatcaatc acctcttcgt ggactacagg
2880

gatcagcagc tgaaccccc a gattggcatg gctgacatca tggatgcccc gaacctcgtc
2940

gcgtcaatct ccgacgtgta ctccgatgcg gtcttcgaga tcccaggcat caactacgag
3000

atctacacgg agctgagcaa ccggtgcag caggcctcgt acctccacat gtctcgcaat
3060

gcgatgcaga acggggactt caattctggc ctggattcat ggaatgcgac tgctggggcc
3120

accgttcagc aggacggcaa caccacttc ctggtgctct ccattggga tgcccaagtt
3180

agccagcagt tccgcgtgca gccgaactgc aagtacgtcc tgagggttac tgctgagaag
3240

gtcggcgggg gcgacggcta cgttaccatc cgcgatggcg ctcaccatac agagactctc
3300

accttcaacg cctgcgacta cgatatcaat ggcacgtacg tgacagacaa cacttacctg
3360

accaaggagg ttgtgttcca cccggagacc cagcatatgt gggtcgaggt tagcgagacg
3420

gagggcgtgt tccacctgga cagcggttgag ttcattggaga cccagcagta g
3471

<210> 13

<211> 3471

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 13

atgaatcaga acaagcacgg catcattggg gcctcgaact gcggctgcgc ctccgacgat
60

gtgggctaagt accctctcgc taacaaccct tactccagcg ccctgaacct caattcctgc
120

cagaactcgt ctatcctcaa ttggattaac atcattgggg acgccgcgaa ggaggctgtc
180

agcatcggca ccacgattgt ttcaactgac acggccccgt ccctgacagg cctcatcagc
240

attgtgtacg acctcatcgg gaagggtcat gcgggggtcat ccggccagag catttcggac
300

ctgtcgatct gcgatctcct gtctatcatt gacctcaggg tgtctcagtc agtcctgaac
360

gacgggatcg ccgatttcaa tggctcagtg ctctgtacc gcaactacct ggaggcgctc
420

gactcctgga acaagaatcc caactccgct agcgctgagg agctgaggac caggttcagg
480

attgccgact ccgagttcga tcgcatcctc actcgcggt cactgaccaa tggcggctcc
540

ctgcgccgcc agaacgctca gatcctcctg ctccctagct tcgcgtcggc tgcccttcttc
600

cacctgctcc tgctcaggga cgctacgcgc tacggaacca actggggcct ctacaatgct
660

actccgttca taaactacca gtccaagctg gtcgagctga tcgagctgta caccgactac
720

tgcgttcact ggtacaatag ggggttcaac gagctgaggc agagggggcac gagcgctaca
780

gcttggtctgg agttccatcg ctaccgcagg gagatgacgc tcatggttct ggacatcgctg
840

gccagcttca gctcgctcga tattactaac taccctatcg agaccgactt ccagctgtcg
900

cgcgatgattt acaccgaccc gatcgggttc gtccaccgct cttcactgag gggcgagtct
960

tggttctcat tcgtcaatag ggcgaaattc tcggacctcg agaattgctat cccgaacccc
1020

cggccatctt ggttcctgaa caatatgac atttccactg gcagcctgac cctccccgtt
1080

tcgccatcta cggatcgcg cagggtgtgg tacggctcac gggaccgcat ctcaccggct
1140

aactcccagt tcattacaga gctgatctct gcccagcaca caactgctac acagactatt
1200

ctgggccgga atatcttccg cgtggacagc caggcctgca atctcaacga taccacgtac
1260

ggcgtcaaca gggctgtttt ctaccatgac gcctcggagg gctctcagcg ctcagtctac
1320

gagggataca tcaggacaac tggcatcgat aatcctcggg tgcagaatat taacacgtac
1380

ctcccgggcg agaacagcga catccctacg ccggaggatt acacacacat tctgtcgacc
1440

acaatcaacc tcaccggcgg gctgaggcag gtcgcttcta acaggcgctc cagcctcggtt
1500

atgtacggct ggactcataa gtcaactggcg cgcaacaaca caatcaaccc tgataggatc
1560

actcagattc cgctcaccaa ggtggacact agggggaccg gcgtgtcgta cgtcaacgat
1620

cccggcttca tcggcggggc cctgctccag cgcaccgacc acgggtccct cggggttctg
1680

cgggtgcagt tccactgca tctccgccag cagtacagga ttgggtccg ctacgcgtcc
1740

acaactaaca tccgcctcag cgtgaatggg tcgttcggca cgatctccca gaacctgccc
1800

agcacaatga ggctcgggga ggacctgcgc tacggctcct tcgccattcg ggagttcaac
 1860

acgagcatcc gcccacagc gtcgccagat cagattaggc tcaactatcg gccaaagcttc
 1920

atcaggcagg aggtctacgt tgaccggatc gagttcattc ctgtcaaccc gacgagggag
 1980

gctaaggagg atctggaggc tgctaagaag gctgtggcta gcctcttcac caggacgagg
 2040

gacggcctgc aagtgaatgt caaggactac caggttgatc aggccgcgaa cctcgtgtcc
 2100

tgcctgagcg acgagcagta cggctacgat aagaagatgc tgctcgaggc ggtcagggct
 2160

gctaagaggc tctccaggga gaggaacctg ctccaggacc ctgatttcaa tacgatcaac
 2220

agcacagagg agaacgggtg gaaggcgtct aatggcgtga ccatttcaga gggcggccca
 2280

ttctacaagg gcagggctat ccagctggct tcggctcggg agaactaccc cacgtacatc
 2340

taccagaagg tcgacgcctc tgagctgaag ccatacacac gctaccgcct ggatggcttc
 2400

gtgaagtcgt ctcaggacct cgagatcgat ctgattcacc atcacaaggt ccacctggtt
 2460

aagaatgtgc ccgacaacct ggtcctcgac acctaccag acgattcctg caatggcatc
 2520

aacaggtgcg acgagcagaa gatggtgaac gccagctcg agaccggaca tcaccatccg
 2580

atggactgct gcgagggcgc tcagacgcac gagttctcat cctacattaa cacaggggac
 2640

ctgaatgcc a gctggatca gggcatctgg gtggtcctca aagtcaggac cacggacggg
 2700

tacgctacgc tgggcaacct ggagctggtt gaagtggggc cactctcggg cgagtctctg
 2760

gagagggagc agagggacaa cgccaagtgg agcgtgagc tgggcaggaa gagggctgag
 2820

accgagaggg tctactacgc cgcgaagcag tcgatcaatc acctcttcgt ggactaccaa
 2880

gatcagcagc tgaacccccca gattggcatg gctgacatca tggatgccca gaacctcgtc
2940

gcgtaaatct ccgacgtgta ctccgatgag gtctgcaga tcccaggcat caactacgag
3000

atctacacgg agctgagcaa ccggctgcag caggcctcgt acctccacac gtctcgcaat
3060

gcatgcaga acggggactt caattctggc ctggattcat ggaatgagac tgctggggcc
3120

accgttcagc aggacggcaa caccacttc ctgggtgctct ccattggga tgcccaagtt
3180

agccagcagt tccgctgca gccgaactgc aagtacgtcc tgagggttac tgctgagaag
3240

gtcggcgagg ggcacggcta cgttaccatc cgcgatggcg ctcaccatac agagactctc
3300

acattcaacg cctgcgacta cgatatcaat ggcacgtacg tgacagacaa cacttacctg
3360

accaaggagg ttgtgttcca cccggagacc cagcatatgt gggtcgaggt tagcgagagc
3420

gagggcgtgt tccacctgga cagcgttgag ttcatggaga cccagcagta g
3471

<210> 14

<211> 5919

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 14

atggacctgg atgggaataa gacagagaca gagaccgaga ttgtgaatgg gagcgagagc
60

agcattgacc cgagcagcgt ttcgtacgct gggaacaata gctactccag cgccctgaac
120

ctcaattcgt gccagaatag gggcatcgct cagtgggtta acacgctggg cggggctatt
180

gggcaggccg tgagcatcgg cacatctatc atttcactcc tggccgcgcc gacactcact
240

gggtctatct cactggcctt caatctcatc aggaggatgg ggaccggctc caacggctcg
300

tctatttccg acctgagcat ctgcgatctc ctgagcatca ttaacctgcg ggtttcgcag
360

gctgtgctca acgacgggat cgctgatttc aatggctccg ttgctgtgta cgacctgtac
420

ctccacgccc tgcgcagctg gaacaataac cctaacgctg ctactgctga ggagctgagg
480

acccgcttca ggatcgccga ttcggagttc gagaggattc tgacgagggg ctcgctcaca
540

catggcggct ccctcgcccg ccaggacgct caggtcctcc tgctcccgtc cttcgttaac
600

gcggcttacc tgcacctgct catcctccgc gatgcttcgc gctacggggc ctcttggggc
660

ctcttcaaca ccacgccgca tatcaattac cccgtgaggc tgacgagct cattggcagc
720

tacacgcact actgcacaca ttggtacaac caggggctga atgagatccg gcagcgcggc
780

aacactgccg tgaattggct cgagttccac cgctaccgcc gcgacatgac gctgatggtc
840

ctcgatgtgg tctcgctgtt ctctgccctc gacacgatcc gctacccgaa cgctacagtt
900

gtgcagctca gccgcactgt ctacaccgat ccgattggct tcgttaaccg cgggtcaggc
960

aataggctgt cctggttcga ctggaggaac caggcgaatt tctctactct cgagtcagag
1020

atgccgaccc cctcatcccc actgagcctc aaccacatgt cgatcttcac tgggcctctg
1080

accctcccag tgtcccctaa cacgcatagg gcccggtct ggtacggcaa ccagaatatg
1140

ttcacaactg ggtcacagaa ctccggccag accacgaact ctattcagaa tatctcaggc
1200

ctggagattt tccgcacga ctctcaggcg tgcaatctca ataacaattc atacggcgtg
1260

aacagggcgg agttcttcca cggggctagc cagggctcgc agcggctctgt ctaccagggg
1320

tacatccgcc agagcggcct ggacaaccct gtcgttatga atctgcagtc tttcctccca
1380

ggcgagaact cagccacccc tacggcgag gattacaccc acattctgtc caaccgggtt
1440

aatatcaggg gcgggctcag gcagattgtg gccgacaggc gtcctccgt ggtcgtttac
1500

ggctggacgc acaagtcctt gagcaggagg tcaactcgtg ctccagacca gatcacccag
1560

gtcccagccg ttaaggcgtc ccctttcttca cattgcacta tcattgccgg cccaggcttc
1620

accggcgggg acctgggtgtc gctccagccc aacggccagc tcgtcatccc gttccagggt
1680

tctgcgcccg agacgaacta ccacattcgc atctgctacg tctcgacgtc tgattgcagc
1740

attaacacaa tctgcaatga cgagacgcat ctgtccacac tcccagcac aacttcagc
1800

ctggagaacc tccagtgcaa tcacctgcat tacttcaacg tgggcacttt caagccaacc
1860

atcgactcga agctgacgct cgtcaacaca tctcctaacg ctaacatcat tatcgacaag
1920

atcgagttca tcccgggtgga taccgcccag cagcagaacg aggacctcga ggccgcgaag
1980

aaggctgtcg cctccctggt cacacgcact agggacggcc tccagggtcaa tgttaaggac
2040

taccagggtg atcagggtgc caacctggtc tcatgcctct ccgacgagca gtacggctac
2100

gataagaaga tgctgctcga ggccgtgagg gctgctaaga ggctgagcag ggagaggaac
2160

ctgctccagg accccgattt caacacaatc aactcgaccg aggagaacgg gtggaaggcg
2220

tcaaattggcg tcaccatctc cgagggcggg ccattctaca agggcagggc tattcagctc
2280

gcgtctgctc gggagaacta cccacatac atctaccaga aggtggatgc ctccgagctg
2340

aagccataca cccgtaccg cctcgacggc ttcgtcaagt cgtctcagga cctggagatt
2400

gatctcatcc accatcaciaa ggtgcacctg gtcaagaacg ttccggacaa tctcgtgagc
2460

atgaacagga acaacccaaa cgagtaagag attattgatg ctccatactg cggctgcccc
2520

tccgatgacg atgtgcgcta cccctcgcc agcgaccga acgccgcgtt ccagaatatg
2580

aactacaagg agtacctgca gacctacgac ggcgattaca cggggtcact gattaatcca
2640

aacctctcca tcaatcctcg cgacgtcctc cagaccggaa tcaacattgt tggccgcacg
2700

ctcggttcc tgggcgtgcc gttcgctggc cagctggta ccttctacac gttcctcctg
2760

aaccagctct ggcctacgaa tgacaacgcg gtgtgggagg cttcatggc gcagatcgag
2820

gagctgattg atcagaagat ctccgctcag gtggtcagga acgccctcga cgatctgacc
2880

ggcctccacg actactacga ggagtacctg gctgctctcg aggagtggct cgagaggcca
2940

aacggcgctc gcgctaattt ggtcacgcag cgcttcgaga acctccatac cgccttcgtg
3000

acgaggatgc cgagcttcgg gacaggcccc gggtcgcaga gggacgcggt tgctctcctg
3060

actgtgtacg cgcaggcggc taacctgcac ctctgctcc tgaaggatgc tgagatctac
3120

ggcgctcggg gggggctgca gcagggccag atcaacctct acttcaatgc ccagcaggag
3180

cgcaacaagga tctacactaa ccactgcgtt gagacctaca atagggggct cgaggacgtg
3240

cggggcacga acacagagag ctggctgaat taccatcgct tccgcaggga gatgacgctc
3300

atggcgatgg acctcgtggc gctcttcccc tactacaacg tccgccagta cccaaatggc
3360

gccaaccctc agctgacaag ggagatctac actgaccga tcgtgtacaa cccaccagct
3420

aatcagggga tctgcaggcg ctggggcaac aatccctaca acaccttctc tgagctggag
3480

aatgcgttca ttaggccacc tcacctcttc gaccgctga acaggctcac catctcccg
3540

aatcgctaca cggctccaac cacgaactcg tacctggatt actggtctgg ccataccctc
3600

cagtcacagt acgcaacaa tcctacaact tacgagacgt cctacggcca gattacaagc
3660

aacactcgcc tcttcaatac cacgaacggc gctaatagta tcgacagcag ggcccggaaac
3720

ttcgggaatc tgtacgcgaa tctctacggc gtctccagcc tgaacatttt cccaaccggc
3780

ggtatgtcag agatcacctc cggccctaac acgtgctggc aggacctcac aactaccgag
3840

gagctgccac tggatgaacaa taacttcaat ctctgtccc acgtcacatt cctgagggtc
3900

aacacgacac agggcggccc actcgcgact gttggcttcg tgcccaccta cgtgtggacg
3960

cggcaggacg tcgatttcaa taacatcatt acaccaaacc gcactactca gattcctgtt
4020

gtgaaggcgt acgagctgtc gtctggcgct acagtcgtta agggcccggg gttcactggc
4080

ggggacgtca ttaggaggac taacaccggc gggttcgggg ctatcagggt gagcgtcaca
4140

ggccccctca ctacgcgcta caggattcgg ttccgctacg cgtcgaccat cgacttogat
4200

ttcttcgtca cgcgcggcgg gactaccatc aataacttcc ggttcacgcg caccatgaac
4260

aggggccagg agagccggta cgagtcgtac cgcaccgtgg agttcacgac accgttcaac
4320

ttcacacaga gccaggacat cattcgcact tctattcagg gcctgtcagg caacggggag
4380

gtttacctcg accggtcga gatcattcca gtgaacccaa ccaggagggc tgaggaggat
4440

ctcgaggctg cgaagaaggc tgtggcctcg ctgttacta ggaccggga cggcctccag
4500

gttaatgtga cggactacca ggtggatcag gctgccaaac tgggtctgtg cctctctgac
4560

gagcagtacg ctacagataa gaagatgtc ctggaggccg tccgcgtgc taagaggctg
4620

tgcagggaga ggaacctcct gcaggacceca gatttcaaca ccattaattc tacggaggag
4680

aacgggtgga aggcgtcgaa tggcgtgacc atctctgagg gcgggccttt ctacaagggc
4740

cgcgctctcc agctggcttc agctagggag aactaccga catacatcta ccagaaggtc
4800

gacgcctccg agctgaagcc ctacacgcgc taccgcctcg atggcttcgt gaagtcatcc
4860

caggacctgg agatcgatct cattcaccat cacaaggctc acctgggtta gaacgtgctg
4920

gacaatctcg tcagcgatac ctaccggac gattcttgct caggcatcaa tcgctgcgag
4980

gagcagcaga tgggtgaacgc ccagctcgag accgagcatc accatccgat ggactgctgc
5040

gaggctgctc agacgcatga gttcagctcg tacatcgaca caggggatct gaactccact
5100

gtggaccagg gcatctgggt catcttcaag gtgcgcacta ccgatgggta cgcgaccctc
5160

ggcaacctcg agctgggtcg ggttgggccg ctctggggcg agccactgga gagggagcag
5220

agggagaacg ctaagtggaa tgctgagctg ggcaggaagc gggctgagac cgaccgcgtc
5280

taccaggatg ctaagcagtc catcaatcac ctgttcgtgg actaccagga tcagcagctc
5340

aaccacaga ttggcatggc tgacatcatg gatgccaga acctgggtcg gtccatcagc
5400

gacgtctact ctgatgccgt tctccagatc cctggcatta attacgagat ctacacagag
5460

ctgtcaaacc ggctccagca ggcctcctac ctgtacacta gccgcaacgc ggtgcagaat
5520

ggggacttca ataacggcct cgattcctgg aatgctacag ctggggctag cgtccagcag
5580

gacggcaaca ctcacttcct cgttctgtct cattgggatg cccaggtttc acagcagttc
5640

aggggtgcagc cgaactgcaa gtatgtgctg agggttaccg ctgagaaggt tggcgggggc
5700

gacggctacg tgacgatcag ggatggcgct caccatacgg agacactgac tttcaacgcc
5760

tgcgactacg atatcaatgg cacctacgtc acggacaaca catacctcac taaggaggtc
5820

ctctttctaca gccacaccga gcatatgtgg gtggagggtca acgagacgga gggcgcttc
5880

cacctcgact cgattgagtt cgtggagacc gagaagtga
5919

<210> 15

<211> 3468

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Gen Cry tổng hợp

<400> 15

atgaatagga ataaccagga cgagtacgag atcattgacg cttcgacttg cgggtgcagc
60

tcagacgatg ttgttcagta cccctcgct cgcgatccca atgccgtgtt ccagaacatg
120

cactacaagg actacctcca gacatacgac ggcgattaca ctgggtcgct cattaaccca
180

aatctgtcta tcaatcctcg cgacgtcctg cagacgggaa tcaacattgt tgggaggctc
240

ctgggcttcc tcggggttcc gttcgctggc cagctgggtga cattctacac tttcctcctg
300

aaccagctct ggcccaccaa cgacaatgcg gtctgggagg cttcatggc gcagatcgag
360

gagctgatta accagcgcat ctccgaggct gtggtcggca ccgctgctga tcacctcacg
420

gggctgcatg acaactacga gctttacgtg gaggctctcg aggagtggct ggagaggcca
480

aacgctgcta ggacgaacct cctgttcaat aggttcacca cgctcgactc cctgttcaca
540

cagttcatgc catcattcgg gactggcccc ggctcccaga actacgccgt ccctctcctg
600

accgtttacg cgcaggcggc taacctccac ctctgtctcc tgaaggacgc tgagatctac
660

ggcgctcgct gggggctgaa ccagaatcag atcaacagct tccacacaag gcagcaggag
720

cggacccagt actacacgaa tcattgcgtg acaacttaca acaccggcct cgataggctg
780

agggggacca atacggagtc ctgggtcaac taccataggt tccgcaggga gatgacctg
840

atggcgatgg acctcgtggc gctgttcccg tactacaacg tccggcagta cccaacggc
900

gctaattccac agctcaccgc cgagatctac acggacccaa tcgtcttcaa cccaccagct
960

aatgttgccc tgtgcaggcg ctgggggaac aatccttaca atcgcttcag cgagctggag
1020

aacgccttca tcaggccacc tcacctgttc gatcgggtca atacactgac tatttcgagg
1080

aaccggttcg acgtcggcag caatttcacg gagccgtggt cgggccatac gctgaggagg
1140

tcctacagca acaattcaac agtgtacgag gattcctacg gccagattac cgctacgagg
1200

accacgatca acctccctgc caatggcaca gggcgggtgg agtctactgc tgtcgacttc
1260

cgctcagccc tggttggcat ctacggggtg aacagggcga gcttcatccc gggcggggtc
1320

ttcagcggca caactcagcc atcgaccggc ggggtgcagg acctctacga ttccagcgac
1380

gagctgccac cagacgagag cacgggctcg ttcgctcaca ggctctctca tgtcacattc
1440

ctgtcattca ccacgaacca ggctggctcg atcgccaatt ctgggagggg tccgacatac
1500

gtgtggactc accgggacgt ggatttcaac aataccatca acccaaatcg catcacgcag
1560

attcctgttg tgaaggcgta cgagctgtcc tccggcgcta ccgtcgtaa gggcccgggg
1620

ttcacgggcg gggacgtgat tcgcaggaca aacatcggcg ggttcggcgc tatcagggtc
1680

tccgttaccg gccccctgac gcagcgctac aggattcggg tccgctacgc ctccactatc
1740

gacttcgatt tcttcgtcac ccggggcggg acaactatca acaatttccg cttcacaagg
1800

actatgaaca ggggccagga gtcacggtac gagtcctacc gcaccgtgga gttcaccacg
1860

cccttcaact tcacgcagtc ccaggacatc attaggacat ccattcaggg cctcagcggc
1920

aacggggagg tttacctgga tcgcatcgag atcattccag tgaaccctac gagggagggt
1980

gaggaggacc tcgaggctgc gaagaaggct gtggctagcc tcttcaccag gacgagggat
2040

ggcctgcaag tgaatgtcac cgattaccag gtggaccagg ctgccaacct cgtctcttgc
2100

ctgtcagatg agcagtacgg ccacgacaag aagatgctcc tggaggccgt ccgcgctgct
2160

aagaggctgt ccagggagag gaacctcctg caggaccag atttcaacac aatcaatagc
2220

actgaggaga atggctggaa ggcgtccaac ggggtgacca tcagcgaggg cgggcctttc
2280

tacaagggcc gcgctctcca gctggcttcg gctagggaga attaccaac ttacatctac
2340

cagaaggcca acgcctctga gctgaagcct tacacgcgct accgcctgga tggcttcgtg
2400

aagtcattccc aggatctcga gatcgacctg attcaccatc acaagggtgca cctcgtcaag
2460

aacgttccgg acaatctggg cagcgatacg tactcggacg gctcgtgctc tgggatgaat
2520

cgctgcgagg agcagcagat ggtgaacgcg cagctcgaga cagagcatca ccatccgatg
2580

gactgctgcg aggccgcgca gactcatgag ttcagctcgt acattaatac cggcgatctg
2640

aactcttcag tcgaccaggg catctgggtg gtctctcaagg ttgcgacaa tgatggctac
2700

gccacgctgg ggaacctcga gctggttgaa gtgggcccac tctccgggga gagcctggag
2760

agggagcaga gggacaacgc gaagtgggtcc gctgagctgg gcaggaagcg ggctgagacc
2820

gatagggttt accaggacgc taagcagtcc atcaatcacc tcttcgtgga ttaccaggac
2880

cagcagctga acccggagat tggcatggct gatatcattg acgcccagaa cctcgtcgcg
2940

tcaatctccg atgtctacag cgacgccgtt ctgcagatcc ccggcattaa ttacgagatc
3000

tacacggagc tgtcgaacag gctgcagcag gcctcgtacc tctacacatc tcggaacgcg
3060

gtgcagaatg gcgatttcaa ctctgggctg gactcatgga atgctaccgg cggggctacg
3120

gtgcagcagg atggcaacac ccacttcctc gtctgtccc attgggacgc tcaggtgagc
3180

cagcagttcc ggggtccagcc gaactgcaag tatgtgctga gggtcactgc tgagaagggt
3240

ggcgggggcg atggctacgt gaccatcagg gacggggctc accatacaga gaagctcact
3300

ttcaacgcct gcgactacga tatcaatggc acctacgtca cggacaacac atacatcact
3360

aaggaggttg tgttctactc gcacactgag catatgtggg tcgagggttc tgagaccgag
3420

ggcgccttcc acctcgactc aattgagttc gtggagaccg agaagtga
3468

<210> 16
<211> 1168
<212> PRT
<213> Bacillus thuringiensis

<400> 16

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

Cys	Gly	Cys	Pro	Ala	Asp	Asp	Val	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	Thr
		35					40					45			

Tyr 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

Thr 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

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

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

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

Leu His Leu Leu Leu Leu Lys Asp Ala Asp 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 Ala 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 Pro 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 His 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 Thr Asn Asn Arg Leu Thr Asn Phe Ser Met Pro Ser Asp
 405 410 415

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

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

Phe Ser Thr Tyr Val Gly Gln Asn Ser Val Leu Pro Gln Cys His Gln
 450 455 460

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Cys Ser Gly Ile Asn Arg Cys Gln Glu Gln Gln Met Val Asn Ala Gln
 850 855 860

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

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

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

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

Gly Glu Pro Leu Glu Arg Glu Gln Arg Glu Asn Ala Lys Trp Asn Ala
 930 935 940

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

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

Asn Pro Glu Ile Gly Met Ala Asp Ile Met Asp Ala Gln Asn Leu Val
 980 985 990

Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly
 995 1000 1005

Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln
 1010 1015 1020

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

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

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

Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys
 1070 1075 1080

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

Tyr Val Thr Ile Arg Asp Asp Ala His His Thr Glu Thr Leu Thr
 1100 1105 1110

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

Asn Thr Tyr Leu Thr Lys Glu Val Val Phe Tyr Ser His Thr Asp
 1130 1135 1140

His Met Trp Val Glu Val Ser Glu Thr Glu Gly Ala Phe His Ile
 1145 1150 1155

Asp Ser Ile Glu Phe Val Glu Thr Glu Lys
 1160 1165

<210> 17
 <211> 1156
 <212> PRT
 <213> Bacillus thuringiensis

<400> 17

Met Asn Gln Asn Lys His Gly Ile Ile Gly Ala Ser Asn Cys Gly Cys
 1 5 10 15

Ala Ser Asp Asp Val Ala Lys Tyr Pro Leu Ala Asn Asn Pro Tyr Ser
 20 25 30

Ser Ala Leu Asn Leu Asn Ser Cys Gln Asn Ser Ser Ile Leu Asn Trp
 35 40 45

Ile Asn Ile Ile Gly Asp Ala Ala Lys Glu Ala Val Ser Ile Gly Thr
 50 55 60

Thr Ile Val Ser Leu Ile Thr Ala Pro Ser Leu Thr Gly Leu Ile Ser
 65 70 75 80

Ile Val Tyr Asp Leu Ile Gly Lys Val Leu Gly Gly Ser Ser Gly Gln
 85 90 95

Ser Ile Ser Asp Leu Ser Ile Cys Asp Leu Leu Ser Ile Ile Asp Leu
 100 105 110

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

Ser Val Leu Leu Tyr Arg Asn Tyr Leu Glu Ala Leu Asp Ser Trp Asn
 130 135 140

Lys Asn Pro Asn Ser Ala Ser Ala Glu Glu Leu Arg Thr Arg Phe Arg
 145 150 155 160

Ile Ala Asp Ser Glu Phe Asp Arg Ile Leu Thr Arg Gly Ser Leu Thr
 165 170 175

Asn Gly Gly Ser Leu Ala Arg Gln Asn Ala Gln Ile Leu Leu Leu Pro
 180 185 190

Ser Phe Ala Ser Ala Ala Phe Phe His Leu Leu Leu Leu Arg Asp Ala
 195 200 205

Thr Arg Tyr Gly Thr Asn Trp Gly Leu Tyr Asn Ala Thr Pro Phe Ile
 210 215 220

Asn Tyr Gln Ser Lys Leu Val Glu Leu Ile Glu Leu Tyr Thr Asp Tyr
 225 230 235 240

Cys Val His Trp Tyr Asn Arg Gly Phe Asn Glu Leu Arg Gln Arg Gly
 245 250 255

Thr Ser Ala Thr Ala Trp Leu Glu Phe His Arg Tyr Arg Arg Glu Met
 260 265 270

Thr Leu Met Val Leu Asp Ile Val Ala Ser Phe Ser Ser Leu Asp Ile
 275 280 285

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

Thr Asp Pro Ile Gly Phe Val His Arg Ser Ser Leu Arg Gly Glu Ser
 305 310 315 320

Trp Phe Ser Phe Val Asn Arg Ala Asn Phe Ser Asp Leu Glu Asn Ala
 325 330 335

Ile Pro Asn Pro Arg Pro Ser Trp Phe Leu Asn Asn Met Ile Ile Ser

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

Arg Val Gln Phe Pro Leu His Leu Arg Gln Gln Tyr Arg Ile Arg Val
 565 570 575

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

Gly Thr Ile Ser Gln Asn Leu Pro Ser Thr Met Arg Leu Gly Glu Asp
 595 600 605

Leu Arg Tyr Gly Ser Phe Ala Ile Arg Glu Phe Asn Thr Ser Ile Arg
 610 615 620

Pro Thr Ala Ser Pro Asp Gln Ile Arg Leu Thr Ile Glu Pro Ser Phe
 625 630 635 640

Ile Arg Gln Glu Val Tyr Val Asp Arg Ile Glu Phe Ile Pro Val Asn
 645 650 655

Pro Thr Arg Glu Ala Lys Glu Asp Leu Glu Ala Ala Lys Lys Ala Val
 660 665 670

Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Lys
 675 680 685

Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp
 690 695 700

Glu Gln Tyr Gly Tyr Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala
 705 710 715 720

Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe
 725 730 735

Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly
 740 745 750

Val Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Ile Gln
 755 760 765

Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val

770																			
Asp	Ala	Ser	Glu	Leu	Lys	Pro	Tyr	Thr	Arg	Tyr	Arg	Leu	Asp	Gly	Phe				
785					790					795					800				
Val	Lys	Ser	Ser	Gln	Asp	Leu	Glu	Ile	Asp	Leu	Ile	His	His	His	Lys				
				805					810					815					
Val	His	Leu	Val	Lys	Asn	Val	Pro	Asp	Asn	Leu	Val	Leu	Asp	Thr	Tyr				
			820					825					830						
Pro	Asp	Asp	Ser	Cys	Asn	Gly	Ile	Asn	Arg	Cys	Asp	Glu	Gln	Lys	Met				
		835					840					845							
Val	Asn	Ala	Gln	Leu	Glu	Thr	Glu	His	His	His	Pro	Met	Asp	Cys	Cys				
	850					855					860								
Glu	Ala	Ala	Gln	Thr	His	Glu	Phe	Ser	Ser	Tyr	Ile	Asn	Thr	Gly	Asp				
865					870					875					880				
Leu	Asn	Ala	Ser	Val	Asp	Gln	Gly	Ile	Trp	Val	Val	Leu	Lys	Val	Arg				
				885					890					895					
Thr	Thr	Asp	Gly	Tyr	Ala	Thr	Leu	Gly	Asn	Leu	Glu	Leu	Val	Glu	Val				
			900					905					910						
Gly	Pro	Leu	Ser	Gly	Glu	Ser	Leu	Glu	Arg	Glu	Gln	Arg	Asp	Asn	Ala				
		915					920					925							
Lys	Trp	Ser	Ala	Glu	Leu	Gly	Arg	Lys	Arg	Ala	Glu	Thr	Glu	Arg	Val				
	930					935					940								
Tyr	Tyr	Ala	Ala	Lys	Gln	Ser	Ile	Asn	His	Leu	Phe	Val	Asp	Tyr	Arg				
945					950					955					960				
Asp	Gln	Gln	Leu	Asn	Pro	Gln	Ile	Gly	Met	Ala	Asp	Ile	Met	Asp	Ala				
				965					970					975					
Gln	Asn	Leu	Val	Ala	Ser	Ile	Ser	Asp	Val	Tyr	Ser	Asp	Ala	Val	Leu				
			980					985					990						

Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg
 995 1000 1005

Leu Gln Gln Ala Ser Tyr Leu His Met Ser Arg Asn Ala Met Gln
 1010 1015 1020

Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Ala
 1025 1030 1035

Gly Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu
 1040 1045 1050

Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro
 1055 1060 1065

Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly
 1070 1075 1080

Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu
 1085 1090 1095

Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr
 1100 1105 1110

Val Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe His Pro
 1115 1120 1125

Glu Thr Gln His Met Trp Val Glu Val Ser Glu Thr Glu Gly Val
 1130 1135 1140

Phe His Ile Asp Ser Val Glu Phe Met Glu Thr Gln Gln
 1145 1150 1155

<210> 18

<211> 1156

<212> PRT

<213> Bacillus thuringiensis

<400> 18

Met Asn Gln Asn Lys His Gly Ile Ile Gly Ala Ser Asn Cys Gly Cys
 1 5 10 15

Ala Ser Asp Asp Val Ala Lys Tyr Pro Leu Ala Asn Asn Pro Tyr Ser
 20 25 30

Ser Ala Leu Asn Leu Asn Ser Cys Gln Asn Ser Ser Ile Leu Asn Trp
 35 40 45

Ile Asn Ile Ile Gly Asp Ala Ala Lys Glu Ala Val Ser Ile Gly Thr
 50 55 60

Thr Ile Val Ser Leu Ile Thr Ala Pro Ser Leu Thr Gly Leu Ile Ser
 65 70 75 80

Ile Val Tyr Asp Leu Ile Gly Lys Val Leu Gly Gly Ser Ser Gly Gln
 85 90 95

Ser Ile Ser Asp Leu Ser Ile Cys Asp Leu Leu Ser Ile Ile Asp Leu
 100 105 110

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

Ser Val Leu Leu Tyr Arg Asn Tyr Leu Glu Ala Leu Asp Ser Trp Asn
 130 135 140

Lys Asn Pro Asn Ser Ala Ser Ala Glu Glu Leu Arg Thr Arg Phe Arg
 145 150 155 160

Ile Ala Asp Ser Glu Phe Asp Arg Ile Leu Thr Arg Gly Ser Leu Thr
 165 170 175

Asn Gly Gly Ser Leu Ala Arg Gln Asn Ala Gln Ile Leu Leu Leu Pro
 180 185 190

Ser Phe Ala Ser Ala Ala Phe Phe His Leu Leu Leu Leu Arg Asp Ala
 195 200 205

Thr Arg Tyr Gly Thr Asn Trp Gly Leu Tyr Asn Ala Thr Pro Phe Ile
 210 215 220

Asn Tyr Gln Ser Lys Leu Val Glu Leu Ile Glu Leu Tyr Thr Asp Tyr
 225 230 235 240

Cys Val His Trp Tyr Asn Arg Gly Phe Asn Glu Leu Arg Gln Arg Gly
 245 250 255

Thr Ser Ala Thr Ala Trp Leu Glu Phe His Arg Tyr Arg Arg Glu Met
 260 265 270

Thr Leu Met Val Leu Asp Ile Val Ala Ser Phe Ser Ser Leu Asp Ile
 275 280 285

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

Thr Asp Pro Ile Gly Phe Val His Arg Ser Ser Leu Arg Gly Glu Ser
 305 310 315 320

Trp Phe Ser Phe Val Asn Arg Ala Asn Phe Ser Asp Leu Glu Asn Ala
 325 330 335

Ile Pro Asn Pro Arg Pro Ser Trp Phe Leu Asn Asn Met Ile Ile Ser
 340 345 350

Thr Gly Ser Leu Thr Leu Pro Val Ser Pro Ser Thr Asp Arg Ala Arg
 355 360 365

Val Trp Tyr Gly Ser Arg Asp Arg Ile Ser Pro Ala Asn Ser Gln Phe
 370 375 380

Ile Thr Glu Leu Ile Ser Gly Gln His Thr Thr Ala Thr Gln Thr Ile
 385 390 395 400

Leu Gly Arg Asn Ile Phe Arg Val Asp Ser Gln Ala Cys Asn Leu Asn
 405 410 415

Asp Thr Thr Tyr Gly Val Asn Arg Ala Val Phe Tyr His Asp Ala Ser
 420 425 430

Glu Gly Ser Gln Arg Ser Val Tyr Glu Gly Tyr Ile Arg Thr Thr Gly
 435 440 445

Ile Asp Asn Pro Arg Val Gln Asn Ile Asn Thr Tyr Leu Pro Gly Glu
 450 455 460

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

Thr Ile Asn Leu Thr Gly Gly Leu Arg Gln Val Ala Ser Asn Arg Arg
 485 490 495

Ser Ser Leu Val Met Tyr Gly Trp Thr His Lys Ser Leu Ala Arg Asn
 500 505 510

Asn Thr Ile Asn Pro Asp Arg Ile Thr Gln Ile Pro Leu Thr Lys Val
 515 520 525

Asp Thr Arg Gly Thr Gly Val Ser Tyr Val Asn Asp Pro Gly Phe Ile
 530 535 540

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

Arg Val Gln Phe Pro Leu His Leu Arg Gln Gln Tyr Arg Ile Arg Val
 565 570 575

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

Gly Thr Ile Ser Gln Asn Leu Pro Ser Thr Met Arg Leu Gly Glu Asp
 595 600 605

Leu Arg Tyr Gly Ser Phe Ala Ile Arg Glu Phe Asn Thr Ser Ile Arg
 610 615 620

Pro Thr Ala Ser Pro Asp Gln Ile Arg Leu Thr Ile Glu Pro Ser Phe
 625 630 635 640

Ile Arg Gln Glu Val Tyr Val Asp Arg Ile Glu Phe Ile Pro Val Asn
 645 650 655

Pro Thr Arg Glu Ala Lys Glu Asp Leu Glu Ala Ala Lys Lys Ala Val
 660 665 670

Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Lys
 675 680 685

Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp
 690 695 700

Glu Gln Tyr Gly Tyr Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala
 705 710 715 720

Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe
 725 730 735

Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly
 740 745 750

Val Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Ile Gln
 755 760 765

Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val
 770 775 780

Asp Ala Ser Glu Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe
 785 790 795 800

Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys
 805 810 815

Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Leu Asp Thr Tyr
 820 825 830

Pro Asp Asp Ser Cys Asn Gly Ile Asn Arg Cys Asp Glu Gln Lys Met
 835 840 845

Val Asn Ala Gln Leu Glu Thr Gly His His His Pro Met Asp Cys Cys
 850 855 860

Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp
 865 870 875 880

Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg
 885 890 895

Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val
 900 905 910

Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn Ala
 915 920 925

Lys Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Glu Arg Val
 930 935 940

Tyr Tyr Ala Ala Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Gln
 945 950 955 960

Asp Gln Gln Leu Asn Pro Gln Ile Gly Met Ala Asp Ile Met Asp Ala
 965 970 975

Gln Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu
 980 985 990

Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg
 995 1000 1005

Leu Gln Gln Ala Ser Tyr Leu His Thr Ser Arg Asn Ala Met Gln
 1010 1015 1020

Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Ala
 1025 1030 1035

Gly Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu
 1040 1045 1050

Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro
 1055 1060 1065

Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly
 1070 1075 1080

Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu
1085 1090 1095

Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr
1100 1105 1110

Val Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe His Pro
1115 1120 1125

Glu Thr Gln His Met Trp Val Glu Val Ser Glu Thr Glu Gly Val
1130 1135 1140

Phe His Ile Asp Ser Val Glu Phe Met Glu Thr Gln Gln
1145 1150 1155

<210> 19
<211> 1152
<212> PRT
<213> Bacillus thuringiensis

<400> 19

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

Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp
20 25 30

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

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

Asn Pro Arg Asp Val Leu Gln Thr Gly Ile Asn Ile Val Gly Arg Ile
65 70 75 80

Leu Gly Phe Leu Gly Val Pro Phe Ala Gly Gln Leu Val Thr Phe Tyr
85 90 95

Thr Phe Leu Leu Asn Gln Leu Trp Pro Thr Asn Asp Asn Ala Val Trp
100 105 110

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

Ala Gln Val Val Arg Asn Ala Leu Asp Asp Leu Thr Gly Leu His Asp
 130 135 140

Tyr Tyr Glu Glu Tyr Leu Ala Ala Leu Glu Glu Trp Leu Glu Arg Pro
 145 150 155 160

Asn Gly Ala Arg Ala Asn Leu Val Thr Gln Arg Phe Glu Asn Leu His
 165 170 175

Thr Ala Phe Val Thr Arg Met Pro Ser Phe Gly Thr Gly Pro Gly Ser
 180 185 190

Gln Arg Asp Ala Val Ala Leu Leu Thr Val Tyr Ala Gln Ala Ala Asn
 195 200 205

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

Gly Leu Gln Gln Gly Gln Ile Asn Leu Tyr Phe Asn Ala Gln Gln Glu
 225 230 235 240

Arg Thr Arg Ile Tyr Thr Asn His Cys Val Glu Thr Tyr Asn Arg Gly
 245 250 255

Leu Glu Asp Val Arg Gly Thr Asn Thr Glu Ser Trp Leu Asn Tyr His
 260 265 270

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

Asn Gln Gly Ile Cys Arg Arg Trp Gly Asn Asn Pro Tyr Asn Thr Phe

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

Arg Arg Thr Asn Thr Gly Gly Phe Gly Ala Ile Arg Val Ser Val Thr
 545 550 555 560
 Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr Ala Ser Thr
 565 570 575
 Ile Asp Phe Asp Phe Phe Val Thr Arg Gly Gly Thr Thr Ile Asn Asn
 580 585 590
 Phe Arg Phe Thr Arg Thr Met Asn Arg Gly Gln Glu Ser Arg Tyr Glu
 595 600 605
 Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe Thr Gln Ser
 610 615 620
 Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn Gly Glu
 625 630 635 640
 Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro Thr Arg Glu
 645 650 655
 Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala Ser Leu Phe
 660 665 670
 Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp Tyr Gln Val
 675 680 685
 Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu Gln Tyr Ala
 690 695 700
 His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala Lys Arg Leu
 705 710 715 720
 Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn Thr Ile Asn
 725 730 735
 Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val Thr Ile Ser
 740 745 750
 Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Leu Gln Leu Ala Ser Ala

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

Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly
 980 985 990

Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln Ala
 995 1000 1005

Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val Gln Asn Gly Asp Phe
 1010 1015 1020

Asn Asn Gly Leu Asp Ser Trp Asn Ala Thr Ala Gly Ala Ser Val
 1025 1030 1035

Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu Ser His Trp Asp
 1040 1045 1050

Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys Tyr
 1055 1060 1065

Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly Gly Asp Gly Tyr
 1070 1075 1080

Val Thr Ile Arg Asp Gly Ala His His Thr Glu Thr Leu Thr Phe
 1085 1090 1095

Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr Val Thr Asp Asn
 1100 1105 1110

Thr Tyr Leu Thr Lys Glu Val Ile Phe Tyr Ser His Thr Glu His
 1115 1120 1125

Met Trp Val Glu Val Asn Glu Thr Glu Gly Ala Phe His Ile Asp
 1130 1135 1140

Ser Ile Glu Phe Val Glu Thr Glu Lys
 1145 1150

<210> 20
 <211> 1155
 <212> PRT
 <213> Bacillus thuringiensis

<400> 20

Met Asn Arg Asn Asn Gln Asp Glu Tyr Glu Ile Ile Asp Ala Ser Thr
 1 5 10 15

Cys Gly Cys Ser Ser Asp Asp Val Val Gln Tyr Pro Leu Ala Arg Asp
 20 25 30

Pro Asn Ala Val Phe Gln Asn Met His Tyr Lys Asp Tyr Leu Gln Thr
 35 40 45

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

Asn Pro Arg Asp Val Leu Gln Thr Gly Ile Asn Ile Val Gly Arg Leu
 65 70 75 80

Leu Gly Phe Leu Gly Val Pro Phe Ala Gly Gln Leu Val Thr Phe Tyr
 85 90 95

Thr Phe Leu Leu Asn Gln Leu Trp Pro Thr Asn Asp Asn Ala Val Trp
 100 105 110

Glu Ala Phe Met Ala Gln Ile Glu Glu Leu Ile Asn Gln Arg Ile Ser
 115 120 125

Glu Ala Val Val Gly Thr Ala Ala Asp His Leu Thr Gly Leu His Asp
 130 135 140

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

Asn Ala Ala Arg Thr Asn Leu Leu Phe Asn Arg Phe Thr Thr Leu Asp
 165 170 175

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

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

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

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

Arg Thr Gln Tyr Tyr Thr Asn His Cys Val Thr Thr Tyr Asn Thr Gly
 245 250 255

Leu Asp Arg Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Asn Tyr His
 260 265 270

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

Asn Val Gly Leu Cys Arg Arg Trp Gly Asn Asn Pro Tyr Asn Arg Phe
 325 330 335

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

Leu Asn Thr Leu Thr Ile Ser Arg Asn Arg Phe Asp Val Gly Ser Asn
 355 360 365

Phe Ile Glu Pro Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Ser Asn
 370 375 380

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

Thr Thr Ile Asn Leu Pro Ala Asn Gly Thr Gly Arg Val Glu Ser Thr
 405 410 415

Ala Val Asp Phe Arg Ser Ala Leu Val Gly Ile Tyr Gly Val Asn Arg
 420 425 430

Ala Ser Phe Ile Pro Gly Gly Val Phe Ser Gly Thr Thr Gln Pro Ser
 435 440 445

Thr Gly Gly Cys Arg Asp Leu Tyr Asp Ser Ser Asp Glu Leu Pro Pro
 450 455 460

Asp Glu Ser Thr Gly Ser Phe Ala His Arg Leu Ser His Val Thr Phe
 465 470 475 480

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

Val Pro Thr Tyr Val Trp Thr His Arg Asp Val Asp Phe Asn Asn Thr
 500 505 510

Ile Asn Pro Asn Arg Ile Thr Gln Ile Pro Val Val Lys Ala Tyr Glu
 515 520 525

Leu Ser Ser Gly Ala Thr Val Val Lys Gly Pro Gly Phe Thr Gly Gly
 530 535 540

Asp Val Ile Arg Arg Thr Asn Ile Gly Gly Phe Gly Ala Ile Arg Val
 545 550 555 560

Ser Val Thr Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr
 565 570 575

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

Ile Asn Asn Phe Arg Phe Thr Arg Thr Met Asn Arg Gly Gln Glu Ser
 595 600 605

Arg Tyr Glu Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe
 610 615 620

Thr Gln Ser Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly
 625 630 635 640

Asn Gly Glu Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro
 645 650 655

Thr Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala
 660 665 670

Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp
 675 680 685

Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu
 690 695 700

Gln Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala
 705 710 715 720

Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn
 725 730 735

Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val
 740 745 750

Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Leu Gln Leu
 755 760 765

Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val Asn
 770 775 780

Ala Ser Glu Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe Val
 785 790 795 800

Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys Val
 805 810 815

His Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser Asp Thr Tyr Ser
 820 825 830

Asp Gly Ser Cys Ser Gly Met Asn Arg Cys Glu Glu Gln Gln Met Val
 835 840 845

Asn Ala Gln Leu Glu Thr Glu His His His Pro Met Asp Cys Cys Glu
 850 855 860

Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp Leu
 865 870 875 880

Asn Ser Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg Thr
 885 890 895

Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val Gly
 900 905 910

Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn Ala Lys
 915 920 925

Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Asp Arg Val Tyr
 930 935 940

Gln Asp Ala Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Gln Asp
 945 950 955 960

Gln Gln Leu Asn Pro Glu Ile Gly Met Ala Asp Ile Ile Asp Ala Gln
 965 970 975

Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln
 980 985 990

Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu
 995 1000 1005

Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val Gln Asn
 1010 1015 1020

Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Gly Gly
 1025 1030 1035

Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu Ser
 1040 1045 1050

His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn
 1055 1060 1065

Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly Gly
 1070 1075 1080

Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu Lys
 1085 1090 1095

Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr Val
 1100 1105 1110

Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe Tyr Ser His
 1115 1120 1125

Thr Glu His Met Trp Val Glu Val Ser Glu Thr Glu Gly Ala Phe
 1130 1135 1140

His Ile Asp Ser Ile Glu Phe Val Glu Thr Glu Lys
 1145 1150 1155

<210> 21
 <211> 1168
 <212> PRT
 <213> Trình tự Nhân tạo

<220>
 <223> BT-0009 đột biến

<400> 21

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

Cys Gly Cys Pro Ala Asp Asp Val 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 Thr
 35 40 45

Tyr 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

Thr 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

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

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

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

Leu His Leu Leu Leu Leu Lys Asp Ala Asp 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 Ala 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 Pro 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 His 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 Thr Asn Asn Arg Leu Thr Asn Phe Ser Met Pro Ser Asp
 405 410 415

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

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

Phe Ser Thr Tyr Val Gly Gln Asn Ser Val Leu Pro Gln Cys His Gln
 450 455 460

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Cys Ser Gly Ile Asn Arg Cys Gln Glu Gln Gln Met Val Asn Ala Gln
850 855 860

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

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

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

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

Gly Glu Pro Leu Glu Arg Glu Gln Arg Glu Asn Ala Lys Trp Asn Ala
930 935 940

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

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

Asn Pro Glu Ile Gly Met Ala Asp Ile Met Asp Ala Gln Asn Leu Val
980 985 990

Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly
995 1000 1005

Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln
1010 1015 1020

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

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

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

Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys
1070 1075 1080

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

Tyr Val Thr Ile Arg Asp Asp Ala His His Thr Glu Thr Leu Thr
1100 1105 1110

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

Asn Thr Tyr Ile Thr Lys Glu Val Val Phe Tyr Ser His Thr Asp
1130 1135 1140

His Met Trp Val Glu Val Ser Glu Thr Glu Gly Ala Phe His Ile
1145 1150 1155

Asp Ser Leu Glu Phe Val Glu Thr Glu Lys
1160 1165

<210> 22
<211> 1156
<212> PRT
<213> Trình tự Nhân tạo

<220>
<223> BT-0012 đột biến

<400> 22

Met Asn Gln Asn Lys His Gly Ile Ile Gly Ala Ser Asn Cys Gly Cys
1 5 10 15

Ala Ser Asp Asp Val Ala Lys Tyr Pro Leu Ala Asn Asn Pro Tyr Ser
20 25 30

Ser Ala Leu Asn Leu Asn Ser Cys Gln Asn Ser Ser Ile Leu Asn Trp
35 40 45

Ile Asn Ile Ile Gly Asp Ala Ala Lys Glu Ala Val Ser Ile Gly Thr
50 55 60

Thr Ile Val Ser Leu Ile Thr Ala Pro Ser Leu Thr Gly Leu Ile Ser
65 70 75 80

Ile Val Tyr Asp Leu Ile Gly Lys Val Ile Gly Gly Ser Ser Gly Gln
85 90 95

Ser Ile Ser Asp Leu Ser Ile Cys Asp Leu Leu Ser Ile Ile Asp Leu
100 105 110

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

Ser Val Leu Leu Tyr Arg Asn Tyr Leu Glu Ala Leu Asp Ser Trp Asn
130 135 140

Lys Asn Pro Asn Ser Ala Ser Ala Glu Glu Leu Arg Thr Arg Phe Arg
145 150 155 160

Ile Ala Asp Ser Glu Phe Asp Arg Ile Leu Thr Arg Gly Ser Leu Thr
 165 170 175

Asn Gly Gly Ser Leu Ala Arg Gln Asn Ala Gln Ile Leu Leu Leu Pro
 180 185 190

Ser Phe Ala Ser Ala Ala Phe Phe His Leu Leu Leu Leu Arg Asp Ala
 195 200 205

Thr Arg Tyr Gly Thr Asn Trp Gly Leu Tyr Asn Ala Thr Pro Phe Ile
 210 215 220

Asn Tyr Gln Ser Lys Leu Val Glu Leu Ile Glu Leu Tyr Thr Asp Tyr
 225 230 235 240

Cys Val His Trp Tyr Asn Arg Gly Phe Asn Glu Leu Arg Gln Arg Gly
 245 250 255

Thr Ser Ala Thr Ala Trp Leu Glu Phe His Arg Tyr Arg Arg Glu Met
 260 265 270

Thr Leu Met Val Leu Asp Ile Val Ala Ser Phe Ser Ser Leu Asp Ile
 275 280 285

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

Thr Asp Pro Ile Gly Phe Val His Arg Ser Ser Leu Arg Gly Glu Ser
 305 310 315 320

Trp Phe Ser Phe Val Asn Arg Ala Asn Phe Ser Asp Leu Glu Asn Ala
 325 330 335

Ile Pro Asn Pro Arg Pro Ser Trp Phe Leu Asn Asn Met Ile Ile Ser
 340 345 350

Thr Gly Ser Leu Thr Leu Pro Val Ser Pro Ser Thr Asp Arg Ala Arg
 355 360 365

Val Trp Tyr Gly Ser Arg Asp Arg Ile Ser Pro Ala Asn Ser Gln Phe
 370 375 380

Ile Thr Glu Leu Ile Ser Gly Gln His Thr Thr Ala Thr Gln Thr Ile
 385 390 395 400

Leu Gly Arg Asn Ile Phe Arg Val Asp Ser Gln Ala Cys Asn Leu Asn
 405 410 415

Asp Thr Thr Tyr Gly Val Asn Arg Ala Val Phe Tyr His Asp Ala Ser
 420 425 430

Glu Gly Ser Gln Arg Ser Val Tyr Glu Gly Tyr Ile Arg Thr Thr Gly
 435 440 445

Ile Asp Asn Pro Arg Val Gln Asn Ile Asn Thr Tyr Leu Pro Gly Glu
 450 455 460

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

Thr Ile Asn Leu Thr Gly Gly Leu Arg Gln Val Ala Ser Asn Arg Arg
 485 490 495

Ser Ser Leu Val Met Tyr Gly Trp Thr His Lys Ser Leu Ala Arg Asn
 500 505 510

Asn Thr Ile Asn Pro Asp Arg Ile Thr Gln Ile Pro Leu Thr Lys Val
 515 520 525

Asp Thr Arg Gly Thr Gly Val Ser Tyr Val Asn Asp Pro Gly Phe Ile
 530 535 540

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

Arg Val Gln Phe Pro Leu His Leu Arg Gln Gln Tyr Arg Ile Arg Val
 565 570 575

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

Gly Thr Ile Ser Gln Asn Leu Pro Ser Thr Met Arg Leu Gly Glu Asp
 595 600 605

Leu Arg Tyr Gly Ser Phe Ala Ile Arg Glu Phe Asn Thr Ser Ile Arg
 610 615 620

Pro Thr Ala Ser Pro Asp Gln Ile Arg Leu Thr Ile Glu Pro Ser Phe
 625 630 635 640

Ile Arg Gln Glu Val Tyr Val Asp Arg Ile Glu Phe Ile Pro Val Asn
 645 650 655

Pro Thr Arg Glu Ala Lys Glu Asp Leu Glu Ala Ala Lys Lys Ala Val
 660 665 670

Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Lys
 675 680 685

Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp
 690 695 700

Glu Gln Tyr Gly Tyr Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala
 705 710 715 720

Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe
 725 730 735

Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly
 740 745 750

Val Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Ile Gln
 755 760 765

Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val
 770 775 780

Asp Ala Ser Glu Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe
 785 790 795 800

Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys
805 810 815

Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Leu Asp Thr Tyr
820 825 830

Pro Asp Asp Ser Cys Asn Gly Ile Asn Arg Cys Asp Glu Gln Lys Met
835 840 845

Val Asn Ala Gln Leu Glu Thr Glu His His His Pro Met Asp Cys Cys
850 855 860

Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp
865 870 875 880

Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg
885 890 895

Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val
900 905 910

Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn Ala
915 920 925

Lys Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Glu Arg Val
930 935 940

Tyr Tyr Ala Ala Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Arg
945 950 955 960

Asp Gln Gln Leu Asn Pro Gln Ile Gly Met Ala Asp Ile Met Asp Ala
965 970 975

Gln Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu
980 985 990

Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg
995 1000 1005

Leu Gln Gln Ala Ser Tyr Leu His Met Ser Arg Asn Ala Met Gln
1010 1015 1020

Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Ala
1025 1030 1035

Gly Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu
1040 1045 1050

Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro
1055 1060 1065

Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly
1070 1075 1080

Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu
1085 1090 1095

Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr
1100 1105 1110

Val Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe His Pro
1115 1120 1125

Glu Thr Gln His Met Trp Val Glu Val Ser Glu Thr Glu Gly Val
1130 1135 1140

Phe His Leu Asp Ser Val Glu Phe Met Glu Thr Gln Gln
1145 1150 1155

<210> 23

<211> 1156

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> BT-0013 đột biến

<400> 23

Met Asn Gln Asn Lys His Gly Ile Ile Gly Ala Ser Asn Cys Gly Cys
1 5 10 15

Ala Ser Asp Asp Val Ala Lys Tyr Pro Leu Ala Asn Asn Pro Tyr Ser
20 25 30

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

Cys Val His Trp Tyr Asn Arg Gly Phe Asn Glu Leu Arg Gln Arg Gly
 245 250 255

Thr Ser Ala Thr Ala Trp Leu Glu Phe His Arg Tyr Arg Arg Glu Met
 260 265 270

Thr Leu Met Val Leu Asp Ile Val Ala Ser Phe Ser Ser Leu Asp Ile
 275 280 285

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

Thr Asp Pro Ile Gly Phe Val His Arg Ser Ser Leu Arg Gly Glu Ser
 305 310 315 320

Trp Phe Ser Phe Val Asn Arg Ala Asn Phe Ser Asp Leu Glu Asn Ala
 325 330 335

Ile Pro Asn Pro Arg Pro Ser Trp Phe Leu Asn Asn Met Ile Ile Ser
 340 345 350

Thr Gly Ser Leu Thr Leu Pro Val Ser Pro Ser Thr Asp Arg Ala Arg
 355 360 365

Val Trp Tyr Gly Ser Arg Asp Arg Ile Ser Pro Ala Asn Ser Gln Phe
 370 375 380

Ile Thr Glu Leu Ile Ser Gly Gln His Thr Thr Ala Thr Gln Thr Ile
 385 390 395 400

Leu Gly Arg Asn Ile Phe Arg Val Asp Ser Gln Ala Cys Asn Leu Asn
 405 410 415

Asp Thr Thr Tyr Gly Val Asn Arg Ala Val Phe Tyr His Asp Ala Ser
 420 425 430

Glu Gly Ser Gln Arg Ser Val Tyr Glu Gly Tyr Ile Arg Thr Thr Gly
 435 440 445

Ile Asp Asn Pro Arg Val Gln Asn Ile Asn Thr Tyr Leu Pro Gly Glu
 450 455 460

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

Thr Ile Asn Leu Thr Gly Gly Leu Arg Gln Val Ala Ser Asn Arg Arg
 485 490 495

Ser Ser Leu Val Met Tyr Gly Trp Thr His Lys Ser Leu Ala Arg Asn
 500 505 510

Asn Thr Ile Asn Pro Asp Arg Ile Thr Gln Ile Pro Leu Thr Lys Val
 515 520 525

Asp Thr Arg Gly Thr Gly Val Ser Tyr Val Asn Asp Pro Gly Phe Ile
 530 535 540

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

Arg Val Gln Phe Pro Leu His Leu Arg Gln Gln Tyr Arg Ile Arg Val
 565 570 575

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

Gly Thr Ile Ser Gln Asn Leu Pro Ser Thr Met Arg Leu Gly Glu Asp
 595 600 605

Leu Arg Tyr Gly Ser Phe Ala Ile Arg Glu Phe Asn Thr Ser Ile Arg
 610 615 620

Pro Thr Ala Ser Pro Asp Gln Ile Arg Leu Thr Ile Glu Pro Ser Phe
 625 630 635 640

Ile Arg Gln Glu Val Tyr Val Asp Arg Ile Glu Phe Ile Pro Val Asn
 645 650 655

Pro Thr Arg Glu Ala Lys Glu Asp Leu Glu Ala Ala Lys Lys Ala Val
 660 665 670

Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Lys
 675 680 685

Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp
 690 695 700

Glu Gln Tyr Gly Tyr Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala
 705 710 715 720

Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe
 725 730 735

Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly
 740 745 750

Val Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Ile Gln
 755 760 765

Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val
 770 775 780

Asp Ala Ser Glu Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe
 785 790 795 800

Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys
 805 810 815

Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Leu Asp Thr Tyr
 820 825 830

Pro Asp Asp Ser Cys Asn Gly Ile Asn Arg Cys Asp Glu Gln Lys Met
 835 840 845

Val Asn Ala Gln Leu Glu Thr Gly His His His Pro Met Asp Cys Cys
 850 855 860

Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp
 865 870 875 880

Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg
 885 890 895

Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val
 900 905 910

Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn Ala
 915 920 925

Lys Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Glu Arg Val
 930 935 940

Tyr Tyr Ala Ala Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Gln
 945 950 955 960

Asp Gln Gln Leu Asn Pro Gln Ile Gly Met Ala Asp Ile Met Asp Ala
 965 970 975

Gln Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu
 980 985 990

Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg
 995 1000 1005

Leu Gln Gln Ala Ser Tyr Leu His Thr Ser Arg Asn Ala Met Gln
 1010 1015 1020

Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Ala
 1025 1030 1035

Gly Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu
 1040 1045 1050

Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro
 1055 1060 1065

Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly
 1070 1075 1080

Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu
 1085 1090 1095

Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr
 1100 1105 1110

Val Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe His Pro
 1115 1120 1125

Glu Thr Gln His Met Trp Val Glu Val Ser Glu Thr Glu Gly Val
 1130 1135 1140

Phe His Leu Asp Ser Val Glu Phe Met Glu Thr Gln Gln
 1145 1150 1155

<210> 24
 <211> 1152
 <212> PRT
 <213> Trình tự Nhân tạo

<220>
 <223> BT-0023 đột biến

<400> 24

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

Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp
 20 25 30

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

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

Asn Pro Arg Asp Val Leu Gln Thr Gly Ile Asn Ile Val Gly Arg Ile
 65 70 75 80

Leu Gly Phe Leu Gly Val Pro Phe Ala Gly Gln Leu Val Thr Phe Tyr
 85 90 95

Thr Phe Leu Leu Asn Gln Leu Trp Pro Thr Asn Asp Asn Ala Val Trp
 100 105 110

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

Ala Gln Val Val Arg Asn Ala Leu Asp Asp Leu Thr Gly Leu His Asp
 130 135 140

Tyr Tyr Glu Glu Tyr Leu Ala Ala Leu Glu Glu Trp Leu Glu Arg Pro
 145 150 155 160

Asn Gly Ala Arg Ala Asn Leu Val Thr Gln Arg Phe Glu Asn Leu His
 165 170 175

Thr Ala Phe Val Thr Arg Met Pro Ser Phe Gly Thr Gly Pro Gly Ser
 180 185 190

Gln Arg Asp Ala Val Ala Leu Leu Thr Val Tyr Ala Gln Ala Ala Asn
 195 200 205

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

Gly Leu Gln Gln Gly Gln Ile Asn Leu Tyr Phe Asn Ala Gln Gln Glu
 225 230 235 240

Arg Thr Arg Ile Tyr Thr Asn His Cys Val Glu Thr Tyr Asn Arg Gly
 245 250 255

Leu Glu Asp Val Arg Gly Thr Asn Thr Glu Ser Trp Leu Asn Tyr His
 260 265 270

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

Asn Gln Gly Ile Cys Arg Arg Trp Gly Asn Asn Pro Tyr Asn Thr Phe
 325 330 335

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

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

Asn Ser Tyr Leu Asp Tyr Trp Ser Gly His Thr Leu Gln Ser Gln Tyr
 370 375 380

Ala Asn Asn Pro Thr Thr Tyr Glu Thr Ser Tyr Gly Gln Ile Thr Ser
 385 390 395 400

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

Arg Ala Arg Asn Phe Gly Asn Leu Tyr Ala Asn Leu Tyr Gly Val Ser
 420 425 430

Ser Leu Asn Ile Phe Pro Thr Gly Val Met Ser Glu Ile Thr Ser Ala
 435 440 445

Pro Asn Thr Cys Trp Gln Asp Leu Thr Thr Thr Glu Glu Leu Pro Leu
 450 455 460

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

Asn Thr Thr Gln Gly Gly Pro Leu Ala Thr Val Gly Phe Val Pro Thr
 485 490 495

Tyr Val Trp Thr Arg Gln Asp Val Asp Phe Asn Asn Ile Ile Thr Pro
 500 505 510

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

Gly Ala Thr Val Val Lys Gly Pro Gly Phe Thr Gly Gly Asp Val Ile
 530 535 540

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

Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr Ala Ser Thr
 565 570 575

Ile Asp Phe Asp Phe Phe Val Thr Arg Gly Gly Thr Thr Ile Asn Asn
 580 585 590

Phe Arg Phe Thr Arg Thr Met Asn Arg Gly Gln Glu Ser Arg Tyr Glu
 595 600 605

Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe Thr Gln Ser
 610 615 620

Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn Gly Glu
 625 630 635 640

Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro Thr Arg Glu
 645 650 655

Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala Ser Leu Phe
 660 665 670

Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp Tyr Gln Val
 675 680 685

Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu Gln Tyr Ala
 690 695 700

His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala Lys Arg Leu
 705 710 715 720

Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn Thr Ile Asn
 725 730 735

Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val Thr Ile Ser
 740 745 750

Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Leu Gln Leu Ala Ser Ala
 755 760 765

Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val Asp Ala Ser Glu
 770 775 780

Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe Val Lys Ser Ser
 785 790 795 800

Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys Val His Leu Val
 805 810 815

Lys Asn Val Leu Asp Asn Leu Val Ser Asp Thr Tyr Pro Asp Asp Ser
 820 825 830

Cys Ser Gly Ile Asn Arg Cys Glu Glu Gln Gln Met Val Asn Ala Gln
 835 840 845

Leu Glu Thr Glu His His His Pro Met Asp Cys Cys Glu Ala Ala Gln
 850 855 860

Thr His Glu Phe Ser Ser Tyr Ile Asp Thr Gly Asp Leu Asn Ser Thr
 865 870 875 880

Val Asp Gln Gly Ile Trp Val Ile Phe Lys Val Arg Thr Thr Asp Gly
 885 890 895

Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val Gly Pro Leu Leu
 900 905 910

Gly Glu Pro Leu Glu Arg Glu Gln Arg Glu Asn Ala Lys Trp Asn Ala
 915 920 925

Glu Leu Gly Arg Lys Arg Ala Glu Thr Asp Arg Val Tyr Gln Asp Ala
 930 935 940

Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Gln Asp Gln Gln Leu
 945 950 955 960

Asn Pro Gln Ile Gly Met Ala Asp Ile Met Asp Ala Gln Asn Leu Val
 965 970 975

Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly
 980 985 990

Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln Ala
 995 1000 1005

Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val Gln Asn Gly Asp Phe
 1010 1015 1020

Asn Asn Gly Leu Asp Ser Trp Asn Ala Thr Ala Gly Ala Ser Val
 1025 1030 1035

Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu Ser His Trp Asp
 1040 1045 1050

Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys Tyr
 1055 1060 1065

Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly Gly Asp Gly Tyr
 1070 1075 1080

Val Thr Ile Arg Asp Gly Ala His His Thr Glu Thr Leu Thr Phe
 1085 1090 1095

Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr Val Thr Asp Asn
 1100 1105 1110

Thr Tyr Leu Thr Lys Glu Val Leu Phe Tyr Ser His Thr Glu His
 1115 1120 1125

Met Trp Val Glu Val Asn Glu Thr Glu Gly Ala Phe His Leu Asp
 1130 1135 1140

Ser Ile Glu Phe Val Glu Thr Glu Lys
 1145 1150

<210> 25

<211> 1155

<212> PRT

<213> Trình tự Nhân tạo

<220>

<223> BT-0067 đột biến

<400> 25

Met Asn Arg Asn Asn Gln Asp Glu Tyr Glu Ile Ile Asp Ala Ser Thr
 1 5 10 15

Cys Gly Cys Ser Ser Asp Asp Val Val Gln Tyr Pro Leu Ala Arg Asp
 20 25 30

Pro Asn Ala Val Phe Gln Asn Met His Tyr Lys Asp Tyr Leu Gln Thr
 35 40 45

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

Asn Pro Arg Asp Val Leu Gln Thr Gly Ile Asn Ile Val Gly Arg Leu
 65 70 75 80

Leu Gly Phe Leu Gly Val Pro Phe Ala Gly Gln Leu Val Thr Phe Tyr
 85 90 95

Thr Phe Leu Leu Asn Gln Leu Trp Pro Thr Asn Asp Asn Ala Val Trp
 100 105 110

Glu Ala Phe Met Ala Gln Ile Glu Glu Leu Ile Asn Gln Arg Ile Ser
 115 120 125

Glu Ala Val Val Gly Thr Ala Ala Asp His Leu Thr Gly Leu His Asp
 130 135 140

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

Asn Ala Ala Arg Thr Asn Leu Leu Phe Asn Arg Phe Thr Thr Leu Asp
 165 170 175

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

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

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

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

Arg Thr Gln Tyr Tyr Thr Asn His Cys Val Thr Thr Tyr Asn Thr Gly
 245 250 255

Leu Asp Arg Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Asn Tyr His
 260 265 270

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

Asn Val Gly Leu Cys Arg Arg Trp Gly Asn Asn Pro Tyr Asn Arg Phe
 325 330 335

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

Leu Asn Thr Leu Thr Ile Ser Arg Asn Arg Phe Asp Val Gly Ser Asn
 355 360 365

Phe Ile Glu Pro Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Ser Asn
 370 375 380

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

Thr Thr Ile Asn Leu Pro Ala Asn Gly Thr Gly Arg Val Glu Ser Thr
 405 410 415

Ala Val Asp Phe Arg Ser Ala Leu Val Gly Ile Tyr Gly Val Asn Arg
420 425 430

Ala Ser Phe Ile Pro Gly Gly Val Phe Ser Gly Thr Thr Gln Pro Ser
435 440 445

Thr Gly Gly Cys Arg Asp Leu Tyr Asp Ser Ser Asp Glu Leu Pro Pro
450 455 460

Asp Glu Ser Thr Gly Ser Phe Ala His Arg Leu Ser His Val Thr Phe
465 470 475 480

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

Val Pro Thr Tyr Val Trp Thr His Arg Asp Val Asp Phe Asn Asn Thr
500 505 510

Ile Asn Pro Asn Arg Ile Thr Gln Ile Pro Val Val Lys Ala Tyr Glu
515 520 525

Leu Ser Ser Gly Ala Thr Val Val Lys Gly Pro Gly Phe Thr Gly Gly
530 535 540

Asp Val Ile Arg Arg Thr Asn Ile Gly Gly Phe Gly Ala Ile Arg Val
545 550 555 560

Ser Val Thr Gly Pro Leu Thr Gln Arg Tyr Arg Ile Arg Phe Arg Tyr
565 570 575

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

Ile Asn Asn Phe Arg Phe Thr Arg Thr Met Asn Arg Gly Gln Glu Ser
595 600 605

Arg Tyr Glu Ser Tyr Arg Thr Val Glu Phe Thr Thr Pro Phe Asn Phe
610 615 620

Thr Gln Ser Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly
625 630 635 640

Asn Gly Glu Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Asn Pro
 645 650 655

Thr Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys Ala Val Ala
 660 665 670

Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn Val Thr Asp
 675 680 685

Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu Ser Asp Glu
 690 695 700

Gln Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala Val Arg Ala Ala
 705 710 715 720

Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro Asp Phe Asn
 725 730 735

Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser Asn Gly Val
 740 745 750

Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg Ala Leu Gln Leu
 755 760 765

Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln Lys Val Asn
 770 775 780

Ala Ser Glu Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp Gly Phe Val
 785 790 795 800

Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His His Lys Val
 805 810 815

His Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser Asp Thr Tyr Ser
 820 825 830

Asp Gly Ser Cys Ser Gly Met Asn Arg Cys Glu Glu Gln Gln Met Val
 835 840 845

Asn Ala Gln Leu Glu Thr Glu His His His Pro Met Asp Cys Cys Glu
850 855 860

Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly Asp Leu
865 870 875 880

Asn Ser Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val Arg Thr
885 890 895

Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu Val Gly
900 905 910

Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn Ala Lys
915 920 925

Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Asp Arg Val Tyr
930 935 940

Gln Asp Ala Lys Gln Ser Ile Asn His Leu Phe Val Asp Tyr Gln Asp
945 950 955 960

Gln Gln Leu Asn Pro Glu Ile Gly Met Ala Asp Ile Ile Asp Ala Gln
965 970 975

Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp Ala Val Leu Gln
980 985 990

Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asn Arg Leu
995 1000 1005

Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val Gln Asn
1010 1015 1020

Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr Gly Gly
1025 1030 1035

Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu Val Leu Ser
1040 1045 1050

His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val Gln Pro Asn
1055 1060 1065

Cys Lys Tyr Val Leu Arg Val Thr Ala Glu Lys Val Gly Gly Gly
 1070 1075 1080

Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Thr Glu Lys
 1085 1090 1095

Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly Thr Tyr Val
 1100 1105 1110

Thr Asp Asn Thr Tyr Ile Thr Lys Glu Val Val Phe Tyr Ser His
 1115 1120 1125

Thr Glu His Met Trp Val Glu Val Ser Glu Thr Glu Gly Ala Phe
 1130 1135 1140

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

<210> 26

<211> 26

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Đoạn mỗi xuôi OAR2613a

<400> 26

aaacatgaac cgaaataatc aaaatg

26

<210> 27

<211> 22

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> Đoạn mỗi ngược OAR2615a

<400> 27

atccgtccct tgtgcgtgta aa

22

<210> 28

<211> 30

<212> ADN
 <213> Trình tự Nhân tạo

<220>
 <223> Đoạn mỗi xuôi OAR2611a-F

<400> 28
 gtttaaacaat gaatcgaaat aatcaaaatg
 30

<210> 29
 <211> 29
 <212> ADN
 <213> Trình tự Nhân tạo

<220>
 <223> Đoạn mỗi ngược OAR2612a-R

<400> 29
 ggcgcgcctt actcttgtgt ttcaataaa
 29

<210> 30
 <211> 29
 <212> ADN
 <213> Trình tự Nhân tạo

<220>
 <223> Đoạn mỗi xuôi OAR2768-F

<400> 30
 gtttaaacaat gaatcaaaat aaacacgga
 29

<210> 31
 <211> 31
 <212> ADN
 <213> Trình tự Nhân tạo

<220>
 <223> Đoạn mỗi ngược OAR2769-R

<400> 31
 ggcgcgcctt actgttgggt ttccatgaac t
 31