(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) **1-0028885**

(51)[8] **C07K 14/325**; C12N 15/82

(13) **B**

(21) 1-2017-01416
(22) 15/10/2015
(86) PCT/US2015/055800 15/10/2015
(87) WO2016/061391 21/04/2016
(30) 62/064,989 16/10/2014 US
(45) 25/07/2021 400
(43) 25/07/2017 352A
(73) MONSANTO TECHNOLOGY LLC (US)
800 N. Lindbergh Boulevard, Mail Zone E1NA, St. Louis, Missouri 63167, United States of America
(72) BAUM, James A. (US); DART, Crystal L. (US); ENGLISH, Leigh H. (US); FU, Xiaoran (US); GUZOV, Victor M. (US); HOWE, Arlene R. (US); CERRUTI, Thomas A. (US); MORGENSTERN, Jay P. (US); ROBERTS, James K. (US); SALVADOR, Sara A. (US); WANG, Jinling (US); FLASINSKI, Stanislaw (US).
(74) Công ty Luật TNHH T&G (TGVN)

(54) PROTEIN KHẢM DIỆT CÔN TRÙNG, POLYNUCLEOTIT MÃ HÓA PROTEIN KHẢM,CHẾ PHẨM ỨC CHẾ CÔN TRÙNG CHỨA PROTEIN KHẢM VÀ PHƯƠNG PHÁP PHÒNG TRỪ SINH VẬT GÂY HẠI BỌ CÁNH VẢY

(57) Sáng chế đề cập đến protein khảm có hoạt tính diệt côn trùng thể hiện hoạt tính ức chế bọ cánh vảy và chế phẩm chứa các protein này. Sáng chế cũng đề xuất polynucleotit mã hóa các protein theo sáng chế và tế bào chủ chứa polynucleotit này. Sáng chế cũng đề xuất tế bào thực vật chuyển gen, cây chuyển gen, các bộ phận của cây, và hạt chứa protein khảm diệt côn trùng, phương pháp sản xuất hạt này và phương pháp phòng trừ sinh vật gây hại bộ cánh vảy.

Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế liên quan đến lĩnh vực các protein ức chế côn trùng. Sáng chế đề cập đến protein khảm diệt côn trùng thể hiện hoạt tính ức chế côn trùng đối với các sinh vật gây hại cho cây trồng và hạt trong nông nghiệp. Cụ thể, nhóm protein được bộc lộ thể hiện hoạt tính diệt côn trùng đối với bộ gồm các côn trùng gây hại là bọ cánh vảy. Cây, các bộ phận của cây, và hạt chứa phân tử axit nucleic tái tổ hợp mã hóa một hoặc nhiều protein của độc tố này cũng được đề xuất.

Tình trạng kỹ thuật của sáng chế

Việc cải thiện năng suất trồng trọt của các cây nông nghiệp quan trọng bao gồm, trong số nhiều cây khác, cây ngô, cây đậu tương, cây mía, cây lúa, cây lúa mỳ, rau, và cây bông, ngày càng trở nên quan trọng. Ngoài nhu cầu gia tăng về các sản phẩm nông nghiệp để cung cấp thức ăn, vải mặc và năng lượng cho dân số ngày càng tăng, tác động và áp lực liên quan đến khí hậu do dân số gia tăng sử dụng đất cho các mục đích phi nông nghiệp được dự đoán làm giảm lượng đất canh tác được để trồng trọt. Các yếu tố này dẫn đến dự báo nghiêm trọng về an ninh lương thực, đặc biệt là khi không có cải thiện đáng kể về công nghệ sinh học thực vật và thực tiễn nông nghiệp. Dưới các áp lực này, cải thiện bền vững về mặt môi trường trong công nghệ, các kỹ thuật nông nghiệp, và phòng diệt sinh vật gây hại là các công cụ sống còn để mở rộng sản lượng trồng trọt trên lượng hạn chế đất canh tác được để trồng trọt.

Côn trùng, đặc biệt là các côn trùng thuộc bộ cánh vảy, được xem nguyên nhân chính phá hoại cây trồng nông nghiệp, do đó làm giảm sản lượng ở các diện tích bị phá hoại. Loài sinh vật gây hại bộ cánh vảy mà ảnh hưởng tiêu cực đến nông nghiệp bao gồm, nhưng không chỉ giới hạn ở, sâu xanh mùa thu (*Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (*Spodoptera exigua*), sâu ăn lá bertha (*Mamestra configurata*), sâu xám (*Agrotis ipsilon*), sâu đo ký sinh trên cải bắp (*Trichoplusia ni*), sâu đo đậu tương (*Chrysodeixis includens*), sâu bướm hại đậu mèo (*Anticarsia gemmatalis*), sâu xanh (*Hypena scabra*), sâu chồi thuốc lá (*Heliothis virescens*), sâu ngài đêm dạng hạt (*Agrotis subterranea*), sâu ăn lá (*Pseudaletia unipuncta*), sâu ngài đêm tây (*Agrotis*

orthogonia), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), bướm hại cam navel (*Amyelois transitella*), sâu kéo màng hại rễ ngô (*Crambus caliginosellus*), sâu kéo màng ở cỏ (*Herpetogramma licarsisalis*), sâu bướm hướng dương (*Homoeosoma electellum*), sâu đục thân ngô nhỏ (*Elasmopalpuslignosellus*), bướm cá tuyết (*Cydia pomonella*), bướm hại cây việt quất (*Endopiza viteana*), bướm hại quả phương đông (*Grapholita molesta*), bướm hại chồi hướng dương (*Suleima helianthana*), sâu tơ (*Plutella xylostella*), sâu hồng đục quả (*Pectinophora gossypiella*), sâu đục thân hồng (*Sesamia inferens*), sâu bướm gypsy (*Lymantria dispar*), sâu hại lá cây bông (*Alabama argillacea)*, sâu cuốn lá cây ăn quả (*Archips argyrospila)*, sâu cuốn lá châu Âu (*Archips rosana)*, sâu đục thân lúa châu Á, hoặc sâu đục thân lúa (*Chilo suppressalis)*, sâu cuốn lá lúa (*Cnaphalocrocis medinalis)*, sâu kéo màng hại rễ ngô (*Crambus caliginosellus)*, sâu kéo màng ở cỏ xanh (*Crambus teterrellus)*, sâu đục thân ngô tây nam (*Diatraea grandiosella)*), sâu đục thân mía (*Diatraea saccharalis)*, sâu đục quả có gai (*Earias insulana)*, sâu đục quả đốm (*Earias vittella)*, sâu hại bông cổ (*Helicoverpa armigera)*, sâu đục trái hại ngô, sâu hại đậu tương hoặc sâu hại bông (*Helicoverpa zea)*, sâu kéo màng ở cỏ (*Herpetogramma licarsisalis)*, bướm hại nho châu Âu (*Lobesia botrana)*, sâu ăn lá quả có múi (*Phyllocnistis citrella)*, bướm trắng to (*Pieris brassicae)*, sâu hại cải bắp nhập khẩu, hoặc bướm trắng nhỏ (*Pieris rapae)*, sâu ngài đêm hại thuốc lá, hoặc sâu bướm cụm (*Spodoptera litura)*, và sâu ăn lá cà chua (*Tuta absoluta)*.

Trước đây, việc sử dụng nhiều chất hóa học tổng hợp diệt côn trùng được tin tưởng làm tác nhân phòng diệt sinh vật gây hại trong nông nghiệp. Ngoài các vấn đề về sự hình thành tính kháng, quan ngại về mặt môi trường và sức khỏe của con người kích thích việc nghiên cứu và phát triển các thuốc sinh học diệt sinh vật gây hại. Cố gắng nghiên cứu này dẫn tới sự phát hiện dần dần và việc sử dụng các loài vi sinh vật khác nhau, kể cả vi khuẩn, gây bệnh cho côn trùng.

Mô hình phòng trừ sinh học dịch chuyển khi tiềm năng của vi khuẩn gây bệnh cho côn trùng, đặc biệt là vi khuẩn thuộc giống *Bacillus*, được phát hiện và được phát triển làm tác nhân sinh học phòng diệt sinh vật gây hại. Chủng vi khuẩn *Bacillus thuringiensis* (*Bt*) được dùng làm nguồn protein diệt côn trùng vì đã phát hiện ra rằng các chủng *Bt* thể hiện độc tính cao đối với côn trùng cụ thể. Đã biết rằng các chủng *Bt* tạo ra nội độc tố delta mà chúng nằm trong các thể vùi tinh thể bào tử vỏ khi bắt đầu

hình thành bào tử và trong giai đoạn sinh trưởng ổn định (ví dụ, các protein Cry), và cũng đã biết rằng chủng này sản sinh protein diệt côn trùng tiết ra được. Khi côn trùng mẫn cảm ăn vào, nội độc tố delta cũng như các độc tố tiết ra tác động đến bề mặt của biểu mô ruột giữa, phá vỡ màng tế bào, dẫn đến sự phá vỡ tế bào và chết. Gen mã hóa các protein diệt côn trùng cũng đã được xác định ở các loài vi khuẩn khác với *Bt*, kể cả *Bacillus* khác và các loài vi khuẩn đa dạng khác, như *Brevibacillus laterosporus, Lysinibacillus sphaericus* ("*Ls*" trước đây được gọi là *Bacillus sphaericus*) và *Paenibacillus popilliae*.

Các độc tố protein diệt côn trùng hòa tan ở dạng tinh thể hoặc dạng tiết ra được là rất đặc hiệu với vật chủ của chúng và đã được chấp nhận rộng rãi trên toàn thế giới làm các phương án thay thế cho hóa chất diệt côn trùng. Ví dụ, các protein độc tố diệt côn trùng đã được sử dụng trong các ứng dụng nông nghiệp khác nhau để bảo vệ các cây trồng nông nghiệp quan trọng khỏi sự phá hoại của côn trùng, làm giảm nhu cầu dùng hóa chất diệt sinh vật gây hại, và tăng sản lượng. Các protein độc tố diệt côn trùng được dùng để phòng trừ các sinh vật gây hại trong nông nghiệp ở cây trồng theo các phương pháp cơ học, như xịt để phân tán các chế phẩm vi khuẩn chứa nhiều chủng vi khuẩn khác nhau lên bề mặt của cây, và bằng cách áp dụng các kỹ thuật biến nạp di truyền để tạo ra các cây chuyển gen và hạt biểu hiện protein độc tố diệt côn trùng.

Việc sử dụng các cây chuyển gen biểu hiện các protein diệt côn trùng đã được làm thích ứng trên toàn cầu. Ví dụ, trong năm 2012, 26,1 triệu hecta được trồng cây chuyển gen biểu hiện các độc tố *Bt* (James, C., *Global Status of Commercialized Biotech/GM Crops*: 2012. ISAAA Brief No. 44). Việc sử dụng trên toàn cầu các cây trồng chuyển gen đã được bảo vệ đối với côn trùng và số lượng hạn chế protein diệt côn trùng được dùng ở các cây trồng này tạo ra áp lực chọn lọc đối với các alen côn trùng hiện có mà truyền tính kháng protein diệt côn trùng hiện dùng.

Việc phát triển tính kháng ở sinh vật gây hại đích đối với các protein diệt côn trùng tạo ra nhu cầu tiếp tục cần khám phá và phát triển các dạng protein diệt côn trùng mới mà là hữu ích để quản lý hiện tượng tăng mức kháng của côn trùng đối với các cây chuyển gen biểu hiện các protein diệt côn trùng. Các protein diệt côn trùng mới với hiệu quả tốt hơn và thể hiện tác dụng phòng trừ đối với phổ rộng hơn gồm các loài côn trùng mẫn cảm sẽ làm giảm số lượng côn trùng sống sót mà có thể hình thành các alen kháng. Ngoài ra, việc sử dụng ở một cây hai hoặc nhiều protein diệt côn trùng

chuyển gen độc đối với cùng một côn trùng gây hại và thể hiện các kiểu hoạt động khác nhau làm giảm xác suất kháng ở loài côn trùng đích bất kỳ.

Bởi vậy, rất cần xác định các protein diệt côn trùng bổ sung với các tính chất diệt côn trùng tốt hơn như hiệu quả cao hơn đối với phổ rộng hơn bao gồm các loài côn trùng gây hại đích và cách thức hoạt động khác nhau so với các độc tố hiện được dùng trong thực tiễn nông nghiệp. Để đáp ứng nhu cầu này, sáng chế bộc lộ các protein khảm diệt côn trùng Cry1 mới mà thể hiện hoạt tính đối với loài sinh vật gây hại bộ cánh vảy đích đáng kể.

Trong lĩnh vực kỹ thuật này, đã biết rằng các thành viên của họ protein tinh thể Cry1 thể hiện hoạt tính sinh học đối với sinh vật gây hại bộ cánh vảy. Dạng tiền chất của các protein tinh thể Cry1 bao gồm hai phân đoạn có kích cỡ gần bằng nhau. Đầu tận cùng carboxyl của protein tiền chất, được gọi là phân đoạn tiền độc tố, làm ổn định quá trình tạo tinh thể và không thể hiện hoạt tính diệt côn trùng. Đầu tận cùng amin của protein tiền chất chứa phân đoạn độc tố của protein Cry1 và, trên cơ sở việc duỗi thẳng trình tự bảo toàn hoặc trình tự cơ bản được bảo toàn ở các thành viên của họ Cry1, có thể được phân chia tiếp nữa thành ba miền cấu trúc, miền I, miền II, và miền III. Miền I chứa khoảng một phần ba đầu tiên của phân đoạn độc tố có hoạt tính và đã được thấy là quan trọng đối với sự hình thành rãnh. Cả miền II và miền III đều liên quan đến liên kết của thụ thể và tính đặc hiệu theo loài côn trùng, tùy theo côn trùng và protein diệt côn trùng được xem xét.

Khả năng tùy ý tạo ra protein khảm với các tính chất tốt hơn từ việc phân loại các cấu trúc miền của nhiều protein diệt côn trùng tự nhiên đã biết trong lĩnh vực kỹ thuật này là mong manh. Đó là do bản chất phức tạp của cấu trúc protein, sự oligome hóa, và sự hoạt hóa (kể cả xử lý phân giải protein chính xác tiền chất khảm, nếu được biểu hiện ở dạng này) cần để giải phóng phân đoạn protein có hoạt tính diệt côn trùng. Chỉ bằng cách chọn lọc cẩn thận các tiền độc tố và các đích cụ thể trong từng protein gốc để tạo ra cấu trúc khảm thì các độc tố khảm chức năng có hoạt tính diệt côn trùng có thể được xây dựng mà thể hiện hoạt tính diệt côn trùng tốt hơn so với các protein gốc mà từ đó tạo ra thể khảm. Trong lĩnh vực kỹ thuật này, đã biết rằng việc ghép lại tiền độc tố và các miền độc tố I, II và III của hai hoặc nhiều độc tố bất kỳ mà là khác nhau thường dẫn đến việc tạo ra các protein mà thể hiện sự hình thành tinh thể lỗi hoặc hoàn toàn không có hoạt tính diệt côn trùng ở mức độ có thể phát hiện được hướng đến các

loài côn trùng gây hại đích được ưu tiên. Chỉ bằng cách thử và lỗi mà các thể khảm diệt côn trùng hữu hiệu được thiết kế, và thậm chí khi đó, người có hiểu biết trung bình về lĩnh vực kỹ thuật này không chắc chắn thu được thể khảm thể hiện hoạt tính diệt côn trùng mà là tương đương hoặc tốt hơn so với protein độc tố gốc đơn lẻ bất kỳ mà từ đó có thể thu được tiền độc tố hoặc các miền độc tố cấu thành thể khảm. Ví dụ, tài liệu chyuyên ngành đã thông báo nhiều ví dụ về cấu trúc hoặc cụm protein khảm từ hai hoặc nhiều tiền chất protein tinh thể. Xem tài liệu: Jacqueline S. Knight, *et al.* "A Strategy for Shuffling Numerous *Bacillus thuringiensis* Crystal Protein Domains." *J. Economic Entomology*, 97 (6) (2004): 1805-1813; Bosch và các đồng tác giả (US 6,204,246); Malvar và Gilmer (US 6,017,534). Ở mỗi ví dụ trong số các ví dụ này, nhiều thể khảm thu được không thể hiện tính chất diệt côn trùng hoặc tính chất tạo ra tinh thể mà là tương đương hoặc tốt hơn so với các protein tiền chất mà từ đó thu được các thành phần của thể khảm.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất phân tử axit nucleic tái tổ hợp mà mã hóa protein khảm diệt côn trùng có độc tính đối với loài sinh vật gây hại là bọ cánh vảy. Mỗi protein khảm diệt côn trùng có thể được sử dụng riêng rẽ hoặc kết hợp với nhau và với các protein diệt côn trùng khác và các tác nhân ức chế côn trùng trong các chế phẩm và trên cây; nhờ đó tạo ra các phương án thay thế cho các protein diệt côn trùng và các hóa chất diệt côn trùng hiện đang được dùng trong các hệ nông nghiệp.

Theo một số phương án nhất định, sáng chế đề xuất protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ đã nêu trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, hoặc 111. Protein khảm diệt côn trùng này thể hiện hoạt tính ức chế đối với loài côn trùng thuộc bộ cánh vảy như, nhưng không chỉ giới hạn ở, *Anticarsia gemmatalis*, *Diatraea saccharalis*, *Elasmopalpus lignosellus*, *Helicoverpa zea*, *Heliothis virescens*, *Chrysodeixis includens*, *Spodoptera cosmioides*, *Spodoptera eridania*, *Spodoptera frugiperda*, *Spodoptera exigua*, *Helicoverpa armigera*, *Spodoptera litura*, *Pectinophora gossypiella*, *Diatraea grandiosella*, *Earias vitella*, *Helicoverpa gelotopeon* và *Rachiplusia nu*.

Theo phương án khác, sáng chế đề xuất polynucleotit mã hóa protein khảm diệt côn trùng, trong đó polynucleotit được liên kết hoạt động với trình tự khởi đầu khác loại và protein khảm diệt côn trùng chứa trình tự axit amin đã nêu ở trình tự bất kỳ trong số SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109 hoặc 111. Polynucleotit mã hóa protein khảm diệt côn trùng, trong đó polynucleotit chứa trình tự nucleotit mà tùy ý: lai trong các điều kiện nghiêm ngặt với thể bổ khuyết ngược của trình tự polynucleotit bất kỳ đã nêu trong số các trình tự nêu trong SEQ ID NO:1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, hoặc 130; hoặc mã hóa protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109 hoặc 111, cũng được dự định.

Theo các phương án khác được bộc lộ, sáng chế đề xuất tế bào chủ chứa polynucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO: 1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129 hoặc 130, trong đó tế bào chủ được chọn từ nhóm bao gồm tế bào chủ vi khuẩn hoặc tế bào chủ thực vật. Vật chủ vi khuẩn được dự định bao gồm *Agrobacterium, Rhizobium*, *Bacillus*, *Brevibacillus*, *Escherichia*, *Pseudomonas*, *Klebsiella*, và *Erwinia*; và trong đó loài *Bacillus* là *Bacillus cereus* hoặc *Bacillus thuringiensis*, *Brevibacillus* là *Brevibacillus laterosperous*, và *Escherichia* là *Escherichia coli.* Các tế bào thực vật được dự định bao gồm loại một lá mầm và hai lá mầm.

Các phương án khác đã được bộc lộ trong bản mô tả này là chế phẩm ức chế côn trùng chứa protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41,

43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109 hoặc 111. Theo các phương án nhất định, chế phẩm ức chế côn trùng còn chứa ít nhất một tác nhân ức chế côn trùng khác với protein khảm diệt côn trùng. Các tác nhân ức chế côn trùng dự định khác với protein khảm diệt côn trùng bao gồm protein ức chế côn trùng, phân tử ARN sợi đôi ức chế côn trùng, và hóa chất ức chế côn trùng. Các tác nhân ức chế côn trùng khác với protein khảm diệt côn trùng này có thể thể hiện hoạt tính đối với một hoặc nhiều loài sinh vật gây hại thuộc bộ cánh vảy, bọ cánh cứng, bọ cánh nửa, bọ cánh giống, hoặc bọ cánh tơ.

Theo phương án khác nữa, sáng chế đề xuất hạt chứa lượng có tác dụng ức chế côn trùng của: protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, hoặc 111; hoặc polynucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO:1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129 hoặc 130.

Các phương pháp phòng trừ sinh vật gây hại bộ cánh vảy bao gồm bước cho sinh vật gây hại bộ cánh vảy tiếp xúc với protein khảm diệt côn trùng theo sáng chế với lượng có tác dụng ức chế cũng được dự định.

Theo phương án khác, sáng chế đề xuất tế bào thực vật, cây hoặc bộ phận của cây chuyển gen chứa protein khảm diệt côn trùng, trong đó: protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, hoặc 111; hoặc protein khảm diệt côn trùng chứa protein có: mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:21, 10; mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:28; mức độ đồng nhất ít nhất 87% so với trình tự SEQ ID NO:7; mức độ đồng nhất ít nhất 90% so với SEQ ID NO:4; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:13; mức độ đồng nhất ít nhất 64% so với trình tự SEQ ID

NO:16; mức độ đồng nhất ít nhất 66% so với trình tự SEQ ID NO:19; mức độ đồng nhất ít nhất 86% so với trình tự SEQ ID NO:23; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:25; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:30; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:33; mức độ đồng nhất ít nhất 64% so với trình tự SEQ ID NO:36; mức độ đồng nhất ít nhất 66% so với trình tự SEQ ID NO:39; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:41; mức độ đồng nhất ít nhất 84% so với trình tự SEQ ID NO:43; mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:45; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO: 47; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:50; hoặc mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:53; hoặc mức độ đồng nhất ít nhất 87% so với các SEQ ID NO:85, 93, 105; hoặc mức độ đồng nhất ít nhất 85% so với các SEQ ID NO:55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79; hoặc mức độ đồng nhất ít nhất 88% so với các SEQ ID NO: 91, 87,89; hoặc mức độ đồng nhất ít nhất 89% so với các SEQ ID NO: 107, 111; hoặc mức độ đồng nhất ít nhất 90% so với SEQ ID NO: 97; mức độ đồng nhất ít nhất 91% so với SEQ ID NO:109; hoặc mức độ đồng nhất ít nhất 93% so với SEQ ID NO:83; hoặc mức độ đồng nhất ít nhất 94% so với các SEQ ID NO:91 hoặc 103; hoặc mức độ đồng nhất ít nhất 95% so với các SEQ ID NO:95, 101; hoặc mức độ đồng nhất ít nhất 98% so với SEQ ID NO:99. Các phương pháp phòng trừ sinh vật gây hại bộ cánh vảy mà bao gồm việc cho sinh vật gây hại tiếp xúc với tế bào của thực vật chuyển gen, cây chuyển gen hoặc bộ phận của cây này, trong đó tế bào thực vật, cây hoặc phần của cây đó biểu hiện protein khảm diệt côn trùng với lượng có tác dụng ức chế bọ cánh vảy cũng được dự định.

Theo các phương án khác của sáng chế, các sản phẩm hàng hóa thu được từ tế bào thực vật, cây, hoặc phần của cây, trong đó sản phẩm chứa protein khảm diệt côn trùng với lượng phát hiện được được đề xuất. Các sản phẩm hàng hóa được dự định bao gồm sinh khối thực vật, dầu, bột xay thô, thức ăn chăn nuôi, bột, cốm, cám, thớ, vỏ, và hạt đã được xử lý.

Phương pháp khác nữa đã được bộc lộ trong bản mô tả này là phương pháp sản xuất hạt chứa protein khảm diệt côn trùng, phương pháp này bao gồm các bước: gieo ít nhất một hạt chứa protein khảm diệt côn trùng; trồng cây từ hạt đó; và thu hoạch hạt từ cây đó, trong đó hạt thu hoạch được chứa protein khảm diệt côn trùng.

Các phân tử polynucleotit tái tổ hợp mã hóa protein khảm diệt côn trùng, chứa trình tự nucleotit được chọn từ nhóm bao gồm các trình tự nêu trong SEQ ID NO: 1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129 hoặc 130; và tùy ý trình tự polynucleotit mã hóa tác nhân ức chế côn trùng khác với protein khảm diệt côn trùng cũng được dự định theo sáng chế.

Phân tử axit nucleic tái tổ hợp khác được dự định theo sáng chế chứa trình tự khởi đầu khác loại được liên kết hoạt động với đoạn polynucleotit mã hóa protein khảm diệt côn trùng, trong đó: protein khảm diệt côn trùng chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO: 21, 10, 28, 7, 4, 13, 16, 19, 23, 25, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, hoặc 111; hoặc protein khảm diệt côn trùng chứa protein có: mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:21, 10; mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:28; mức độ đồng nhất ít nhất 87% so với trình tự SEQ ID NO:7; mức độ đồng nhất ít nhất 90% so với SEQ ID NO:4; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:13; mức độ đồng nhất ít nhất 64% so với trình tự SEQ ID NO:16; mức độ đồng nhất ít nhất 66% so với trình tự SEQ ID NO:19; mức độ đồng nhất ít nhất 86% so với trình tự SEQ ID NO:23; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:25; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:30; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:33; mức độ đồng nhất ít nhất 64% so với trình tự SEQ ID NO:36; mức độ đồng nhất ít nhất 66% so với trình tự SEQ ID NO:39; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO:41; mức độ đồng nhất ít nhất 84% so với trình tự SEQ ID NO:43; mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:45; mức độ đồng nhất ít nhất 94% so với trình tự SEQ ID NO: 47; mức độ đồng nhất ít nhất 91% so với trình tự SEQ ID NO:50; hoặc mức độ đồng nhất ít nhất 93% so với trình tự SEQ ID NO:53; hoặc mức độ đồng nhất ít nhất 87% so với các SEQ ID NO:85, 93, 105; hoặc mức độ đồng nhất ít nhất 85% so với các SEQ ID NO:55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79; hoặc mức độ đồng nhất ít nhất 88% so với các SEQ ID NO:91, 87, 89; hoặc mức độ đồng nhất ít nhất 89% so với các SEQ ID

NO:107, 111; hoặc mức độ đồng nhất ít nhất 90% so với SEQ ID NO:97; hoặc mức độ đồng nhất ít nhất 91% so với SEQ ID NO:109; hoặc mức độ đồng nhất ít nhất 93% so với SEQ ID NO:83; hoặc mức độ đồng nhất ít nhất 94% so với các SEQ ID NO:91,103; hoặc mức độ đồng nhất ít nhất 95% so với các SEQ ID NO:95, 101; hoặc mức độ đồng nhất ít nhất 98% so với SEQ ID NO:99; hoặc đoạn polynucleotit lai với polynucleotit có trình tự nucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO: 1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129 hoặc 130.

Các phương án, dấu hiệu, và các ưu điểm khác theo sáng chế sẽ rõ ràng khi tham khảo phần mô tả chi tiết, ví dụ thực hiện và yêu cầu bảo hộ bên dưới.

Mô tả vắn tắt các trình tự

SEQ ID NO: 1 là trình tự ADN tái tổ hợp mã hóa TIC1100 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 2 là trình tự ADN tổng hợp mã hóa TIC1100 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 3 là trình tự ADN tổng hợp mã hóa TIC1100 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 4 là trình tự axit amin của TIC1100.

SEQ ID NO: 5 là trình tự ADN tái tổ hợp mã hóa TIC860 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 6 là trình tự ADN tổng hợpmã hóa TIC860 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 7 là trình tự axit amin của TIC860.

SEQ ID NO: 8 là trình tự ADN tái tổ hợp mã hóa TIC867 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 9 là trình tự ADN tổng hợp mã hóa TIC867 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 10 là trình tự axit amin của TIC867.

SEQ ID NO: 11 là trình tự ADN tái tổ hợp mã hóa TIC867_20 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 12 là trình tự ADN tổng hợpmã hóa TIC867_20 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 13 là trình tự axit amin của TIC867_20.

SEQ ID NO: 14 là trình tự ADN tái tổ hợp mã hóa TIC867_21 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 15 là trình tự ADN tổng hợpmã hóa TIC867_21 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 16 là trình tự axit amin của TIC867_21.

SEQ ID NO: 17 là trình tự ADN tái tổ hợp mã hóa TIC867_22 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 18 là trình tự ADN tổng hợp mã hóa TIC867_22 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 19 là trình tự axit amin của TIC867_22.

SEQ ID NO: 20 là trình tự ADN tổng hợp mã hóa TIC867_23 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 21 là trình tự axit amin của TIC867_23.

SEQ ID NO: 22 là trình tự ADN tổng hợp mã hóa TIC867_24 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 23 là trình tự axit amin của TIC867_24.

SEQ ID NO: 24 là trình tự ADN tổng hợp mã hóa TIC867_24 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 25 là trình tự axit amin của TIC867_25.

SEQ ID NO: 26 là trình tự ADN tái tổ hợp mã hóa TIC868 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 27 là trình tự ADN tổng hợp mã hóa TIC868 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 28 là trình tự axit amin của TIC868.

SEQ ID NO: 29 là trình tự ADN tổng hợp mã hóa TIC868_9 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 30 là trình tự axit amin của TIC868_9.

SEQ ID NO: 31 là trình tự ADN tái tổ hợp mã hóa TIC868_10 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 32 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa biến thể của TIC868, là TIC868_10.

SEQ ID NO: 33 là trình tự axit amin của TIC868_10.

SEQ ID NO: 34 là trình tự ADN tái tổ hợp mã hóa TIC868_11 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 35 là trình tự ADN tổng hợp mã hóa TIC868_11 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 36 là trình tự axit amin của TIC868_11.

SEQ ID NO: 37 là trình tự ADN tái tổ hợp mã hóa TIC868_12 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 38 là trình tự ADN tổng hợp mã hóa TIC868_12 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 39 là trình tự axit amin của TIC868_12.

SEQ ID NO: 40 là trình tự ADN tổng hợp mã hóa TIC868_13 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 41 là trình tự axit amin của TIC868_13.

SEQ ID NO: 42 là trình tự ADN tổng hợp mã hóa TIC868_14 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 43 là trình tự axit amin của TIC868_14.

SEQ ID NO: 44 là trình tự ADN tổng hợp mã hóa TIC868_15 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 45 là trình tự axit amin của TIC868_15.

SEQ ID NO: 46 là trình tự ADN tổng hợp mã hóa TIC868_29 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 47 là trình tự axit amin của TIC868_29.

SEQ ID NO: 48 là trình tự ADN tái tổ hợp mã hóa TIC869 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 49 là trình tự ADN tổng hợp mã hóa TIC869 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 50 là trình tự axit amin của TIC869.

SEQ ID NO: 51 là trình tự ADN tái tổ hợp mã hóa TIC836 được sử dụng để biểu hiện ở tế bào vi khuẩn.

SEQ ID NO: 52 là trình tự ADN tổng hợp mã hóa TIC836 để biểu hiện ở tế bào thực vật.

SEQ ID NO: 53 là trình tự axit amin của TIC836.

SEQ ID NO: 54 là trình tự ADN mã hóa trình tự axit amin TIC713 khảm.

SEQ ID NO:55 là trình tự axit amin TIC713 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 54.

SEQ ID NO: 56 là trình tự ADN mã hóa trình tự axit amin TIC843 khảm.

SEQ ID NO: 57 là trình tự axit amin TIC843 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 56.

SEQ ID NO: 58 là trình tự ADN mã hóa trình tự axit amin TIC862 khảm.

SEQ ID NO: 59 là trình tự axit amin TIC862 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 58.

SEQ ID NO: 60 là trình tự ADN mã hóa trình tự axit amin TIC1099 khảm.

SEQ ID NO: 61 là trình tự axit amin TIC1099 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 60.

SEQ ID NO: 62 là trình tự ADN mã hóa trình tự axit amin TIC1099-T507E khảm.

SEQ ID NO: 63 là trình tự axit amin TIC1099-T507E đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 62.

SEQ ID NO: 64 là trình tự ADN mã hóa trình tự axit amin TIC1099-R522K khảm.

SEQ ID NO: 65 là trình tự axit amin TIC1099-R522K đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 64.

SEQ ID NO: 66 là trình tự ADN mã hóa trình tự axit amin TIC1099-K490S khảm.

SEQ ID NO: 67 là trình tự axit amin TIC1099-K490S đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 66.

SEQ ID NO: 68 là trình tự ADN mã hóa trình tự axit amin TIC1099-T562R khảm.

SEQ ID NO: 69 là trình tự axit amin TIC1099-T562R đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 68.

SEQ ID NO: 70 là trình tự ADN mã hóa trình tự axit amin TIC1099-S553R khảm.

SEQ ID NO: 71 là trình tự axit amin TIC1099-S553R đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 70.

SEQ ID NO: 72 là trình tự ADN mã hóa trình tự axit amin TIC1099-G498D khảm.

SEQ ID NO: 73 là trình tự axit amin TIC1099-G498D đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 72.

SEQ ID NO: 74 là trình tự ADN mã hóa trình tự axit amin TIC1099-K490A khảm.

SEQ ID NO: 75 là trình tự axit amin TIC1099-K490A đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 74.

SEQ ID NO: 76 là trình tự ADN mã hóa trình tự axit amin TIC1099-E564A khảm.

SEQ ID NO: 77 là trình tự axit amin TIC1099-E564A đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 76.

SEQ ID NO: 78 là trình tự ADN mã hóa trình tự axit amin TIC1103 khảm.

SEQ ID NO: 79 là trình tự axit amin TIC1103đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 78.

SEQ ID NO: 80 là trình tự ADN mã hóa trình tự axit amin TIC1101 khảm.

SEQ ID NO: 81 là trình tự axit amin TIC1101 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 80.

SEQ ID NO: 82 là trình tự ADN mã hóa trình tự axit amin TIC845 khảm.

SEQ ID NO: 83 là trình tự axit amin TIC845 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 82.

SEQ ID NO: 84 là trình tự ADN mã hóa trình tự axit amin TIC846 khảm.

SEQ ID NO: 85 là trình tự axit amin TIC846 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 84.

SEQ ID NO: 86 là trình tự ADN mã hóa trình tự axit amin TIC858 khảm.

SEQ ID NO: 87 là trình tự axit amin TIC858 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 86.

SEQ ID NO: 88 là trình tự ADN mã hóa trình tự axit amin TIC865 khảm.

SEQ ID NO: 89 là trình tự axit amin TIC865 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 88.

SEQ ID NO: 90 là trình tự ADN mã hóa trình tự axit amin TIC866 khảm.

SEQ ID NO: 91 là trình tự axit amin TIC866 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 90.

SEQ ID NO: 92 là trình tự ADN mã hóa trình tự axit amin TIC838 khảm.

SEQ ID NO: 93 là trình tự axit amin TIC838 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 92.

SEQ ID NO: 94 là trình tự ADN mã hóa trình tự axit amin TIC839 khảm.

SEQ ID NO: 95 là trình tự axit amin TIC839 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 94.

SEQ ID NO: 96 là trình tự ADN mã hóa trình tự axit amin TIC841 khảm.

SEQ ID NO: 97 là trình tự axit amin TIC841 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 96.

SEQ ID NO: 98 là trình tự ADN mã hóa trình tự axit amin TIC842 khảm.

SEQ ID NO: 99 là trình tự axit amin TIC842 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 98.

SEQ ID NO: 100 là trình tự ADN mã hóa trình tự axit amin TIC850 khảm.

SEQ ID NO: 101 là trình tự axit amin TIC850 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 100.

SEQ ID NO: 102 là trình tự ADN mã hóa trình tự axit amin TIC859 khảm.

SEQ ID NO: 103 là trình tự axit amin TIC859 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 102.

SEQ ID NO: 104 là trình tự ADN mã hóa trình tự axit amin TIC861 khảm.

SEQ ID NO: 105 là trình tự axit amin TIC861 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 104.

SEQ ID NO: 106 là trình tự ADN mã hóa trình tự axit amin TIC848 khảm.

SEQ ID NO: 107 là trình tự axit amin TIC848 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 106.

SEQ ID NO: 108 là trình tự ADN mã hóa trình tự axit amin TIC849 khảm.

SEQ ID NO: 109 là trình tự axit amin TIC849 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 108.

SEQ ID NO: 110 là trình tự ADN mã hóa trình tự axit amin TIC847 khảm.

SEQ ID NO: 111 là trình tự axit amin TIC847 đã được dịch mã từ khung đọc mở nêu trong trình tự SEQ ID NO: 110.

SEQ ID NO: 112 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC713.

SEQ ID NO: 113 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC713.

SEQ ID NO: 114 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC843.

SEQ ID NO: 115 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC862.

SEQ ID NO: 116 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC1099.

SEQ ID NO: 117 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC1103.

SEQ ID NO: 118 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC845.

SEQ ID NO: 119 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC846.

SEQ ID NO: 120 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC858.

SEQ ID NO: 121 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC866.

SEQ ID NO: 122 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC838.

SEQ ID NO: 123 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC841.

SEQ ID NO: 124 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC842.

SEQ ID NO: 125 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC850.

SEQ ID NO: 126 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC859.

SEQ ID NO: 127 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC861.

SEQ ID NO: 128 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC848.

SEQ ID NO: 129 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC849.

SEQ ID NO: 130 là trình tự ADN tổng hợp để biểu hiện ở tế bào thực vật mã hóa TIC847.

Mô tả chi tiết sáng chế

Vấn đề trong lĩnh vực phòng diệt sinh vật gây hại nông nghiệp có thể được xác định là nhu cầu về các protein diệt côn trùng mới mà là hiệu nghiệm đối với các sinh vật gây hại đích, thể hiện phổ độc tính rộng đối với các loài sinh vật gây hại đích, có khả năng được biểu hiện ở thực vật mà không gây ra các vấn đề nông học không mong muốn, và là phương thức tác động thay thế so với các độc tố thông dụng mà được sử dụng thương mại ở thực vật. Các protein khảm diệt côn trùng mới được bộc lộ trong bản mô tả này, và đáp ứng từng nhu cầu trong số các nhu cầu này, đặc biệt là đối với phổ rộng côn trùng gây hại là bọ cánh vảy.

Nhằm để tránh sự phát triển của, hoặc phá vỡ tính kháng của côn trùng đối với các protein diệt côn trùng hiện đang được sử dụng, các protein diệt côn trùng mới với cách thức hoạt động (MOA) khác, cũng như phổ rộng và hiệu quả, là cần thiết để phòng trừ bọ cánh vảy. Một cách để đáp ứng nhu cầu này là phát hiện các protein diệt côn trùng mới từ các nguồn sinh học khác nhau, tốt hơn là từ vi khuẩn, nấm mốc hoặc thực vật. Cách khác là trao đổi các phân đoạn giữa các protein khác nhau của *Bt* mà thường thể hiện sự tương tự về cấu trúc để tạo ra các protein khảm mới của *Bt* có tính chất ức chế côn trùng. Trong lĩnh vực kỹ thuật này, đã biết rằng khả năng tạo ra protein khảm với các tính chất tốt hơn từ việc phân loại lại các cấu trúc miền của nhiều protein tinh thể tự nhiên có hoạt tính diệt côn trùng đã biết trong lĩnh vực kỹ thuật này là xa vời. Ví dụ, xem tài liệu: Jacqueline S. Knight, et al. "A Strategy for Shuffling Numerous *Bacillus thuringiensis* Crystal Protein Domains." *J. Economic Entomology*, 97 (6) (2004): 1805-1813.

Sáng chế này bộc lộ các trình tự phân tử axit nucleic tái tổ hợp mà mã hóa các protein khảm diệt côn trùng mới. Các protein diệt côn trùng này đáp ứng nhu cầu hiện tại trong lĩnh vực kỹ thuật này xử lý protein diệt côn trùng có độc tính bổ sung với các tính chất diệt côn trùng tốt hơn như hiệu quả cao hơn đối với phổ rộng hơn bao gồm các loài côn trùng gây hại đích và cách thức hoạt động khác. Các thành viên của nhóm protein này, kể cả các protein làm ví dụ đã được bộc lộ trong bản mô tả này, thể hiện hoạt tính diệt côn trùng đối với côn trùng loài sinh vật gây hại bộ cánh vảy.

Các thuật ngữ "phân đoạn" hoặc "đoạn" theo sáng chế chỉ các trình tự axit amin hoặc axit nucleic liên tục mà là ngắn hơn trình tự axit amin hoặc axit nucleic hoàn chỉnh mô tả protein khảm diệt côn trùng đã được bộc lộ. Phân đoạn hoặc đoạn thể hiện hoạt tính ức chế côn trùng cũng được bộc lộ theo sáng chế nếu việc duỗi thẳng của phân đoạn hoặc đoạn này, với phần tương ứng của protein khảm diệt côn trùng, dẫn đến mức độ tương tự về trình tự axit amin theo tỷ lệ phần trăm nằm trong khoảng từ 65 đến 100 phần trăm giữa phân đoạn hoặc đoạn và phần tương ứng của protein khảm diệt côn trùng.

Các thuật ngữ "có hoạt tính" hoặc "hoạt tính", "hoạt tính diệt sinh vật gây hại" hoặc "diệt sinh vật gây hại", hoặc "hoạt tính diệt côn trùng", "ức chế côn trùng" hoặc "diệt côn trùng" theo sáng chế chỉ hiệu quả của tác nhân có độc tính, như protein diệt côn trùng, trong việc ức chế (ức chế sự sinh trưởng, ăn, sinh sản, hoặc sức sống), ngăn chặn (ngăn chặn sự sinh trưởng, ăn, sinh sản, hoặc sức sống), phòng trừ (phòng trừ sự phá hoại của sinh vật gây hại, phòng diệt sinh vật gây hại ăn cây trồng cụ thể chứa protein diệt côn trùng với lượng hữu hiệu) hoặc diệt (gây bệnh, chết, hoặc giảm mức độ sinh sản) sinh vật gây hại. Các thuật ngữ này được dự định bao gồm kết quả của việc cung cấp protein diệt côn trùng với lượng có tác dụng hữu hiệu diệt sinh vật gây hại đối với sinh vật gây hại khi sự tiếp cận của sinh vật gây hại với protein diệt côn trùng dẫn đến bệnh tật, chết, giảm sinh sản, hoặc còi cọc. Các thuật ngữ này còn bao gồm cả sự sợ hãi của sinh vật gây hại đối với thực vật, mô thực vật, bộ phận của cây, hạt, các tế bào thực vật, hoặc vị trí địa lý cụ thể nơi thực vật đó có thể sinh trưởng, là kết quả của việc cung cấp protein diệt côn trùng với lượng có tác dụng hữu hiệu diệt sinh vật gây hại ở hoặc trên thực vật. Nói chung, hoạt tính diệt côn trùng là khả năng của protein diệt côn trùng hữu hiệu trong việc ức chế sự sinh trưởng, phát triển, sức sống, tập tính ăn, tập tính giao phối, sức sinh sản, hoặc mức độ giảm bất kỳ đo được về

tác dụng phụ gây ra do côn trùng ăn protein này, đoạn protein, phân đoạn protein hoặc polynucleotit của sinh vật gây hại đích cụ thể, bao gồm, nhưng không chỉ giới hạn ở, côn trùng thuộc bộ cánh vảy. Protein diệt côn trùng có thể được tạo ra bằng thực vật hoặc có thể được dùng cho thực vật hoặc môi trường trong khu vực của thực vật đó. Các thuật ngữ "hoạt tính sinh học", "hiệu quả", "hữu hiệu" hoặc các biến thể của chúng cũng là các thuật ngữ được dùng thay thế lẫn nhau trong bản mô tả này để mô tả tác động của các protein khảm diệt côn trùng theo sáng chế đối với các côn trùng gây hại đích.

Lượng có tác dụng hữu hiệu diệt sinh vật gây hại của tác nhân có độc tính, khi được cung cấp trong chế độ dinh dưỡng của sinh vật gây hại đích, thể hiện hoạt tính diệt côn trùng khi tác nhân có độc tính tiếp xúc sinh vật gây hại. Tác nhân có độc tính có thể là protein diệt côn trùng hoặc một hoặc nhiều tác nhân hóa học đã biết trong lĩnh vực kỹ thuật này. Các tác nhân hóa học có hoạt tính diệt côn trùng và các tác nhân protein có hoạt tính diệt côn trùng có thể được sử dụng riêng rẽ hoặc kết hợp với nhau. Các tác nhân hóa học bao gồm, nhưng không chỉ giới hạn ở, các phân tử ARN sợi đôi hướng đích đến gen cụ thể để ngăn chặn ở sinh vật gây hại đích, clorua hữu cơ, phosphat hữu cơ, carbamat, pyrethroit, neonicotinoit, và ryanoit. Các tác nhân protein có hoạt tính diệt côn trùng bao gồm các protein khảm diệt côn trùng nêu trong bản mô tả này, cũng như các tác nhân có bản chất protein có độc tính khác kể cả các tác nhân mà hướng đích đến loài sinh vật gây hại bộ cánh vảy, cũng như các độc tố protein mà được dùng để phòng diệt sinh vật gây hại thực vật khác như các protein Cry sẵn có trong lĩnh vực kỹ thuật này để dùng trong việc phòng trừ các loài bọ cánh cứng, bọ cánh tơ, bọ cánh nửa và bọ cánh giống.

Dự định rằng việc đề cập đến sinh vật gây hại, đặc biệt là sinh vật gây hại cho cây trồng, nghĩa là các côn trùng gây hại cây trồng, đặc biệt là các côn trùng gây hại là bọ cánh vảy mà được phòng trừ bởi các protein khảm diệt côn trùng đã được bộc lộ. Tuy nhiên, việc đề cập đến sinh vật gây hại còn có thể bao gồm các côn trùng gây hại thực vật thuộc bọ cánh cứng, bọ cánh nửa và bọ cánh giống, cũng như giun tròn và nấm mốc khi các tác nhân có độc tính hướng đích đến các sinh vật gây hại này ở cùng khu vực hoặc có mặt cùng với protein khảm diệt côn trùng, hoặc protein mà đồng nhất 65 đến 100 phần trăm với protein khảm diệt côn trùng.

Các protein khảm diệt côn trùng đã được bộc lộ theo sáng chế thể hiện hoạt tính diệt côn trùng đối với các côn trùng gây hại thuộc loài côn trùng bọ cánh vảy, kể cả con trưởng thành, con nhộng, ấu trùng, và con mới sinh, cũng như loài côn trùng bọ cánh nửa, kể cả con trưởng thành và con nhộng. Côn trùng thuộc bộ cánh vảy bao gồm, nhưng không chỉ giới hạn ở, sâu ăn lá, sâu ngài đêm, sâu đo, và bướm đêm thuộc họ Noctuidae, ví dụ, sâu xanh mùa thu (*Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (*Spodoptera exigua*), sâu ăn lá bertha (*Mamestra configurata*), sâu xám (*Agrotis ipsilon*), sâu đo ký sinh trên cải bắp (*Trichoplusia ni*), sâu đo đậu tương (*Pseudoplusia includens*), sâu bướm hại đậu mèo (*Anticarsia gemmatalis*), sâu xanh (*Hypena scabra*), sâu chồi thuốc lá (*Heliothis virescens*), sâu ngài đêm dạng hạt (*Agrotis subterranea*), sâu ăn lá (*Pseudaletia unipuncta*), sâu ngài đêm tây (*Agrotis orthogonia*); sâu đục thân, sâu túi, sâu kéo màng, sâu hình nón, sâu hại cải bắp và sâu gặm thịt lá thuộc họ Pyralidae, ví dụ, sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), bướm hại cam navel (*Amyelois transitella*), sâu kéo màng hại rễ ngô (*Crambus caliginosellus*), sâu kéo màng ở cỏ (*Herpetogramma licarsisalis*), sâu bướm hướng dương (*Homoeosoma electellum*), sâu đục thân ngô nhỏ (*Elasmopalpus lignosellus*); sâu cuốn lá, sâu đục chồi, sâu hại hạt, và sâu hại quả thuộc họ Tortricidae, ví dụ, bướm cá tuyết (*Cydia pomonella*), bướm hại cây việt quất (*Endopiza viteana*), bướm hại quả phương đông (*Grapholita molesta*), bướm hại chồi hướng dương (*Suleima helianthana*); và nhiều bọ cánh vảy khác quan trọng về mặt kinh tế, ví dụ, sâu tơ (*Plutella xylostella*), sâu hồng đục quả (*Pectinophora gossypiella*) và sâu bướm gypsy (*Lymantria dispar*). Các côn trùng gây hại khác thuộc bộ cánh vảy bao gồm, ví dụ, *Alabama argillacea* (sâu hại lá cây bông), *Archips argyrospila* (sâu cuốn lá cây ăn quả), *Archips rosana* (sâu cuốn lá châu Âu) và loài *Archips* khác, *Chilo suppressalis* (sâu đục thân lúa châu Á, hoặc sâu đục thân lúa), *Cnaphalocrocis medinalis* (sâu cuốn lá lúa), *Crambus caliginosellus* (sâu kéo màng hại rễ ngô), *Crambus teterrellus* (sâu kéo màng cỏ blue), *Diatraea grandiosella* (sâu đục thân ngô tây nam), *Diatraea saccharalis* (sâu đục thân mía), *Earias insulana* (sâu đục quả có gai), *Earias vittella* (sâu đục quả đốm), *Helicoverpa armigera* (sâu đục quả Mỹ), *Helicoverpa zea* (sâu hại ngô hoặc sâu hại bông), *Heliothis virescens* (sâu chồi thuốc lá), *Herpetogramma licarsisalis* (sâu kéo màng ở cỏ), *Lobesiabotrana* (bướm hại nho châu Âu), *Phyllocnistis citrella* (sâu ăn lá quả có múi), *Pieris brassicae* (bướm trắng to), *Pieris*

rapae (sâu hại cải bắp nhập khẩu, hoặc bướm trắng nhỏ), *Plutella xylostella* (sâu tơ), *Spodoptera exigua* (sâu hại củ cải đường), *Spodoptera litura* (sâu hại thuốc lá, sâu bướm cụm), và *Tuta absoluta* (sâu ăn lá cà chua).

Thuật ngữ "phân tử ADN phân lập", hoặc thuật ngữ hoặc cụm từ tương đương, được dùng trong bản mô tả này để chỉ phân tử ADN là loại mà có mặt riêng rẽ hoặc kết hợp với các thành phần khác, nhưng không trong môi trường tự nhiên của nó. Ví dụ, các thành phần axit nucleic như trình tự mã hóa, trình tự intron, trình tự dẫn đầu chưa dịch mã, trình tự của trình tự khởi đầu, trình tự kết thúc phiên mã, và thành phần tương tự, mà vốn tìm thấy ở ADN của hệ gen của sinh vật không được xem là "đã được phân lập" chừng nào thành phần đó ở trong hệ gen của sinh vật này và tại vị trí trong hệ gen nơi nó được tìm thấy trong tự nhiên. Tuy nhiên, mỗi thành phần trong số các thành phần này, và phần phụ của các thành phần này, sẽ là "đã được phân lập" trong phạm vi của bản mô tả này chừng nào thành phần đó không trong hệ gen của sinh vật này và tại vị trí trong hệ gen nơi nó được tìm thấy trong tự nhiên. Tương tự, trình tự nucleotit mã hóa protein diệt côn trùng hoặc biến thể bất kỳ có hoạt tính diệt côn trùng trong tự nhiên của protein đó sẽ là trình tự nucleotit phân lập chừng nào trình tự nucleotit không nằm trong ADN của vi khuẩn nơi mà trình tự mã hóa protein đó được tìm thấy trong tự nhiên. Trình tự nucleotit tổng hợp mã hóa trình tự axit amin của protein có hoạt tính diệt côn trùng có trong tự nhiên được xem là được phân lập nhằm mục đích của sáng chế. Nhằm mục đích của sáng chế, trình tự nucleotit chuyển gen bất kỳ, tức là trình tự nucleotit của ADN được cài xen vào hệ gen của các tế bào của thực vật hoặc vi khuẩn, hoặc có mặt ở vectơ ngoài nhiễm sắc thể, xem là trình tự nucleotit phân lập dù nó có mặt trong plasmit hoặc cấu trúc tương tự được dùng để biến nạp tế bào, trong hệ gen của thực vật hoặc vi khuẩn, hoặc có mặt với lượng phát hiện được ở mô, cây con, các mẫu sinh học hoặc các sản phẩm hàng hóa thu được từ thực vật hoặc vi khuẩn đó.

Như được mô tả rõ hơn trong các ví dụ thực hiện sáng chế, thông qua cố gắng tạo khảm khoảng tám trăm và bốn mươi bốn (844) trình tự nucleotit mà mã hóa các protein khảm diệt côn trùng được xây dựng từ miền của tiền độc tố và miền của độc tố của các độc tố diệt côn trùng đã biết (trong bản mô tả này còn được gọi là "các protein gốc"), và được biểu hiện và được thử nghiệm trong thử nghiệm sinh học về hoạt tính đối với bọ cánh vảy. Số lượng ít protein khảm diệt côn trùng xây dựng được thể hiện

hoạt tính tốt hơn đối với bọ cánh vảy hoặc phổ nâng cao đối với bọ cánh vảy so với các protein gốc mà từ đó thu được các thành phần độc tố.

Các protein khảm diệt côn trùng mới có hoạt tính tốt hơn đối với bọ cánh vảy hoặc phổ nâng cao đối với bọ cánh vảy này được xây dựng từ miền của tiền độc tố và miền của độc tố của protein gốc diệt côn trùng dưới đây: Cry1Ah (miền I), Cry1Bb1 (miền I và miền II), Cry 1Be2 (miền I và miền II), Cry1Ja1 (miền I và miền II), Cry1Fa1 (miền I và miền II), Cry1Ac (miền II và tiền độc tố), Cry1Ca (miền III và tiền độc tố), Cry1Ka (miền III và tiền độc tố), Cry1Jx (miền III), Cry1Ab (miền III), Cry1Ab3 (tiền độc tố), Cry1Da1 (tiền độc tố), Cry4 (tiền độc tố), Cry9 (tiền độc tố), Cry1Be (tiền độc tố) và Cry1Ka (tiền độc tố).

Đặc biệt, các protein khảm diệt côn trùng mới theo sáng chế có hoạt tính tốt hơn đối với bọ cánh vảy hoặc phổ nâng cao đối với bọ cánh vảy bao gồm các tổ hợp tiền độc tố và miền dưới đây: TIC1100/SEQ ID NO:4 (miền I- Cry1Ah, miền II- Cry1Ac, miền III- Cry1Ca, tiền độc tố- Cry1Ac), TIC860/SEQ ID NO:7 (miền I- Cry1Bb1, miền II- Cry1BB1, miền III- Cry1Ca, tiền độc tố- Cry1Ac), TIC867/SEQ ID NO:10 (miền I- Cry1Be2, miền II- Cry1Be2, miền III-Cry1Ka, tiền độc tố- Cry1Ab3), TIC868/SEQ ID NO:28 (miền I- Cry1Be2, miền II-Cry1Be2, và miền III- Cry1Ca, tiền độc tố- Cry1Ab3), TIC869/SEQ ID NO:50 (miền I-Cry1Ja1, miền II- Cry1Ja1, miền III- Cry1Jx, tiền độc tố-Cry1Ab3) và TIC836/SEQ ID NO:53 (miền I-Cry1Fa1, miền II-Cry1Fa1, miền III- Cry1Ab, tiền độc tố-Cry1Ac).

Các biến thể trong đó có đưa vào các thay thế axit amin hoặc các miền xen kẽ của tiền độc tố cũng được xây dựng cho các protein khảm diệt côn trùng TIC867 và TIC868. Đặc biệt, các biến thể này của TIC867 và TIC868 bao gồm các thay thế axit amin hoặc các miền xen kẽ của tiền độc tố dưới đây: TIC867_20/SEQ ID NO:13 (miền xen kẽ của tiền độc tố Cry1Da1), TIC867_21/SEQ ID NO:16 (miền xen kẽ của tiền độc tố Cry4), TIC867_22/SEQ ID NO:19 (miền xen kẽ của tiền độc tố Cry9), TIC867_23/SEQ ID NO:21 (miền xen kẽ của tiền độc tố Cry1Be), TIC867_24/SEQ ID NO:23 (miền xen kẽ của tiền độc tố Cry1Ka), TIC867_25/SEQ ID NO: 25 (miền xen kẽ của tiền độc tố Cry1Ka), TIC868_9/SEQ ID NO:30 (cải biến axit amin N240S_Y343Q_N349T), TIC868_10/SEQ ID NO:33 (miền xen kẽ của tiền độc tố Cry1Da1), TIC868_11/SEQ ID NO:36 (miền xen kẽ của tiền độc tố Cry4), TIC868_12/SEQ ID NO:39 (miền xen kẽ của tiền độc tố Cry 9), TIC868_13/SEQ ID

NO:41 (miền xen kẽ của tiền độc tố Cry1Be), TIC868_14/SEQ ID NO:43 (miền xen kẽ của tiền độc tố Cry1Ka), TIC868_15/SEQ ID NO:45 (miền xen kẽ của tiền độc tố Cry1Ca), và TIC868_29/SEQ ID NO:47 (cải biến axit amin Q136Y_Y343Q_N349T).

Như được thể hiện trong các ví dụ thực hiện sáng chế, mỗi biến thể trong số các biến thể TIC867 và TIC868 này thay đổi hoạt tính đối với bọ cánh vảy và/hoặc làm giảm phổ hoạt tính đối với bọ cánh vảy của protein khảm diệt côn trùng gốc, do đó cho thấy rằng miền xen kẽ của tiền độc tố và thay thế axit amin có tác động trực tiếp đến hoạt tính diệt côn trùng và phổ của các protein khảm diệt côn trùng TIC867 và TIC868.

Nhiều protein khảm diệt côn trùng thể hiện hoạt tính diệt côn trùng đối với nhiều loài côn trùng sinh vật gây hại bộ cánh vảy. Đặc biệt, các protein khảm diệt côn trùng mới được bộc lộ trong bản mô tả này thể hiện hoạt tính đối với một hoặc nhiều côn trùng gây hại là bọ cánh vảy dưới đây, sâu bướm nhung hại đỗ (VBC, *Anticarsia gemmatalis*), sâu đục thân mía (SCB, *Diatraea saccharalis*), sâu đục thân ngô nhỏ (LSCB, *Elasmopalpus lignosellus*), sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu vỏ đậu tương (SPW, *Helicoverpa zea*), sâu hại bông (CBW, *Helicoverpa zea*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu đen ăn lá (BLAW, *Spodoptera cosmioides*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (BAW, *Spodoptera exigua*), sâu đục quả cựu thế giới (OBW, *Helicoverpa armigera*), sâu ăn lá phương đông (OLW, *Spodoptera litura*), sâu hồng đục quả (PBW, *Pectinophora gossypiella*), sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*), sâu đục quả đốm (SBW, *Earias vitella*), sâu đục quả Mỹ (SABW, *Helicoverpa gelotopeon*), và sâu đo hướng dương (SFL, *Rachiplusia nu*). Do đó, các protein làm ví dụ đã được bộc lộ trong bản mô tả này có liên quan theo chức năng chung và thể hiện hoạt tính diệt côn trùng đối với các côn trùng gây hại thuộc các loài côn trùng bọ cánh vảy kể cả con trưởng thành, ấu trùng và con nhộng.

Các protein mà giống các protein khảm diệt côn trùng có thể được xác định bằng cách so sánh với nhau bằng cách áp dụng các thuật toán khác nhau trên máy tính đã biết trong lĩnh vực kỹ thuật này. Ví dụ, nhận diện trình tự axit amin của các protein liên quan đến các protein khảm diệt côn trùng có thể được phân tích bằng cách áp dụng cách duỗi thẳng Clustal W bằng cách sử dụng các thông số mặc định này: ma

trận trọng lượng: blosum, điểm trừ cho việc mở khoảng cách: 10,0, điểm trừ cho việc kéo dài khoảng cách: 0,05, các khoảng cách ưa nước: bật, các gốc ưa nước: GPSNDQERK, điểm trừ khoảng cách đặc hiệu đối với gốc: bật (Thompson, et al (1994) Nucleic Acids Research, 22:4673-4680). Phần trăm nhận diện axit amin còn được tính theo kết quả của 100% nhân với (số nhận diện axit amin/chiều dài của protein đối tượng). Các thuật toán duỗi thẳng khác cũng sẵn có trong lĩnh vực kỹ thuật này, tạo ra kết quả tương tự như thu được bằng cách áp dụng cách duỗi thẳng Clustal W và được dự định theo sáng chế.

Dự định rằng protein nghi ngờ thể hiện hoạt tính ức chế côn trùng được bộc lộ trong bản mô tả này nếu việc duỗi thẳng protein nghi ngờ đó với các protein khảm diệt côn trùng đối tượng nêu trong trình tự SEQ ID NO: 4, 7, 10, 13, 16, 19, 21, 23, 25, 28, 30, 33, 36, 39, 41, 43, 45, 47, 50, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, và 111 và dẫn đến mức độ tương tự về trình tự axit amin ít nhất khoảng 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, hoặc khoảng 100% (hoặc phân đoạn tỷ lệ phần trăm bất kỳ trong khoảng này) giữa protein nghi ngờ và protein đối tượng.

Như được mô tả rõ hơn trong các ví dụ thực hiện sáng chế, các trình tự tổng hợp hoặc nhân tạo mã hóa các protein khảm diệt côn trùng được thiết kế để dùng ở thực vật. Các trình tự nucleotit tổng hợp làm ví dụ mà được thiết kế để dùng ở thực vật được thể hiện ở trình tự SEQ ID NO: 2 và 3 (TIC1100), SEQ ID NO:6 (TIC860), SEQ ID NO:9 (TIC867), SEQ ID NO:12 (TIC867_20), SEQ ID NO:15 (TIC867_21), SEQ ID NO:18 (TIC867_22), SEQ ID NO:20 (TIC867_23), SEQ ID NO:22 (TIC867_24), SEQ ID NO: 24 (TIC867_25), SEQ ID NO:27 (TIC868), SEQ ID NO:29 (TIC868_9), SEQ ID NO:32 (TIC868_10), SEQ ID NO:35 (TIC868_11), SEQ ID NO:38 (TIC868_12), SEQ ID NO:40 (TIC868_13), SEQ ID NO:42 (TIC868_14), SEQ ID NO:44 (TIC868_15), SEQ ID NO:46 (TIC868_29), SEQ ID NO:49 (TIC869) và SEQ ID NO:52 (TIC836), SEQ ID NO:112 và 113 (TIC713), SEQ ID NO:114 (TIC843), SEQ ID NO:115 (TIC862), SEQ ID NO:116 (TIC1099), SEQ ID NO:117 (TIC1103), SEQ ID NO:118 (TIC845), SEQ ID NO:119 (TIC846), SEQ ID NO:120 (TIC858), SEQ ID NO:121 (TIC866), SEQ ID NO:122 (TIC838), SEQ ID NO:123 (TIC841),

SEQ ID NO:124 (TIC842), SEQ ID NO:125 (TIC850), SEQ ID NO:126 (TIC859), SEQ ID NO:127 (TIC861), SEQ ID NO:128 (TIC848), SEQ ID NO:129 (TIC849), và SEQ ID NO:130 (TIC847).

Để biểu hiện ở các tế bào thực vật, các protein khảm diệt côn trùng có thể được biểu hiện để cư trú trong dung dịch bào tương hoặc được hướng đích đến các hạt cơ quan khác nhau của tế bào thực vật. Ví dụ, hướng đích protein đến thể lạp lục có thể làm tăng nồng độ của protein đã được biểu hiện ở thực vật chuyển gen trong khi ngăn chặn hiện tượng mất kiểu hình. Việc hướng đích còn có thể làm tăng hiệu lực kháng của sinh vật gây hại trong trường hợp chuyển gen. Peptit đích hoặc peptit chuyển tiếp là mạch peptit ngắn (dài 3 đến 70 axit amin) mà chi phối sự vận chuyển của protein đến vùng cụ thể trong tế bào, kể cả nhân, ty thể, lưới nội chất (ER), lạp lục, không bào, thể peroxi và màng huyết tương. Một số peptit đích được phân cắt từ protein nhờ peptiđaza tín hiệu sau khi các protein được vận chuyển. Để hướng đích đến lạp lục, các protein chứa các peptit chuyển tiếp mà khoảng 40 đến 50 axit amin. Xem US 5,188,642 và US 5,728,925 mô tả về việc sử dụng peptit chuyển tiếp ở lạp lục. Nhiều protein nằm ở lạp lục được biểu hiện từ gen ở nhân như tiền chất và được hướng đích đến lạp lục nhờ peptit chuyển tiếp ở lạp lục (CTP). Các ví dụ về các protein ở lạp lục đã được phân lập này bao gồm, nhưng không chỉ giới hạn ở, các protein liên quan đến cấu trúc siêu phân tử nhỏ (SSU) của ribuloza-1,5-bisphosphat carboxylaza, feređoxin, feređoxin oxiđoređuctaza, protein I và protein II phức thu ánh sáng, thioređoxin F, enolpyruvyl shikimat phosphat synthaza (EPSPS), và các peptit chuyển tiếp đã được bộc lộ trong US 7,193,133. Đã được thể hiện *in vivo* và *in vitro* rằng các protein không ở lạp lục có thể được hướng đích đến lạp lục bằng cách sử dụng thể dung hợp protein với CTP khác nguồn gốc và rằng CTP là đủ để hướng đích protein đến lạp lục. Đã thấy rằng việc kết hợp peptit chuyển tiếp thích hợp ở lạp lục như EPSPS CTP (CTP2) của *Arabidopsis thaliana* (xem tài liệu: Klee et al., Mol. Gen. Genet. 210:437-442, 1987) hoặc EPSPS CTP (CTP4) của Petunia hybrida (xem tài liệu: della-Cioppa et al., Proc. Natl. Acad. Sci. USA 83:6873-6877, 1986) hướng đích các trình tự protein EPSPS khác nguồn gốc đến lạp lục ở các cây chuyển gen (xem US 5,627,061; US 5,633,435; và US 5,312,910; và EP 0218571; EP 189707; EP 508909; và EP 924299). Để hướng đích các protein khảm diệt côn trùng đến lạp lục, trình tự mã hóa peptit chuyển tiếp ở lạp lục được đặt ở đầu 5' trong liên kết hoạt động và trong khung với trình tự mã hóa

tổng hợp mã hóa protein khảm diệt côn trùng mà đã được thiết kế để biểu hiện tối ưu ở các tế bào thực vật.

Các catxet biểu hiện và các vectơ chứa các trình tự nucleotit tổng hợp hoặc nhân tạo này được xây dựng và đưa vào các tế bào ngô, bông, và đậu tương theo các phương pháp biến nạp và các kỹ thuật mà là đã biết trong lĩnh vực kỹ thuật này. Các tế bào đã biến nạp được tái sinh thành cây biến nạp mà quan sát thấy biểu hiện protein khảm diệt côn trùng. Để thử nghiệm hoạt tính diệt côn trùng, các thử nghiệm sinh học được thực hiện với sự có mặt của ấu trùng sinh vật gây hại bộ cánh vảy bằng cách sử dụng phiến lá thực vật thu được từ cây đã biến nạp. Các chế phẩm phân tử axit nucleic tái tổ hợp mà mã hóa các protein khảm diệt côn trùng được dự định theo sáng chế. Ví dụ, các protein khảm diệt côn trùng có thể được biểu hiện với cấu trúc ADN tái tổ hợp, trong đó phân tử polynucleotit với ORF mã hóa protein khảm diệt côn trùng được liên kết hoạt động với yếu tố biểu hiện di truyền như gen khởi đầu và thành phần điều hòa bất kỳ khác cần thiết để biểu hiện ở hệ mà cấu trúc này được dự định. Các ví dụ không giới hạn phạm vi của sáng chế bao gồm trình tự khởi đầu có chức năng ở thực vật được liên kết hoạt động với protein khảm diệt côn trùng tổng hợp mã hóa các trình tự để biểu hiện protein khảm diệt côn trùng ở thực vật hoặc trình tự khởi đầu chức năng của *Bt* được liên kết hoạt động với trình tự mã hóa protein khảm diệt côn trùng để biểu hiện protein ở vi khuẩn *Bt* hoặc loài *Bacillus* khác. Các thành phần khác có thể được liên kết hoạt động với các các trình tự mã hóa protein khảm diệt côn trùng bao gồm, nhưng không chỉ giới hạn ở, gen tăng cường, intron, trình tự dẫn đầu chưa được dịch mã, đuôi đánh dấu cố định protein đã mã hóa (đuôi đánh dấu HIS), các peptit dịch chuyển (tức là peptit chuyển tiếp thể hạt, peptit tín hiệu), các trình tự polypeptit để cải biến sau dịch mã enzym, vị trí liên kết ở ribosom, và vị trí đích ở ARN can thiệp.

Các phân tử polynucleotit tái tổ hợp làm ví dụ theo sáng chế bao gồm, nhưng không chỉ giới hạn ở, trình tự khởi đầu khác loại được liên kết hoạt động với polynucleotit như SEQ ID NO: 1, 2, 3, 5, 6, 8, 9, 11, 12, 14, 15, 17, 18, 20, 22, 24, 26, 27, 29, 31, 32, 34, 35, 37, 38, 40, 42, 44, 46, 48, 49, 51, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, và 130; mà mã hóa polypeptit hoặc protein có trình tự axit amin đã nêu trong SEQ ID NO: 4 (TIC1100), 7 (TIC860), 10 (TIC867), 13 (TIC867_20), 16

(TIC867_21), 19 (TIC867_22), 21 (TIC867_23), 23 (TIC867_24), 25 (TIC867_25), 28 (TIC868), 30 (TIC868_9), 33 (TIC868_10), 36 (TIC868_11), 39 (TIC867_12), 41 (TIC867_13), 43 (TIC867_14), 45 (TIC867_15), 47 (TIC867_29), 50 (TIC869), 53 (TIC836), 55 (TIC713), 57 (TIC843), 59 (TIC862), 61 (TIC1099), 63 (TIC1099-T507E), 65 (TIC1099-R522K), 67 (TIC1099-K490S), 69 (TIC1099-T562R), 71 (TIC1099-S533R), 73 (TIC1099-G498D), 75 (TIC1099-K490A), 77 (TIC1099-E564A), 79 (TIC1103), 81 (TIC1101), 83 (TIC845), 85 (TIC846), 87 (TIC858), 89 (TIC865), 91 (TIC866), 93 (TIC838), 95 (TIC839), 97 (TIC841), 99 (TIC842), 101 (TIC850), 103 (TIC859), 105 (TIC861), 107 (TIC848), 109 (TIC849), và 111 (TIC847). Trình tự khởi đầu khác loại cũng có thể được liên kết hoạt động với các trình tự mã hóa ADN tổng hợp mã hóa protein khảm diệt côn trùng hướng đích đến thể hạt và protein khảm diệt côn trùng không hướng đích. Dự định rằng các codon của phân tử axit nucleic tái tổ hợp mã hóa protein khảm diệt côn trùng đã được bộc lộ trong bản mô tả này có thể được thế bằng codon cùng nghĩa (đã biết trong lĩnh vực kỹ thuật này là thay thế câm).

Phân tử hoặc cấu trúc ADN tái tổ hợp chứa trình tự mã hóa protein khảm diệt côn trùng còn có thể chứa vùng của ADN mà mã hóa một hoặc nhiều tác nhân có độc tính mà có thể được tạo cấu hình để biểu hiện đồng thời hoặc cùng biểu hiện với trình tự ADN mã hóa protein khảm diệt côn trùng, protein khác với protein khảm diệt côn trùng, phân tử ARN sợi đôi ức chế côn trùng, hoặc protein phụ. Các protein phụ bao gồm, nhưng không chỉ giới hạn ở, các đồng tác nhân, các enzym, các đối tác liên kết, hoặc các tác nhân khác mà thực hiện chức năng trợ giúp có hiệu quả của tác nhân ức chế côn trùng, ví dụ, bằng cách trợ giúp sự biểu hiện của nó, ảnh hưởng đến tính ổn định ở thực vật, tối ưu hóa năng lượng tự do cho việc tạo oligome, làm gia tăng độc tính của nó, và làm tăng phổ hoạt tính của nó. Protein phụ có thể tạo điều kiện thuận lợi cho sự hấp thụ một hoặc nhiều tác nhân phòng diệt côn trùng, ví dụ, hoặc tạo tiềm năng cho tác dụng độc của tác nhân có độc tính.

Phân tử hoặc cấu trúc ADN tái tổ hợp có thể được lắp ráp sao cho tất cả các protein hoặc các phân tử ARN sợi đôi được biểu hiện từ một trình tự khởi đầu hoặc từng protein hoặc phân tử ARN sợi đôi chịu sự kiểm soát của trình tự khởi đầu riêng hoặc một vài cách kết hợp của chúng. Các protein theo sáng chế có thể được biểu hiện từ hệ biểu hiện nhiều gen, trong đó protein khảm diệt côn trùng được biểu hiện từ đoạn

nucleotit chung mà nó còn chứa các khung đọc mở và các gen khởi đầu khác, tùy theo loại hệ biểu hiện đã chọn. Ví dụ, hệ biểu hiện nhiều gen của vi khuẩn có thể sử dụng một trình tự khởi đầu để dẫn sự biểu hiện của các khung đọc mở đa liên kết/nối tiếp từ bên trong một operon (tức là biểu hiện đa cistron). Theo ví dụ khác, hệ biểu hiện nhiều gen của thực vật có thể sử dụng các catxet biểu hiện đa bội không liên kết, mỗi catxet biểu hiện một protein khác hoặc tác nhân có độc tính khác như một hoặc nhiều phân tử ARN sợi đôi.

Các phân tử axit nucleic tái tổ hợp hoặc các cấu trúc ADN tái tổ hợp chứa trình tự mã hóa protein khảm diệt côn trùng có thể được phân phối đến các tế bào chủ nhờ vectơ, ví dụ, plasmit, baculovirut, nhiễm sắc thể tổng hợp, virion, cosmit, phagomit, thể thực khuẩn, hoặc vectơ virut. Các vectơ này có thể được dùng để đạt được sự biểu hiện ổn định hoặc tạm thời của trình tự mã hóa protein khảm diệt côn trùng ở tế bào chủ, hoặc sự biểu hiện tiếp theo của polypeptit được mã hóa. Polynucleotit tái tổ hợp hoặc cấu trúc ADN tái tổ hợp ngoại sinh mà chứa trình tự mã hóa trình tự protein khảm diệt côn trùng và được đưa vào tế bào chủ trong bản mô tả này được gọi là "gen chuyển".

Vi khuẩn chuyển gen, các tế bào thực vật chuyển gen, các cây chuyển gen, và bộ phận của cây chuyển gen mà chứa polynucleotit mà mã hóa một hoặc nhiều protein khảm diệt côn trùng bất kỳ được bộc lộ trong bản mô tả này. Thuật ngữ "tế bào vi khuẩn" hoặc "vi khuẩn" có thể bao gồm, nhưng không chỉ giới hạn ở, tế bào *Agrobacterium*, *Bacillus*, *Escherichia*, *Salmonella*, *Pseudomonas*, hoặc *Rhizobium*. Thuật ngữ "tế bào thực vật" hoặc "thực vật" có thể bao gồm, nhưng không chỉ giới hạn ở, tế bào hai lá mầm hoặc tế bào một lá mầm. Các cây và các tế bào thực vật đã dự định bao gồm, nhưng không chỉ giới hạn ở, cỏ linh lăng, chuối, lúa mạch, đỗ, bông cải xanh, cải bắp, cải bẹ xanh, ca rốt, sắn, thầu dầu, sup lơ, cần tây, đậu gà, cải thảo, cây có múi, dừa, cà phê, ngô, cỏ ba lá, cây bông, cây thuộc họ bầu bí, cây dưa chuột, cây thông Douglas, cà tím dài, khuynh điệp, cây lanh, tỏi, nho, hublong, tỏi tây, rau diếp, cây thông, hạt kê, quả dưa, quả hạch, yến mạch, ôliu, hành, cây cảnh, cọ, cỏ cảnh, đậu Hà lan, lạc, hạt tiêu, đậu triều, cây thông, khoai tây, cây dương, bí ngô, cây thông Radiata, củ cải, hạt cải dầu, gạo, cây thân rễ, lúa mạch đen, cây rum, cây ngải đắng, lúa miến, cây thông nam, đậu tương, rau bina, cây bí, dâu, củ cải đường, mía đường, hướng dương, ngô ngọt, gôm ngọt, khoai lang, cỏ mềm, chè, thuốc lá, cà chua, tiểu hắc

mạch, cỏ mặt đất, dưa hấu, và tế bào hoặc cây lúa mỳ. Theo các phương án nhất định, các cây chuyển gen và bộ phận của cây chuyển gen tái sinh được từ tế bào của cây chuyển gen được đề xuất. Theo các phương án nhất định, các cây chuyển gen có thể thu được từ hạt chuyển gen, bằng cách cắt, kẹp vỡ, nghiền nát hoặc phân tách theo cách khác phần này từ cây. Theo các phương án nhất định, phần của cây có thể là hạt, quả nang, lá, hoa, thân, rễ, hoặc phần bất kỳ của chúng, hoặc phần không thể tái sinh của bộ phận của cây chuyển gen. Phần "không thể tái sinh" của bộ phận của cây chuyển gen được dùng trong bản mô tả này là phần mà không thể kích thích được để tạo ra cả cây hoặc không thể kích thích được để tạo ra cả cây mà có khả năng sinh sản hữu tính và/hoặc sinh sản vô tính. Theo các phương án nhất định, phần không thể tái sinh của bộ phận của cây là phần của hạt chuyển gen, quả nang, lá, hoa, thân hoặc rễ.

Các phương pháp tạo ra các cây chuyển gen mà chứa các protein khảm diệt côn trùng với lượng có tác dụng ức chế bọ cánh vảy được đề xuất. Các cây này có thể được tạo thành bởi đưa polynucleotit mà mã hóa các protein khảm diệt côn trùng theo sáng chế vào tế bào thực vật, và chọn cây thu được từ tế bào thực vật đó mà biểu hiện protein khảm diệt côn trùng với lượng có tác dụng ức chế côn trùng hoặc bọ cánh vảy. Cây có thể thu được từ các tế bào thực vật nhờ tái sinh, hạt, phấn hoa, hoặc các kỹ thuật biến nạp mô phân sinh. Các phương pháp biến nạp thực vật là đã biết trong lĩnh vực kỹ thuật này. Ví dụ, việc biến nạp nhờ *Agrobacterium* được bộc lộ trong US 2009/0138985 A1 (cây đậu tương), US 2008/0280361A1 (cây đậu tương), US 2009/0142837A1 (cây ngô), US 2008/0282432 (cây bông), và US 2008/0256667 (cây bông).

Các cây biểu hiện các protein khảm diệt côn trùng có thể được lai bằng cách nhân giống với các thành phần chuyển gen biểu hiện các protein diệt côn trùng khác và/hoặc biểu hiện các tính trạng chuyển gen khác như tính trạng phòng diệt côn trùng khác, gen dung nạp thuốc trừ cỏ, gen tạo năng suất hoặc tính trạng chịu đựng stress, và tính trạng khác, hoặc các tính trạng này có thể được kết hợp vào một vectơ sao cho tất cả các tính trạng này được liên kết.

Các sản phẩm thực vật đã xử lý, trong đó sản phẩm đã xử lý chứa lượng phát hiện được của protein khảm diệt côn trùng, phân đoạn hoặc đoạn của chúng có tác dụng phòng diệt côn trùng, hoặc phần khác biệt bất kỳ của chúng, cũng được bộc lộ trong bản mô tả này. Theo các phương án nhất định, sản phẩm đã xử lý được chọn từ

nhóm bao gồm các bộ phận của cây, sinh khối thực vật, dầu, bột xay thô, đường, thức ăn chăn nuôi, bột, cốm, cám, thớ, vỏ, hạt đã xử lý, và hạt. Theo các phương án nhất định, sản phẩm đã xử lý không có khả năng tái sinh. Sản phẩm thực vật có thể là hàng hóa hoặc sản phẩm thương mại khác thu được từ cây chuyển gen hoặc bộ phận của cây chuyển gen, khi hàng hóa hoặc các sản phẩm khác có thể theo dõi được trong thương mại bằng cách phát hiện các đoạn nucleotit hoặc ARN đã được biểu hiện hoặc các protein mà mã hóa hoặc bao gồm phần khác biệt của protein khảm diệt côn trùng.

Các phương pháp phòng trừ côn trùng, cụ thể là sinh vật gây hại bộ cánh vảy cây trồng, bằng các protein khảm diệt côn trùng cũng được bộc lộ trong bản mô tả này. Các phương pháp này có thể bao gồm việc trồng cây chứa protein khảm diệt côn trùng với lượng ức chế côn trùng hoặc bọ cánh vảy. Theo các phương án nhất định, các phương pháp này còn có thể bao gồm một hoặc nhiều bước bất kỳ: (i) dùng chế phẩm bất kỳ chứa hoặc mã hóa protein khảm diệt côn trùng cho cây hoặc hạt để tạo ra cây; và (ii) biến nạp cây hoặc tế bào thực vật mà tạo ra cây với polynucleotit mã hóa protein khảm diệt côn trùng. Nói chung, dự định rằng protein khảm diệt côn trùng có thể được cung cấp ở dạng chế phẩm, cung cấp ở dạng vi sinh vật, hoặc cung cấp ở dạng thực vật chuyển gen để truyền hoạt tính ức chế côn trùng đối với côn trùng bọ cánh vảy.

Theo các phương án nhất định, protein khảm diệt côn trùng là thành phần có hoạt tính diệt côn trùng của chế phẩm ức chế côn trùng tạo ra được bằng cách nuôi cấy *Bacillus* tái tổ hợp hoặc tế bào vi khuẩn tái tổ hợp bất kỳ khác đã được biến nạp để biểu hiện protein khảm diệt côn trùng trong các điều kiện thích hợp để biểu hiện. Chế phẩm này có thể được tạo ra bằng cách làm khô, làm đông khô nhanh, đồng nhất hóa, chiết, lọc, ly tâm, kết lắng, hoặc cô dịch nuôi cấy tế bào tái tổ hợp này biểu hiện/tạo ra protein khảm diệt côn trùng. Quy trình này có thể tạo ra dịch chiết tế bào, huyền phù tế bào, dịch đồng nhất tế bào, dịch phân giải tế bào, dịch nổi tế bào, dịch lọc tế bào, hoặc viên tế bào của *Bacillus* hoặc vi khuẩn khác gây bệnh cho côn trùng. Nhờ nhận được protein khảm diệt côn trùng tạo ra được theo cách đó, chế phẩm mà chứa protein khảm diệt côn trùng có thể chứa tế bào vi khuẩn, bào tử vi khuẩn, và thể vùi của bào tử vỏ và có thể được tạo ra cho các mục đích sử dụng khác nhau, bao gồm làm sản phẩm xịt ức chế côn trùng trong nông nghiệp hoặc làm chế phẩm ức chế côn trùng trong thử nghiệm sinh học về chế độ ăn.

Hợp chất hoặc chế phẩm nêu trên còn có thể chứa chất mang nông dụng, như mồi, bột, bụi, viên, hạt nhỏ, bụi nước, nhũ tương, huyền phù thể keo, dung dịch nước, bào tử hoặc chế phẩm dạng tinh thể của *Bacillus* hoặc chế phẩm hạt xử lý. Hợp chất hoặc chế phẩm này còn có thể chứa thêm tế bào thực vật tái tổ hợp, mô thực vật, hạt hoặc cây đã được biến nạp để biểu hiện một hoặc nhiều protein; hoặc vi khuẩn đã được biến nạp để biểu hiện một hoặc nhiều protein. Tùy theo mức độ ức chế côn trùng hoặc ức chế diệt côn trùng vốn có ở polypeptit tái tổ hợp và lượng của hợp chất hoặc chế phẩm sẽ được dùng cho cây hoặc thử nghiệm về chế độ dinh dưỡng, hợp chất hoặc chế phẩm này có thể chứa trọng lượng khác nhau của polypeptit tái tổ hợp, ví dụ nằm trong khoảng từ 0,0001% đến 0,001% đến 0,01% đến 1% đến 99% trọng lượng polypeptit tái tổ hợp.

Theo một phương án, để làm giảm khả năng hình thành tính kháng, chế phẩm ức chế côn trùng hoặc cây chuyển gen chứa protein khảm diệt côn trùng còn có thể chứa ít nhất một tác nhân bổ sung có độc tính mà thể hiện hoạt tính ức chế côn trùng đối với cùng một loài côn trùng bọ cánh vảy, nhưng là khác với protein khảm diệt côn trùng. Các tác nhân bổ sung có độc tính có khả năng đối với chế phẩm này bao gồm protein ức chế côn trùng và phân tử ARN sợi đôi ức chế côn trùng. Một ví dụ về việc sử dụng các trình tự ribonucleotit này để phòng trừ các côn trùng gây hại được bộc lộ trong tài liệu của Baum và các đồng tác giả (US 2006/0021087 A1). (Các) polypeptit bổ sung như vậy để phòng trừ sinh vật gây hại bộ cánh vảy có thể được chọn từ nhóm bao gồm protein ức chế côn trùng như, nhưng không chỉ giới hạn ở, Cry1A (US 5,880,275), Cry1Ab, Cry1Ac, Cry1A,105, Cry1Ae, Cry1B (US 10/525,318), Cry1C (US 6,033,874), Cry1D, Cry1E, Cry1F, và các thể khảm Cry1A/F (US 7,070,982; US 6,962,705; và US 6,713,063), Cry1G, Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab (US 7,064,249), Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, Cry51Aa1, ET66, TIC400, TIC800, TIC834, TIC1415, Vip3A, VIP3Ab, VIP3B, AXMI-001, AXMI-002, AXMI-030, AXMI-035, và AXMI-045 (US 2013-0117884 A1), AXMI-52, AXMI-58, AXMI-88, AXMI-97, AXMI-102, AXMI-112, AXMI-117, AXMI-100 (US 2013-0310543 A1), AXMI-115, AXMI-113, AXMI-005 (US 2013-0104259 A1), AXMI-134 (US 2013-0167264 A1), AXMI-150 (US 2010-0160231 A1), AXMI-184 (US 2010-0004176 A1), AXMI-196, AXMI-204, AXMI-207, AXMI-209 (US 2011-0030096 A1), AXMI-218, AXMI-220 (US 2014-0245491

A1), AXMI-221z, AXMI-222z, AXMI-223z, AXMI-224z, AXMI-225z (US 2014-0196175 A1), AXMI-238 (US 2014-0033363 A1), AXMI-270 (US 2014-0223598 A1), AXMI-345 (US 2014-0373195 A1), DIG-3 (US 2013-0219570 A1), DIG-5 (US 2010-0317569 A1), DIG-11 (US 2010-0319093 A1), AfIP-1A và các chất dẫn xuất của nó (US 2014-0033361 A1), AfIP-1B và các chất dẫn xuất của nó (US 2014-0033361 A1), PIP-1APIP-1B (US 2014-0007292 A1), PSEEN3174 (US 2014-0007292 A1), AECFG-592740 (US 2014-0007292 A1), Pput_1063 (US 2014-0007292 A1), Pput_1064 (US 2014-0007292 A1), GS-135 và các chất dẫn xuất của nó (US 2012-0233726 A1), GS153 và các chất dẫn xuất của nó (US 2012-0192310 A1), GS154 và các chất dẫn xuất của nó (US 2012-0192310 A1), GS155 và các chất dẫn xuất của nó (US 2012-0192310 A1), SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2012-0167259 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2012-0047606 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2011-0154536 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2011-0112013 A1, SEQ ID NO:2 và 4 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0192256 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0077507 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0077508 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2009-0313721 A1, SEQ ID NO:2 hoặc 4 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0269221 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7,772,465 B2, CF161_0085 và các chất dẫn xuất của nó như đã được bộc lộ trong WO2014/008054 A2, các protein có độc tính đối với bọ cánh vảy và các chất dẫn xuất của nó như đã được bộc lộ trong US 2008-0172762 A1, US 2011-0055968 A1, và US 2012-0117690 A1; SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7510878 B2, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7812129 B1; và các polypeptit tương tự.

Theo các phương án khác, chế phẩm ức chế côn trùng hoặc cây chuyển gen còn có thể chứa ít nhất một tác nhân bổ sung có độc tính thể hiện hoạt tính ức chế côn trùng đối với côn trùng gây hại mà không bị ức chế bởi các protein khảm diệt côn trùng theo sáng chế (như sinh vật gây hại là bọ cánh cứng, bọ cánh nửa và cánh giống), để mở rộng phổ ức chế côn trùng thu được.

Tác nhân bổ sung có độc tính này để phòng diệt sinh vật gây hại là bọ cánh cứng có thể được chọn từ nhóm bao gồm protein ức chế côn trùng như, nhưng không chỉ giới hạn ở, Cry3Bb (US 6,501,009), các biến thể của Cry1C, các biến thể của Cry3A, Cry3, Cry3B, Cry34/35, 5307, AXMI134 (US 2013-0167264 A1) AXMI-184 (US 2010-0004176 A1), AXMI-205 (US 2014-0298538 A1), axmi207 (US 2013-0303440 A1), AXMI-218, AXMI-220 (US 20140245491A1), AXMI-221z, AXMI-223z (US 2014-0196175 A1), AXMI-279 (US 2014-0223599 A1), AXMI-R1 và các biến thể của nó (US 2010-0197592 A1, TIC407, TIC417, TIC431, TIC807, TIC853, TIC901, TIC1201, TIC3131, DIG-10 (US 2010-0319092 A1), eHIP (US 2010/0017914), IP3 và các biến thể của nó (US 2012-0210462 A1), và α-Hexatoxin-Hv1a (US 2014-0366227 A1).

Tác nhân bổ sung có độc tính như vậy để phòng diệt sinh vật gây hại là bọ cánh nửa có thể được chọn từ nhóm bao gồm các protein có hoạt tính đối với bọ cánh nửa như, nhưng không chỉ giới hạn ở, TIC1415 (US 2013-0097735 A1), TIC807 (US 8609936), TIC834 (US 2013-0269060 A1), AXMI-036 (US 2010-0137216 A1), và AXMI-171 (US 2013-0055469 A1). Các polypeptit bổ sung để phòng diệt côn trùng gây hại là bọ cánh cứng, bọ cánh vảy, và bọ cánh nửa có thể được tìm thấy trên trang web về danh pháp độc tố của *Bacillus thuringiensis* duy trì bởi Neil Crickmore (trên trang web toàn cầu tại địa chỉ btnomenclature.info).

Các trình tự mã hóa protein khảm diệt côn trùng và các trình tự có mức độ đồng nhất đáng kể tính theo phần trăm với các protein khảm diệt côn trùng có thể được xác định bằng cách áp dụng các phương pháp mà người có hiểu biết trung bình về lĩnh vực kỹ thuật này đã biết như phản ứng chuỗi polymeraza (PCR), khuếch đại bằng nhiệt và lai. Ví dụ, các protein khảm diệt côn trùng có thể được dùng để tạo ra các kháng thể mà liên kết đặc hiệu với protein liên quan, và có thể được dùng để sàng lọc và tìm ra các protein khác mà liên quan gần.

Hơn nữa, các trình tự nucleotit mã hóa các protein khảm diệt côn trùng có thể được dùng làm mẫu dò và các đoạn mồi để sàng lọc nhằm xác định các thành viên khác của nhóm này bằng cách áp dụng việc khuếch đại theo chu kỳ nhiệt hoặc đẳng nhiệt và các phương pháp lai. Ví dụ, các oligonucleotit thu được từ các trình tự đã nêu trong SEQ ID NO:2 có thể được dùng để xác định sự có mặt hoặc sự không có mặt của gen chuyển khảm diệt côn trùng ở mẫu axit đeoxyribonucleic thu được từ sản phẩm

thương mại. Nhờ độ nhạy của một số phương pháp phát hiện axit nucleic nhất định mà sử dụng các oligonucleotit, dự định rằng các oligonucleotit thu được từ các trình tự bất kỳ nêu trong SEQ ID NO:2 có thể được dùng để phát hiện protein khảm diệt côn trùng tương ứng ở các sản phẩm hàng hóa thu được từ các nguồn đã gom khi chỉ một phân đoạn của sản phẩm hàng hóa thu được từ cây chuyển gen chứa trình tự bất kỳ nêu trong SEQ ID NO:2.

Ví dụ thực hiện sáng chế

Trên cơ sở nội dung nêu trên, người có hiểu biết trung bình về lĩnh vực kỹ thuật này cần phải hiểu rằng các phương án được bộc lộ dưới đây chỉ là đại diện cho sáng chế, mà có thể bao gồm các dạng khác nhau. Do đó, các chi tiết cụ thể về cấu trúc và chức năng đã được bộc lộ trong bản mô tả này không nhằm giới hạn phạm vi yêu cầu bảo hộ của sáng chế.

Ví dụ 1. Tạo và tách dòng trình tự mã hóa protein khảm mới có hoạt tính diệt côn trùng đối với bọ cánh vảy

Ví dụ này minh họa cách tạo ra các protein khảm mới diệt côn trùng và tách dòng và biểu hiện các protein khảm diệt côn trùng này.

Các trình tự axit nucleic tái tổ hợp được xây dựng từ gen protein Cry đã biết để tạo ra các trình tự polynucleotit mã hóa các protein khảm mới diệt côn trùng. Các trình tự polynucleotit thu được được tách dòng vào vectơ plasmit biểu hiện *Bacillus thuringiensis* (*Bt*). Sau khi xác nhận trình tự polynucleotit, plasmit biểu hiện được biến nạp vào *Bt* và được biểu hiện. Các chế phẩm chứa các protein khảm mới đã biểu hiện được thử nghiệm về hoạt tính đối với các sinh vật gây hại bộ cánh vảy khác nhau.

Nhiều trình tự polynucleotit mã hóa các protein khảm diệt côn trùng được tạo ra và được thử nghiệm trong thử nghiệm sinh học. Không phải tất cả các protein khảm diệt côn trùng đều thể hiện hoạt tính. Chỉ một vài protein trong số các protein khảm diệt côn trùng được chọn lọc dựa trên hoạt tính của chúng đối với bọ cánh vảy cụ thể được thể hiện trong thử nghiệm sinh học. Các biến thể axit amin, trong đó các thay thế axit amin, hoặc miền xen kẽ của tiền độc tố, được đưa vào cũng được tạo ra dựa trên các protein khảm gốc có hoạt tính diệt côn trùng TIC867 và TIC868. Các thành phần của các protein khảm diệt côn trùng (các miền I, II và III và tiền độc tố) theo sáng chế

được thể hiện trong Bảng 1. Các thay thế axit amin ở biến thể của TIC868 so với trình tự protein TIC868 gốc cũng được thể hiện.

Bảng 1. Các protein khảm mới diệt sinh vật gây hại và thành phần của chúng.

Độc tố	PRT SEQ ID NO:	Miền 1	Miền 2	Miền 3	Tiền độc tố	Cải biến axit amins*
TIC1100	4	Cry1Ah	Cry1Ac	Cry1Ca	Cry1Ac	
TIC860	7	Cry1Bb1	Cry1Bb1	Cry1Ca	Cry1Ac	
TIC867	10	Cry1Be2	Cry1Be2	Cry1Ka	Cry1Ab3	
TIC867_20	13	Cry1Be2	Cry1Be2	Cry1Ka	Cry1Da1	
TIC867_21	16	Cry1Be2	Cry1Be2	Cry1Ka	Cry4	
TIC867_22	19	Cry1Be2	Cry1Be2	Cry1Ka	Cry9	
TIC867_23	21	Cry1Be2	Cry1Be2	Cry1Ka	Cry1Be	
TIC867_24	23	Cry1Be2	Cry1Be2	Cry1Ka	Cry1Ka	
TIC867_25	25	Cry1Be2	Cry1Be2	Cry1Ka	Cry1Ca	
TIC868	28	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Ab3	
TIC868_9	30	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Ab3	N240S_Y343Q_N349T
TIC868_10	33	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Da1	
TIC868_11	36	Cry1Be2	Cry1Be2	Cry1Ca	Cry4	
TIC868_12	39	Cry1Be2	Cry1Be2	Cry1Ca	Cry9	
TIC868_13	41	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Be	
TIC868_14	43	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Ka	
TIC868_15	45	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Ca	
TIC868_29	47	Cry1Be2	Cry1Be2	Cry1Ca	Cry1Ab3	Q136Y_Y343Q_N349T
TIC869	50	Cry1Ja1	Cry1Ja1	Cry1Jx	Cry1Ab3	
TIC836	53	Cry1Fa1	Cry1Fa1	Cry1Ab	Cry1Ac	

*Các đột biến axit amin được xác định bằng cách sử dụng mã axit amin chuẩn theo IUPAC. Xem tài liệu: IUPAC-IUB Joint Commission on Biochemical Nomenclature. Nomenclature and Symbolism for Amino Acids and Peptides. *Eur. J. Biochem.* 138:9-37(1984). Chữ viết tắt trình tự axit amin thứ nhất biểu thị gốc axit amin trong protein khung đã định, chữ số thể hiện vị trí của axit amin, và chữ viết tắt trình tự axit amin thứ hai biểu thị axit amin được đặt ở vị trí đó trong biến thể protein cải tiến.

Ví dụ 2. Protein khảm mới có hoạt tính diệt côn trùng thể hiện hoạt tính đối với sinh vật gây hại bộ cánh vảy

Ví dụ này minh họa việc thử nghiệm các protein khảm diệt côn trùng được bộc lộ trong Ví dụ 1 và hoạt tính đối với bọ cánh vảy quan sát được đối với các protein khảm diệt côn trùng.

Các trình tự polynucleotit mã hóa các protein khảm diệt côn trùng được biểu hiện ở *Bt*. Sau đó, các protein khảm diệt côn trùng đã biểu hiện được thử nghiệm đối với nhiều bọ cánh vảy đã biết là các sinh vật gây hại cho cây ngô, cây mía đường, cây đậu tương và cây bông, cũng như các cây trồng khác. Đặc biệt, các protein diệt côn trùng được thử nghiệm về hoạt tính đối với sâu bướm hại đậu mèo (VBC, *Anticarsia gemmatalis*), sâu đục thân mía (SCB, *Diatraea saccharalis*), sâu đục thân ngô nhỏ (LSCB, *Elasmopalpus lignosellus*), sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu đen ăn lá (BLAW, *Spodoptera cosmioides*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (BAW, *Spodoptera exigua*), sâu đục quả cựu thế giới (OBW, *Helicoverpa armigera*), sâu ăn lá phương đông (OLW, *Spodoptera litura*), sâu hồng đục quả (PBW, *Pectinophora gossypiella*), sâu xám (BCW, *Agrotis ipsilon*), sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*), sâu đục quả đốm (SBW, *Earias vitella*), và sâu đục thân ngô châu Âu (ECB, *Ostrinia nubilalis*). Sâu đục trái hại ngô (CEW, *Helicoverpa zea*) còn được gọi là sâu vỏ đậu tương (SPW) và sâu hại bông (CBW). Hoạt tính được xác định nhờ việc kết hợp mức độ gây chết và mức độ còi cọc cũng như mức MIC50. MIC50 là nồng độ ức chế lột xác, trong đó cả ấu trùng chết và ấu trùng L1 (ấu trùng mà không lột xác được lần thứ hai) đều được tính đến ở kết quả. Bảng 2 thể hiện hoạt tính của từng protein khảm diệt côn trùng. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể.

Bảng 2. Hoạt tính theo thử nghiệm sinh học đối với bọ cánh vảy chọn lọc.

Độc tố	PRT SEQ ID NO:	Côn trùng													
		VBC	SCB	LSCB	CEW SPW CBW	BLAW	TBW	SBL	SAW	FAW	BAW	OBW	OLW	PBW	BCW	SWCB	ECB	SBW
TIC1100	4	+	+			+			+		+		+	+				
TIC860	7	+	+	+		+	+	+	+		+		+	+		+		+
TIC867	10	+	+			+		+		+	+	+	+			+		
TIC867_20	13																	
TIC867_21	16				+													
TIC867_22	19				+					+								
TIC868	28	+	+			+		+		+	+		+	+		+		
TIC868_10	33									+								
TIC868_11	36									+								
TIC868_12	39									+								
TIC869	50	+	+					+		+						+		
TIC836	53	+				+		+	+	+								

Như có thể thấy trong Bảng 2 trên đây, hầu hết các protein khảm diệt côn trùng thể hiện hoạt tính đối với một hoặc nhiều loài sinh vật gây hại bộ cánh vảy.

Ví dụ 3. Tổng hợp gen mã hóa protein khảm diệt côn trùng và để biểu hiện ở thực vật

Ví dụ này minh họa quy trình tổng hợp các polynucleotit mã hóa các protein khảm diệt côn trùng để biểu hiện ở thực vật.

Các trình tự mã hóa tổng hợp được xây dựng để dùng nhằm biểu hiện các protein khảm diệt côn trùng ở thực vật. Các trình tự tổng hợp này được thiết kế và tổng hợp theo các phương pháp đã được bộc lộ chung trong US 5,500,365, tránh một số trình tự nhất định có vấn đề về tính độc hại như ATTTA và các trình tự polyađenyl hóa ở thực vật giàu A/T trong khi bảo toàn trình tự axit amin của protein khảm diệt côn trùng. Các trình tự nucleotit cho các gen này mã hóa các protein khảm diệt côn trùng để biểu hiện ở thực vật được liệt kê trong Bảng 3.

Bảng 3. Các trình tự polynucleotit mã hóa các protein khảm diệt côn trùng được thiết kế để dùng ở thực vật.

Protein diệt côn trùng	ADN SEQ ID NO:	PRT SEQ ID NO:
TIC1100	2	4
TIC1100	3	4
TIC860	6	7
TIC867	9	10
TIC867_20	12	13
TIC867_21	15	16
TIC867_22	18	19
TIC867_23	20	21
TIC867_24	22	23
TIC867_25	24	25
TIC868	27	28
TIC868_9	29	30
TIC868_10	32	33
TIC868_11	35	36
TIC868_12	38	39
TIC868_13	40	41

Protein diệt côn trùng	ADN SEQ ID NO:	PRT SEQ ID NO:
TIC868_14	42	43
TIC868_15	44	45
TIC868_29	46	47
TIC869	49	50
TIC836	52	53

Ví dụ 4. Catxet biểu hiện để biểu hiện protein khảm diệt côn trùng ở thực vật

Ví dụ này minh họa việc xây dựng catxet biểu hiện chứa trình tự polynucleotit được thiết kế để dùng ở thực vật mà mã hóa protein khảm diệt côn trùng.

Nhiều loại catxet biểu hiện ở thực vật được xây dựng với các trình tự polynucleotit mã hóa các protein khảm diệt côn trùng được thiết kế để biểu hiện ở thực vật được thể hiện trong Bảng 3. Các catxet biểu hiện này là hữu ích để biểu hiện tạm thời trong thể nguyên sinh thực vật hoặc biến nạp tế bào thực vật. Các catxet biểu hiện điển hình được thiết kế liên quan đến việc cuối cùng đưa protein vào trong tế bào. Một bộ catxet biểu hiện được thiết kế theo cách để cho phép protein được dịch mã và ở lại trong dung dịch bào tương. Một bộ catxet biểu hiện khác được thiết kế để có peptit chuyển tiếp liền kề với protein của độc tố này để cho phép hướng đích đến hạt cơ quan của tế bào như lạp lục hoặc thể hạt. Tất cả các catxet biểu hiện được thiết kế để bắt đầu ở đầu tận cùng 5' với gen khởi đầu, mà có thể bao gồm nhiều thành phần trình tự khởi đầu, các thành phần gen tăng cường, hoặc các thành phần biểu hiện khác mà người có hiểu biết trung bình về lĩnh vực kỹ thuật này đã biết là được liên kết hoạt động để thúc đẩy sự biểu hiện của gen chuyển. Ngay tiếp sau trình tự của trình tự khởi đầu thường là một hoặc nhiều trình tự dẫn đầu 3' so với trình tự khởi đầu. Trình tự intron thường nằm ở phía 3' so với trình tự dẫn đầu để cải thiện mức độ biểu hiện của gen chuyển. Trình tự mã hóa cho độc tố hoặc peptit chuyển tiếp và trình tự mã hóa cho độc tố thường nằm ở phía 3' so với trình tự khởi đầu được liên kết hoạt động, gen dẫn đầu và cấu hình intron. Trình tự 3'UTR thường nằm ở phía 3' của trình tự mã hóa để tạo điều kiện thuận lợi cho việc kết thúc phiên mã và để tạo ra các trình tự quan trọng đối với việc polyađenyl hóa sản phẩm phiên mã thu được. Tất cả các thành phần nêu trên được liên kết hoạt động và được bố trí liên tục, thường cùng với các trình tự bổ sung được cung cấp để xây dựng catxet biểu hiện.

Ví dụ 5. Hoạt tính đối với bọ cánh vảy của protein khảm diệt côn trùng ở cây ngô đã được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế thể hiện bởi các protein khảm diệt côn trùng đối với sinh vật gây hại bộ cánh vảy khi được biểu hiện ở cây ngô và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại cây ngô tương ứng.

Giống ngô LH244 được chuyển nhiễm với các vectơ biến nạp lưỡng phân được bộc lộ trong Ví dụ 4 bằng cách áp dụng phương pháp biến nạp gián tiếp qua *Agrobacterium*. Tế bào đã biến nạp được kích thích để tạo ra cây theo các phương pháp đã biết trong lĩnh vực kỹ thuật này. Các thử nghiệm sinh học bằng cách sử dụng phiến lá thực vật được thực hiện tương tự các thử nghiệm đã được bộc lộ trong US 8,344,207. Cây LH244 không biến nạp được dùng để thu mô nhằm dùng làm đối chứng âm tính. Nhiều thể biến nạp từ từng vectơ lưỡng phân được đánh giá đối với sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu xám (BCW, *Agrotis ipsilon*) và sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*).

Thử nghiệm sinh học phiến lá được thực hiện trên các cây chuyển gen thế hệ R_0 và F_1. Ngoài ra, mức độ tổn thương lá được đánh giá cho toàn bộ cây F_1 chuyển gen biểu hiện một số protein khảm diệt côn trùng nhất định đã bị phá hại bởi côn trùng sinh vật gây hại bộ cánh vảy. Các thành phần chuyển gen F_1 biểu hiện TIC860 và TIC868 cũng được đánh giá về hoạt tính trên thực địa đối với FAW, CEW, và SWCB. Các kết quả thử nghiệm này được thể hiện trong Bảng 4. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể. Như có thể thấy trong Bảng 4, hầu hết các protein khảm diệt côn trùng và nhiều biến thể trong số các biến thể của protein khảm diệt côn trùng thể hiện hoạt tính đối với một hoặc nhiều loài sinh vật gây hại bộ cánh vảy.

Bảng 4. Hoạt tính theo thử nghiệm sinh học của protein khảm diệt côn trùng từ mô lá ngô được chuyển nhiễm một cách ổn định

Bộ tế	PRT SEQ ID NO:	Côn trùng														
		VBC	SCB	LSCB	CEW SPW CBW	BLAW	TBW	SBL	SAW	FAW	BAW	OBW	OLW	PBW	BCW	SWCB	ECB	SBW
TIC1100	4	+	+			+			+		+		+	+				
TIC860	7	+	+	+		+	+	+	+	+	+		+	+		+		+
TIC867	10	+	+			+		+		+	+	+	+			+		
TIC867_20	13	không thử	không g thử	không g thử		không g thử	không thử	không thử	không thử		không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC867_21	16	không thử	không g thử	không g thử	+	không g thử	không thử	không thử	không thử		không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC867_22	19	không thử	không g thử	không g thử	+	không g thử	không thử	không thử	không thử	+	không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC868	28	+	+			+		+	+	+	+		+	+		+		+
TIC868_10	33	không thử	không g thử	không g thử	+	không g thử	không thử	không thử	không thử	+	không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC868_11	36	không thử	không g thử	không g thử	+	không g thử	không thử	không thử	không thử	+	không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC868_12	39	không thử	không g thử	không g thử	+	không g thử	không thử	không thử	không thử	+	không thử	không thử	không thử	không thử	không thử	không thử	không thử	không thử
TIC869	50	+	+					+	+	+						+		
TIC836	53	+				+		+	+	+								

Ví dụ 6. Hoạt tính đối với bọ cánh vảy của các protein khảm diệt côn trùng ở cây đậu tương đã được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế thể hiện bởi các protein khảm diệt côn trùng đối với sinh vật gây hại bộ cánh vảy khi được biểu hiện ở cây đậu tương và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại tương ứng.

Các trình tự mã hóa cho các protein khảm diệt côn trùng đã chọn được thiết kế lại để biểu hiện ở thực vật, đã được tách dòng vào vectơ biến nạp lưỡng phân thực vật, và sử dụng để biến nạp các tế bào thực vật của cây đậu tương. Các vectơ biến nạp thực vật bao gồm catxet gen chuyển thứ nhất để biểu hiện protein khảm diệt côn trùng như được mô tả trong Ví dụ 4 và catxet gen chuyển thứ hai để chọn lọc các tế bào thực vật đã biến nạp bằng cách áp dụng chọn lọc bằng spectinomyxin. Trong một số trường hợp, như trong trường hợp của TIC1100, TIC860 và TIC836, peptit chuyển tiếp ở lạp lục mã hóa trình tự được liên kết hoạt động với trình tự khảm diệt côn trùng mã hóa. Các thử nghiệm được thực hiện với TIC1100, TIC860 và TIC836 hướng đích và không hướng đích đến thể hạt. Bảng 5 dưới đây thể hiện protein khảm diệt côn trùng và biến thể của protein khảm diệt côn trùng TIC867 và các trình tự mã hóa kèm theo được dùng để biểu hiện ở cây đậu tương đã được chuyển nhiễm một cách ổn định.

Các tế bào thực vật của cây đậu tương được chuyển nhiễm bằng cách sử dụng các vectơ biến nạp lưỡng phân nêu trên bằng cách biến nạp qua trung gian *Agrobacterium*. Các tế bào thực vật đã biến nạp thu được được kích thích để tạo ra cả cây đậu tương. Mô lá được thu gom và sử dụng trong thử nghiệm sinh học như được mô tả trong Ví dụ 5 hoặc theo cách khác, mô đã đông khô nhanh được sử dụng trong chế độ dinh dưỡng cho côn trùng để thử nghiệm sinh học. Thử nghiệm sinh học được thực hiện đối với FAW, sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu vỏ đậu tương (SPW, *Helicoverpa zea*), sâu bướm hại đậu mèo (VBC, *Anticarsia gemmatalis*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu đen ăn lá (BLAW, *Spodoptera cosmioides*), sâu đục thân ngô nhỏ (LSCB, *Elasmopalpus lignosellus*) và sâu đục quả cựu thế giới (OBW, *Helicoverpa armigera*).

Bảng 5 thể hiện hoạt tính đối với các loài bọ cánh vảy chọn lọc của từng protein diệt côn trùng ở các cây thế hệ R_0, trong đó '+' biểu thị hoạt tính. Như có thể thấy

trong Bảng 5, mỗi protein trong số các protein khảm diệt côn trùng được biểu hiện ở cây đậu tương đã được chuyển nhiễm một cách ổn định thể hiện hoạt tính đối với nhiều loài bọ cánh vảy. Đặc biệt lưu ý rằng biến thể của TIC867 là TIC867_23 thể hiện hoạt tính đối với SPW.

Bảng 5. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô lá đậu tương R_0 đã được chuyển nhiễm một cách ổn định.

Protein diệt côn trùng	FAW	SAW	SBL	SPW	VBC	TBW	BLAW	LSCB	OBW
TIC1100	+	+	+		+		+	+	+
TIC860	+	+	+		+			+	
TIC867	+	+	+		+	+		+	
TIC867_20		+	+						
TIC867_21		+	+						
TIC867_22		+	+						
TIC867_23	+	+	+	+					
TIC867_24		+	+						
TIC867_25		+	+						
TIC868	+		+		+		+	+	
TIC869			+		+	+		+	
TIC836	+	+	+		+	+	+		+

Các thể biến nạp chọn lọc được để tự thụ phấn và hạt thu được được gieo. Mô lá được thu gom từ các cây thế hệ R_1 và được dùng trong thử nghiệm sinh học cho ăn. Các cây R_1 biểu hiện TIC1100, TIC860, TIC867, TIC868, TIC869 và TIC836 được thử nghiệm về hoạt tính đối với SAW, SBL, SPW và VBC. Bảng 6 thể hiện hoạt tính quan sát được ở các thử nghiệm này. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể. Như được thể hiện trong Bảng 6, hầu hết các protein khảm diệt côn trùng đã được biểu hiện từ các cây thế hệ R_1 thể hiện hoạt tính đối với một hoặc nhiều loài bọ cánh vảy.

Bảng 6. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô lá đậu tương thế hệ R_1 đã được chuyển nhiễm một cách ổn định.

Độc tố	SAW	SBL	SPW	VBC
TIC1100	+	+		+
TIC860	+	+		+
TIC867	+			
TIC868	+	+		+
TIC869	+	+		+
TIC836	+	+		+

Bảng 7 thể hiện kết quả của thử nghiệm thực địa được thực hiện trong nhà kính với cây đậu tương thế hệ R_1 đã được chuyển nhiễm một cách ổn định biểu hiện TIC1100, TIC860, và TIC836. Các loài được sử dụng để gây hại cây trong nhà kính bao gồm SAW, SBL và SPW. Mức độ kháng được xác định là rụng lá thấp hơn hoặc bằng mười lăm phần trăm ở cây đậu tương. Mức độ kháng quan sát được ở các thử nghiệm trong lồng phù hợp với mức độ kháng quan sát được ở thử nghiệm trên mô lá đậu tương thế hệ R_1 thể hiện trong Bảng 6. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể.

Bảng 7. Biên dạng hoạt tính của TIC1100, TIC860 và TIC836 được biểu hiện ở cây đậu tương thế hệ R_1 được thử nghiệm trong thử nghiệm thực địa trong nhà kính.

Độc tố	SAW	SBL	SPW
TIC1100	+	+	
TIC860	+	+	
TIC836	+	+	

Các thử nghiệm thực địa trong nhà kính với các cây đậu tương thế hệ R_1 đã được chuyển nhiễm một cách ổn định biểu hiện TIC867 và TIC869 cũng được thực hiện tại hai nơi khác nhau ở Argentina là Acevedo và Fontezuela. Các loài được sử dụng để gây hại cây trong nhà kính bao gồm sâu đục quả Nam Mỹ (SABW, *Helicoverpa gelotopeon*), VBC, BLAW, và sâu đo hướng dương (SFL, *Rachiplusia nu*). Mức độ kháng được xác định là rụng lá thấp hơn hoặc bằng mười lăm phần trăm ở cây đậu tương. Bảng 8 dưới đây thể hiện tính kháng quan sát được. Ký hiệu '+' biểu thị hoạt

tính quan sát được đối với côn trùng gây hại cụ thể. Như được thể hiện trong Bảng 8, các cây đậu tương chuyển gen biểu hiện TIC867 thể hiện tính kháng BLAW và VBC. Các cây đậu tương chuyển gen biểu hiện TIC869 thể hiện tính kháng SABW, SFL, BLAW, và VBC.

Bảng 8. Biên dạng hoạt tính của TIC867 và TIC869 được biểu hiện ở cây đậu tương thế hệ R_1 được thử ở thử nghiệm thực địa trong nhà kính.

	Acevedo			Fontezuela		
Độc tố	SABW	SFL	VBC	SABW	BLAW	VBC
TIC867			+		+	+
TIC869		+	+	+	+	+

Ví dụ 7. Hoạt tính đối với bọ cánh vảy của các protein khảm diệt côn trùng ở cây bông đã được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế thể hiện bởi các protein khảm diệt côn trùng đối với sinh vật gây hại bộ cánh vảy khi được biểu hiện ở cây bông và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại tương ứng.

Các trình tự mã hóa cho các protein khảm diệt côn trùng chọn lọc được thiết kế lại để biểu hiện ở thực vật, tách dòng vào vectơ biến nạp lưỡng phân thực vật, và sử dụng để biến nạp các tế bào thực vật của cây bông. Các vectơ lưỡng phân thu được là tương tự các vật truyền đã được bộc lộ trong Ví dụ 4 và được sử dụng để biểu hiện TIC860 hướng đích và không hướng đích đến thể hạt (trình tự mã hóa: SEQ ID NO: 6; trình tự protein: SEQ ID NO: 7), TIC867 (trình tự mã hóa: SEQ ID NO: 9; trình tự protein: SEQ ID NO: 10), TIC868 (trình tự mã hóa: SEQ ID NO: 27; trình tự protein: SEQ ID NO: 28) và TIC867_23 (trình tự mã hóa: SEQ ID NO: 20; trình tự protein: SEQ ID NO: 23).

Các tế bào thực vật của cây bông được biến nạp theo phương pháp biến nạp gián tiếp qua *Agrobacterium*. Các tế bào đã biến nạp của cây bông được kích thích để tạo ra cả cây. Mô của lá cây bông được sử dụng trong thử nghiệm sinh học như được mô tả trong Ví dụ 5 đối với sâu đục quả bông (CBW, *Helicoverpa zea*), FAW, TBW và SBL. Bảng 9 thể hiện hoạt tính quan sát được đối với các loài bọ cánh vảy này của TIC860, TIC867, và TIC868 ở cây bông thế hệ R_0 đã được chuyển nhiễm một cách ổn định,

trong đó '+' biểu thị hoạt tính. Như có thể thấy trong Bảng 9, TIC860, TIC867, và TIC868 thể hiện hoạt tính đối với hai hoặc nhiều loài sinh vật gây hại bộ cánh vảy ở cây bông thế hệ R_0 đã được chuyển nhiễm một cách ổn định.

Bảng 9. Hoạt tính theo thử nghiệm sinh học của TIC860, TIC867 và TIC868 từ mô của lá cây bông thế hệ R_0 đã được chuyển nhiễm một cách ổn định.

Độc tố	CBW	FAW	TBW	SBL
TIC860		+		+
TIC867	+	+	+	không thử
TIC868		+		+

Các thể biến nạp chọn lọc được dùng để tạo ra hạt thế hệ R_1. Các cây thế hệ R_1 biểu hiện TIC860, TIC867, và TIC868 được thử nghiệm về tính kháng CBW, FAW, TBW, và SBL. Mô lá, mô vuông và mô quả nang được dùng trong thử nghiệm. Bảng 10 thể hiện hoạt tính quan sát được ở các thử nghiệm này. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể. Như được thể hiện trong Bảng 10, TIC860 thể hiện hoạt tính đối với FAW ở mô lá. Ngoài ra, protein khảm diệt côn trùng TIC867 thể hiện hoạt tính đối với CBW và FAW ở mô lá, mô vuông và mô quả nang, cũng như TBW và SBL ở lá. Protein khảm diệt côn trùng TIC868 thể hiện hoạt tính đối với FAW ở mô lá, mô vuông và mô quả nang, cũng như TBW và SBL ở lá.

Bảng 10. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô của lá cây bông thế hệ R_1 đã được chuyển nhiễm một cách ổn định.

	CBW			FAW			TBW	SBL
Độc tố	Lá	Vuông	Quả nang	Lá	Vuông	Quả nang	Lá	Lá
TIC860				+				
TIC867	+	+	+	+	+	+	+	+
TIC868				+	+	+	+	+

Ví dụ 8. Tách dòng trình tự mã hóa protein khảm mới diệt côn trùng có hoạt tính đối với bọ cánh vảy

Các trình tự axit nucleic tái tổ hợp được xây dựng từ gen protein Cry đã biết sử dụng để tạo ra các trình tự mã hóa mã hóa các protein khảm mới diệt côn trùng. Các

trình tự mã hóa thu được được tách dòng vào vectơ plasmit biểu hiện *Bacillus thuringiensis* (Bt). Sau khi xác nhận trình tự đã được tách dòng, plasmit biểu hiện được biến nạp vào Bt và được biểu hiện. Các chế phẩm chứa các protein Cry khảm mới đã biểu hiện được thử nghiệm về hoạt tính đối với các sinh vật gây hại bộ cánh vảy khác nhau. Nhiều trình tự mã hóa mã hóa các protein khảm diệt côn trùng được tạo ra và được thử nghiệm trong thử nghiệm sinh học. Không phải tất cả các protein khảm diệt côn trùng đều thể hiện hoạt tính. Chỉ một vài trong số các protein khảm diệt côn trùng được chọn lọc dựa trên hoạt tính của chúng đối với bọ cánh vảy cụ thể, và được thể hiện trong Bảng 11 dưới đây.

Bảng 11. Các protein khảm mới diệt sinh vật gây hại và trình tự mã hóa vi khuẩn tương ứng của chúng và trình tự protein.

Protein diệt sinh vật gây hại	ADN của vi khuẩn SEQ ID NO:	Protein SEQ ID NO:
TIC713	54	55
TIC843	56	57
TIC862	58	59
TIC1099	60	61
TIC1103	78	79
TIC1101	80	81
TIC845	82	83
TIC846	84	85
TIC858	86	87
TIC865	88	89
TIC866	90	91
TIC838	92	93
TIC839	94	95
TIC841	96	97
TIC842	98	99

Protein diệt sinh vật gây hại	ADN của vi khuẩn SEQ ID NO:	Protein SEQ ID NO:
TIC850	100	101
TIC859	102	103
TIC861	104	105
TIC848	106	107
TIC849	108	109
TIC847	110	111

Ví dụ 9. Các protein khảm mới có hoạt tính diệt côn trùng thể hiện hoạt tính đối với sinh vật gây hại bộ cánh vảy

Các trình tự mã hóa mã hóa các protein khảm diệt côn trùng được biểu hiện ở *Bt*. Sau đó, các protein đã được biểu hiện được thử nghiệm đối với nhiều loại bọ cánh vảy đã biết là các sinh vật gây hại cho cây ngô, cây mía, cây đậu tương và cây bông, cũng như các cây trồng khác. Các protein diệt côn trùng được thử nghiệm về hoạt tính đối với sâu bướm hại đậu mèo (VBC, *Anticarsia gemmatalis*), sâu đục thân mía (SCB, *Diatraea saccharalis*), sâu đục thân ngô nhỏ (LSCB, *Elasmopalpus lignosellus*), sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu đen ăn lá (BLAW, *Spodoptera cosmioides*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (BAW, *Spodoptera exigua*), sâu đục quả cựu thế giới (OBW, *Helicoverpa armigera*), sâu ăn lá phương đông (OLW, *Spodoptera litura*), sâu hồng đục quả (PBW, *Pectinophora gossypiella*), sâu xám (BCW, *Agrotis ipsilon*), sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*) và sâu đục thân ngô châu Âu (ECB, *Ostrinia nubilalis*). sâu đục trái hại ngô (CEW, *Helicoverpa zea*) còn được gọi là sâu vỏ đậu tương (SPW) và sâu hại quả bông (CBW). Hoạt tính được xác định nhờ việc kết hợp tỷ lệ chết và tỷ lệ còi cọc cũng như mức MIC50. MIC50 chỉ nồng độ ức chế lột xác trong đó cả ấu trùng chết và ấu trùng L1 (các ấu trùng không lột xác lần thứ hai) đều được tính đến ở kết quả. Bảng 12 dưới đây thể hiện hoạt tính của từng protein khảm diệt côn trùng. Ký hiệu '+' biểu thị hoạt tính quan sát được đối với côn trùng gây hại cụ thể.

Bảng 12. Hoạt tính theo thử nghiệm sinh học đối với bọ cánh vảy chọn lọc.

	Côn trùng															
PRT SEQ ID NO:	VBC	SCB	LSCB	CEW SPW CBW	TBW	SBL	BLAW	SAW	FAW	BAW	OBW	OLW	PBW	BCW	SWCB	ECB
55	+	+	+	+	+	+	+		+	+	+	+	+		+	+
57	+	+	+	+	+	+	+			+		+	+		+	
59	+	+	+	+	+	+		+	+						+	
61	+	+		+		+	+		+	+	+	+	+		+	+
79	+	+		+	+	+	+			+	+	+	+	+	+	+
81	+	+				+	+									
83	+	+				+		+		+			+		+	
85	+	+				+		+				+				
87	+	+				+		+		+		+	+		+	
89	+	+	+		+	+	+	+	+						+	
91	+	+				+			+			+			+	
93	+	+	+		+	+										
95	+	+				+	+	+							+	
97	+	+				+	+						+		+	
99	+	+	+		+	+	+	+	+	+	+	+		+	+	
101	+					+	+	+	+	+		+		+	+	
103	+	+			+	+	+	+	+	+		+		+	+	
105	+	+				+		+							+	
107	+	+		+	+	+	+		+	+		+	+			
109	+	+		+	+	+	+		+	+		+	+			
111	+	+			+	+		+	+			+	+		+	

Như có thể thấy trong Bảng 12 trên đây, tất cả các protein khảm diệt côn trùng thể hiện hoạt tính đối với nhiều loài bọ cánh vảy.

Ví dụ 10. Các biến thể axit amin của TIC1099 thể hiện hoạt tính đối với bọ cánh vảy

Trình tự mã hóa cho TIC1099 (SEQ ID NO: 60) được cải biến theo các phương pháp đã biết trong lĩnh vực kỹ thuật này để làm thay đổi các axit amin đặc hiệu từ trình tự protein TIC1099 gốc (SEQ ID NO: 61). Các biến thể thu được được thử nghiệm về hoạt tính đối với bọ cánh vảy chọn lọc. Bảng 13 dưới đây thể hiện các biến thể của TIC1099 và các SEQ ID NO của ADN và protein tương ứng.

Bảng 13. Các biến thể của TIC1099

Biến thể của TIC1099	ADN SEQ ID NO:	PRT SEQ ID NO:
TIC1099-T507E	62	63
TIC1099-R522K	64	65
TIC1099-K490S	66	67
TIC1099-T562R	68	69
TIC1099-S553R	70	71
TIC1099-G498D	72	73
TIC1099-K490A	74	75
TIC1099-E564A	76	77

Mỗi biến thể được thử nghiệm về tỷ lệ gây chết và còi cọc đối với sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu xám (BCW, *Agrotis ipsilon*), sâu đo đậu tương (SBL, *Chrysodeixis includens*) và sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*). Bảng 14 dưới đây thể hiện hoạt tính đối với từng sinh vật gây hại bộ cánh vảy. Hoạt tính được xếp loại '+' đến '++++' dựa vào phần trăm gây chết.

Bảng 14. Hoạt tính đối với bọ cánh vảy của các biến thể của TIC1099.

Biến thể của TIC1099	PRT SEQ ID NO:	FAW	CEW	BCW	SBL	SWCB
TIC1099	61	++++	+++	++++	++++	+++
TIC1099-T507E	63	++++	++++	++++	++++	++++

Biến thể của TIC1099	PRT SEQ ID NO:	FAW	CEW	BCW	SBL	SWCB
TIC1099-R522K	65	++++	++++	++++	+++	++++
TIC1099-K490S	67	++++	+++	++++	++	+
TIC1099-T562R	69	++++	++++	++++	++++	++++
TIC1099-S553R	71	++++	++++	++++	++++	++++
TIC1099-G498D	73	++++	++++	++++	++++	++++
TIC1099-K490A	75	++++	++++	++++	+++	+++
TIC1099-E564A	77	+++	++	+++	+	+

Tất cả các biến thể của TIC1099 đều thể hiện hoạt tính đối với năm loài bọ cánh vảy được thử nghiệm. Một vài biến thể thể hiện hoạt tính thấp hơn so với các biến thể khác đối với các sinh vật gây hại cụ thể.

Ví dụ 11. Các protein khảm diệt côn trùng được biểu hiện ở cây ngô đã được chuyển nhiễm một cách ổn định thể hiện hoạt tính đối với sinh vật gây hại bộ cánh vảy

Các trình tự mã hóa tổng hợp được xây dựng để dùng nhằm biểu hiện protein đã được mã hóa ở thực vật, được tách dòng vào vectơ biến nạp lưỡng phân thực vật, và được sử dụng để biến nạp các tế bào thực vật của cây ngô. Các vectơ biến nạp thực vật bao gồm catxet gen chuyển thứ nhất để biểu hiện protein khảm diệt côn trùng mà bao gồm trình tự khởi đầu cơ định, được liên kết hoạt động ở đầu 5' với gen dẫn đầu, được liên kết hoạt động ở đầu 5' với intron, được liên kết hoạt động ở đầu 5' với trình tự mã hóa protein khảm diệt côn trùng, mà về phần mình được liên kết hoạt động ở đầu 5' với đầu 3' UTR và; catxet gen chuyển thứ hai để chọn lọc các tế bào thực vật đã biến nạp bằng cách áp dụng chọn lọc theo glyphosat. Bảng 15 dưới đây thể hiện các trình tự mã hóa protein khảm diệt côn trùng được dùng để biểu hiện ở cây ngô và các SEQ ID NO tương ứng.

Bảng 15. Các trình tự mã hóa protein diệt côn trùng và các SEQ ID NO tương ứng.

Protein diệt côn trùng	ADN SEQ ID NO:	PRT SEQ ID NO:
TIC713	112	55
TIC713	113	55
TIC843	114	57
TIC862	115	59
TIC1099	116	61
TIC1103	117	79
TIC845	118	83
TIC846	119	85
TIC858	120	87
TIC866	121	91
TIC838	122	93
TIC841	123	97
TIC842	124	99
TIC850	125	101
TIC859	126	103
TIC861	127	105
TIC848	128	107
TIC849	129	109
TIC847	130	111

Giống ngô LH244 được chuyển nhiễm với các vectơ biến nạp lưỡng phân nêu trên bằng cách áp dụng phương pháp biến nạp gián tiếp qua *Agrobacterium*. Tế bào đã biến nạp được kích thích để tạo ra cây theo các phương pháp đã biết trong lĩnh vực kỹ thuật này. Các thử nghiệm sinh học bằng cách sử dụng phiến lá thực vật được thực hiện tương tự các thử nghiệm đã được bộc lộ trong US 8,344,207. Cây LH244 không biến nạp được dùng để thu mô nhằm dùng làm đối chứng âm tính. Nhiều thể biến nạp từ từng vectơ lưỡng phân được đánh giá đối với sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu xám (BCW, *Agrotis ipsilon*) và sâu đục thân ngô tây nam (SWCB, *Diatraea grandiosella*). Các kết

quả thử nghiệm này được thể hiện trong Bảng 16 dưới đây, trong đó '+' biểu thị hoạt tính.

Bảng 16. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô lá của cây ngô được chuyển nhiễm một cách ổn định.

Protein diệt côn trùng	CEW	FAW	BCW	SWCB
TIC713	+	+		+
TIC843	+	+		+
TIC862	+	+		+
TIC1099	+	+	+	+
TIC1103	+	+		+
TIC845	+	+		+
TIC846	+	+		+
TIC858	+	+		+
TIC866	+	+		+
TIC838		+		+
TIC841	+	+		+
TIC842	+	+	+	+
TIC850	+	+		+
TIC859	+	+		+
TIC861	+	+		+
TIC848	+	+		+
TIC849	+	+		+
TIC847		+		+

Như có thể thấy trong Bảng 16 trên đây, tất cả các protein khảm diệt côn trùng thể hiện hoạt tính đối với hai hoặc nhiều loài bọ cánh vảy.

Ví dụ 12. Các protein khảm diệt côn trùng được biểu hiện ở cây đậu tương đã được chuyển nhiễm ổn định thể hiện hoạt tính đối với sinh vật gây hại bộ cánh vảy

Các trình tự mã hóa cho các protein khảm diệt côn trùng chọn lọc được thiết kế lại để biểu hiện ở thực vật, được tách dòng vào vectơ biến nạp lưỡng phân thực vật, và sử dụng để biến nạp các tế bào thực vật của cây đậu tương. Các vectơ biến nạp thực vật bao gồm catxet gen chuyển thứ nhất để biểu hiện protein khảm diệt côn trùng mà bao gồm trình tự khởi đầu cơ định, được liên kết hoạt động ở đầu 5' với gen dẫn đầu, được liên kết hoạt động ở đầu 5' với trình tự mã hóa protein khảm diệt côn trùng, mà về phần mình được liên kết hoạt động ở đầu 5' với đầu 3' UTR và; catxet gen chuyển thứ hai để chọn lọc các tế bào thực vật đã biến nạp bằng cách áp dụng chọn lọc bằng spectinomyxin. Bảng 17 dưới đây thể hiện các trình tự mã hóa protein diệt côn trùng được dùng để biểu hiện ở cây đậu tương và các SEQ ID NO tương ứng.

Bảng 17. Các trình tự mã hóa protein diệt côn trùng và các SEQ ID NO tương ứng.

Protein diệt côn trùng	ADN SEQ ID NO:	PRT SEQ ID NO:
TIC1103	117	79
TIC866	121	91
TIC842	124	99
TIC849	129	109

Các tế bào thực vật của cây đậu tương được chuyển nhiễm bằng cách sử dụng các vectơ biến nạp lưỡng phân nêu trên bằng cách biến nạp qua trung gian *Agrobacterium*. Các tế bào thực vật đã biến nạp thu được được kích thích để tạo ra cả cây đậu tương. Mô lá được thu gom và sử dụng trong thử nghiệm sinh học như được mô tả trong Ví dụ 11 hoặc theo cách khác, mô đã đông khô nhanh được sử dụng trong chế độ dinh dưỡng cho côn trùng để thử nghiệm sinh học. Thử nghiệm sinh học được thực hiện đối với sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu vỏ đậu tương (SPW, *Helicoverpazea*), sâu bướm hại đậu mèo (VBC, *Anticarsiagemmatalis*), sâu đo đậu tương (SBL, *Chrysodeixis includens*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu đen ăn lá (BLAW, *Spodoptera cosmioides*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu đục thân ngô nhỏ (LSCB, *Elasmopalpus lignosellus*) và sâu đục quả cựu thế giới (OBW, *Helicoverpa armigera*).

Bảng 18 dưới đây thể hiện hoạt tính đối với các loài bọ cánh vảy chọn lọc của từng protein diệt côn trùng ở các cây thế hệ R_0.

Bảng 18. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô lá của cây đậu tương thế hệ R_0 đã được chuyển nhiễm một cách ổn định.

Protein diệt côn trùng	FAW	SAW	SBL	SPW	VBC	TBW	BLAW	SAW	LSCB	OBW
TIC1103		+	+	+	+	+			+	+
TIC866	+		+		+				+	+
TIC842	+	+	+		+	+	+	+	+	+
TIC849		+	+	+	+	+			+	

Như có thể thấy trong Bảng 18, mỗi protein trong số các protein khảm diệt côn trùng đã được biểu hiện ở cây đậu tương đã được chuyển nhiễm một cách ổn định thể hiện hoạt tính đối với nhiều loài bọ cánh vảy.

Các thể biến nạp chọn lọc được để tự thụ phấn và hạt thu được được gieo. Mô lá được thu gom từ các cây thế hệ R1 và được sử dụng trong thử nghiệm sinh học cho ăn. Bảng 19 thể hiện hoạt tính quan sát được đối với các loài bọ cánh vảy chọn lọc của các protein diệt côn trùng được biểu hiện ở mô lá của cây đậu tương thế hệ R1.

Bảng 19. Hoạt tính theo thử nghiệm sinh học của các protein khảm diệt côn trùng từ mô lá của cây đậu tương thế hệ R1 đã được chuyển nhiễm một cách ổn định.

Độc tố	SAW	SBL	SPW	VBC
TIC1103		+	+	+
TIC866	+	+		+
TIC842		+	+	
TIC849		+	+	

Như có thể thấy trong Bảng 19 nêu trên, các protein khảm diệt côn trùng đã được biểu hiện từ các cây thế hệ R1 thể hiện hoạt tính đối với hai hoặc nhiều loài bọ cánh vảy.

Tất cả các chế phẩm đã được bộc lộ và được yêu cầu bảo hộ theo đơn này có thể được tạo ra và được thực hiện không cần thử nghiệm quá mức dựa vào bản mô tả này. Trong khi các chế phẩm theo sáng chế đã được bộc lộ theo các phương án minh họa nêu trên, người có hiểu biết trung bình về lĩnh vực kỹ thuật này biết rõ rằng các biến đổi, thay đổi, cải biến và sửa đổi có thể được áp dụng cho chế phẩm được bộc lộ trong bản mô tả này, mà không vượt quá ý tưởng, và phạm vi thực của sáng chế. Cụ thể hơn, rõ ràng rằng một số tác nhân nhất định mà liên quan cả về mặt hóa học và về mặt sinh lý học có thể được thay thế cho các tác nhân được bộc lộ trong bản mô tả này trong khi vẫn đạt được cùng một kết quả hoặc các kết quả tương tự. Tất cả các thay thế và các cải biến dễ nhận thấy đối với người có hiểu biết trung bình về lĩnh vực kỹ thuật này được xem là nằm trong phạm vi, tinh thần và ý tưởng của sáng chế như được xác định bởi các điểm yêu cầu bảo hộ kèm theo.

Tất cả các tài liệu công bố và các tài liệu sáng chế đã công bố trích dẫn trong bản mô tả này được đưa vào đây bằng cách viện dẫn ở cùng một mức độ như từng tài liệu công bố hoặc từng đơn yêu cầu cấp bằng độc quyền sáng chế được chỉ rõ và được đưa độc lập vào đây bằng cách viện dẫn.

YÊU CẦU BẢO HỘ

1. Protein khảm diệt côn trùng chứa SEQ ID NO: 28, trong đó protein khảm diệt côn trùng này thể hiện hoạt tính ức chế đối với loài côn trùng thuộc bộ cánh vảy.

2. Protein khảm diệt côn trùng theo điểm 1, trong đó loài côn trùng được chọn từ nhóm bao gồm: *Anticarsia gemmatalis*, *Diatraea saccharalis*, *Elasmopalpus lignosellus*, *Helicoverpa zea*, *Heliothis virescens*, *Chrysodeixis includens*, *Spodoptera cosmioides*, *Spodoptera eridania*, *Spodoptera frugiperda*, *Spodoptera exigua*, *Spodoptera litura*, *Pectinophora gossypiella*, *Diatraea grandiosella*, *Earias vitella*, *Helicoverpa gelotopeon* và *Rachiplusia nu*.

3. Polynucleotit mã hóa protein khảm diệt côn trùng theo điểm 1, trong đó polynucleotit này được liên kết hoạt động với trình tự khởi đầu khác loại.

4. Polynucleotit mã hóa protein khảm diệt côn trùng, trong đó polynucleotit này chứa trình tự nucleotit:

a. nêu trong SEQ ID NO: 27; hoặc

b. mã hóa protein khảm diệt côn trùng theo điểm 1.

5. Tế bào chủ chứa polynucleotit theo điểm 4, trong đó polynucleotit đã nêu chứa trình tự có độ dài đầy đủ như nêu trong SEQ ID NO: 27, trong đó tế bào chủ này được chọn từ nhóm bao gồm tế bào chủ vi khuẩn và tế bào chủ thực vật.

6. Tế bào chủ theo điểm 5, trong đó tế bào chủ vi khuẩn được chọn từ nhóm bao gồm *Agrobacterium, Rhizobium*, *Bacillus*, *Brevibacillus*, *Escherichia*, *Pseudomonas*, *Klebsiella*, và *Erwinia.*

7. Tế bào chủ theo điểm 5, trong đó tế bào chủ thực vật được chọn từ nhóm thực vật bao gồm thực vật một lá mầm và thực vật hai lá mầm.

8. Chế phẩm ức chế côn trùng chứa protein khảm diệt côn trùng theo điểm 1.

9. Chế phẩm ức chế côn trùng theo điểm 8, trong đó chế phẩm này còn chứa ít nhất một tác nhân ức chế côn trùng khác với protein khảm diệt côn trùng.

10. Chế phẩm ức chế côn trùng theo điểm 9, trong đó ít nhất một tác nhân ức chế côn trùng được chọn từ nhóm bao gồm protein ức chế côn trùng, và phân tử dsARN ức chế côn trùng.

11. Chế phẩm ức chế côn trùng theo điểm 9, trong đó ít nhất một tác nhân diệt sinh vật gây hại khác thể hiện hoạt tính chống lại một hoặc nhiều loài sinh vật gây hại thuộc bộ cánh vảy, bộ cánh cứng, bộ cánh nửa, bộ cánh giống, hoặc bộ cánh tơ.

12. Hạt chứa lượng có tác dụng ức chế côn trùng của:

a. protein khảm diệt côn trùng theo điểm 1; hoặc

b. polynucleotit nêu trong SEQ ID NO: 27.

13. Phương pháp phòng trừ sinh vật gây hại bộ cánh vảy, phương pháp này bao gồm bước cho sinh vật gây hại bộ cánh vảy tiếp xúc với protein khảm diệt côn trùng theo điểm 1 với lượng có tác dụng ức chế.

14. Tế bào thực vật chuyển gen, cây chuyển gen hoặc bộ phận của cây chuyển gen chứa protein khảm diệt côn trùng, trong đó protein khảm diệt côn trùng này chứa protein nêu trong SEQ ID NO: 28.

15. Phương pháp phòng trừ sinh vật gây hại bộ cánh vảy, phương pháp này bao gồm việc cho sinh vật gây hại tiếp xúc với cây chuyển gen hoặc bộ phận của cây chuyển gen theo điểm 14, trong đó cây hoặc bộ phận của cây này biểu hiện protein khảm diệt côn trùng với lượng có tác dụng ức chế bọ cánh vảy.

16. Sản phẩm hàng hóa thu được từ cây chuyển gen hoặc bộ phận của cây chuyển gen theo điểm 14, trong đó sản phẩm này chứa lượng phát hiện được của protein khảm diệt côn trùng.

17. Sản phẩm hàng hóa theo điểm 16, trong đó sản phẩm được chọn từ nhóm bao gồm sinh khối thực vật, dầu, bột xay thô, thức ăn chăn nuôi, bột, cốm, cám, thớ, vỏ, và hạt đã qua xử lý.

18. Phương pháp tạo ra hạt chứa protein khảm diệt côn trùng theo điểm 1, phương pháp này bao gồm bước:

a. gieo ít nhất một hạt chứa protein khảm diệt côn trùng theo điểm 1;

b. trồng cây từ hạt này; và

c. thu hoạch hạt từ cây, trong đó hạt thu được chứa protein khảm diệt côn trùng theo điểm 1.

19. Phân tử polynucleotit tái tổ hợp mã hóa protein khảm diệt côn trùng theo điểm 1, chứa trình tự nucleotit có độ dài đầy đủ nêu trong SEQ ID NO: 27; và trình tự polynucleotit mã hóa tác nhân ức chế côn trùng khác với protein khảm diệt côn trùng.

20. Phân tử axit nucleic tái tổ hợp chứa trình tự khởi đầu khác loại được liên kết hoạt động với đoạn polynucleotit mã hóa protein khảm diệt côn trùng, trong đó:

a. protein khảm diệt côn trùng này chứa SEQ ID NO: 28; hoặc

b. đoạn polynucleotit chứa SEQ ID NO: 27;

trong đó protein khảm diệt côn trùng thể hiện hoạt tính ức chế chống lại loài côn trùng thuộc bộ cánh vảy.

DANH MỤC TRÌNH TỰ

<110> Monsanto Technology LLC
Baum, James A
Cerruti, Thomas A
Dart, Crystal L
English, Leigh H
Flasinski, Stanislaw
Fu, Xiaoran
Guzov, Victor M
Howe, Arlene R
Morgenstern, Jay P
Roberts, James K
Salvador, Sara A
Wang, Jinling

<120> Protein khảm diệt côn trùng, polynucleotit mã hóa protein khảm, chế phẩm ức chế côn trùng chứa protein khảm và phương pháp phòng trừ sinh vật gây hại bộ cánh vảy

<130> P34230WO00/0022270.00098

<150> US 62/064989
<151> 2014-10-16

<160> 130

<170> PatentIn phiên bản 3.5

<210> 1
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC1100.

<400> 1
atggagatag tgaataatca gaatcaatgc gtgccttata attgtttgaa taatcccgaa 60

atcgaaatat tagaaggcgg aagaatatca gttggtaata ccccaattga tatttctctt 120

tcgcttactc agtttctttt gagtgaattt gtcccaggtg cggggtttgt attaggatta 180

attgatttaa tatggggatt tgtaggtcct tcccaatggg acgcatttct tgctcaagtg 240

gaacagttaa ttaaccaaag aatagcagaa gctgtaagaa atacagcaat tcaggaatta 300

gagggaatgg cacgggttta tagaacctat gctactgctt ttgctgagtg ggaaaaagct 360

cctgatgacc cagagctaag agaagcacta cgtacacaat ttacagcaac tgagacttat 420

ataagtggaa gaatatccgt tttaaaaatt caaacttttg aagtacagct gttatcagtg 480

tttgcccaag ctgcaaattt acatttatct ttattaagag acgttgtgtt ttttgggcaa 540

agatggggtt tttcaacgac aaccgtaaat aattactaca atgatttaac agaagggatt 600

agtacctata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attaatcaaa tacctttagt gaaaggattt 1440

agagtttggg ggggcacctc tgtcattaca ggaccaggat ttacaggagg ggatatcctt 1500

cgaagaaata cctttggtga ttttgtatct ctacaagtca atattaattc accaattacc 1560

caaagatacc gtttaagatt tcgttacgct tccagtaggg atgcacgagt tatagtatta 1620

acaggagcgg catccacagg agtgggaggc caagttagtg taaatatgcc tcttcagaaa 1680

actatggaaa taggggagaa cttaacatct agaacattta gatataccga ttttagtaat 1740

ccttttttcat ttagagctaa tccagatata attgggataa gtgaacaacc tctatttggt 1800

gcaggttcta ttagtagcgg tgaactttat atagataaaa ttgaaattat tctagcagat 1860

gcaacatttg aagcagaatc tgatttagaa agagcgcaga aggcggtgaa tgcgctgttt 1920

acgtctacaa accaactagg gctaaaaaca aatgtaacgg attatcatat tgatcaagtg 1980

tccaatttag ttacgtattt atcggatgaa ttttgtctgg atgaaaagcg agaattgtcc 2040

gagaaagtca aacatgcgaa gcgactcagt gatgaacgca atttactcca agattcaaat 2100

ttcaaagaca ttaataggca accagaacgt gggtggggcg gaagtacagg gattaccatc 2160

caaggagggg atgacgtatt taaagaaaat tacgtcacac tatcaggtac ctttgatgag 2220

tgctatccaa catatttgta tcaaaaaatc gatgaatcaa aattaaaagc ctttacccgt 2280

tatcaattaa gagggtatat cgaagatagt caagacttag aaatctattt aattcgctac 2340

aatgcaaaac atgaaacagt aaatgtgcca ggtacgggtt ccttatggcc gctttcagcc 2400

caaagtccaa tcggaaagtg tggagagccg aatcgatgcg cgccacacct tgaatggaat 2460

cctgacttag attgttcgtg tagggatgga gaaaagtgtg cccatcattc gcatcatttc 2520

tccttagaca ttgatgtagg atgtacagac ttaaatgagg acctaggtgt atgggtgatc 2580

tttaagatta agacgcaaga tgggcacgca agactaggga atctagagtt tctcgaagag 2640

aaaccattag taggagaagc gctagctcgt gtgaaaagag cggagaaaaa atggagagac 2700

aaacgtgaaa aattggaatg ggaaacaaat atcgtttata aagaggcaaa agaatctgta 2760

gatgctttat ttgtaaactc tcaatatgat caattacaag cggatacgaa tattgccatg 2820

attcatgcgg cagataaacg tgttcatagc attcgagaag cttatctgcc tgagctgtct 2880

gtgattccgg gtgtcaatgc ggctattttt gaagaattag aagggcgtat tttcactgca 2940

ttctccctat atgatgcgag aaatgtcatt aaaaatggtg attttaataa tggcttatcc 3000

tgctggaacg tgaaagggca tgtagatgta gaagaacaaa acaaccaacg ttcggtcctt 3060

gttgttccgg aatgggaagc agaagtgtca caagaagttc gtgtctgtcc gggtcgtggc 3120

tatatccttc gtgtcacagc gtacaaggag ggatatggag aaggttgcgt aaccattcat 3180

gagatcgaga acaatacaga cgaactgaag tttagcaact gcgtagaaga ggaaatctat 3240

ccaaataaca cggtaacgtg taatgattat actgtaaatc aagaagaata cggaggtgcg 3300

tacacttctc gtaatcgagg atataacgaa gctccttccg taccagctga ttatgcgtca 3360

gtctatgaag aaaaatcgta tacagatgga cgaagagaga atccttgtga atttaacaga 3420

gggtataggg attacacgcc actaccagtt ggttatgtga caaaagaatt agaatacttc 3480

ccagaaaccg ataaggtatg gattgagatt ggagaaacgg aaggaacatt tatcgtggac 3540

agcgtggaat tactccttat ggaggaatga 3570

<210> 2
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC1100.

<400> 2
atggagattg tgaacaacca gaaccagtgc gttccttaca actgcttgaa caaccctgag 60

attgagattc ttgagggtgg tagaatttct gttggcaaca ctcctattga catctctttg 120

agtttgactc aattcttgtt gagtgagttc gttcctggtg ctggtttcgt cttgggtttg 180

attgatttga tttggggttt cgttggtcct agtcaatggg atgctttctt ggctcaagtt 240

gagcaattga ttaaccagag gatcgctgag gctgtgagga acactgctat tcaagagttg 300

gagggtatgg ctagagttta cagaacttac gctactgctt tcgctgagtg ggagaaggct 360

cctgatgacc ctgagttgag ggaggctttg agaactcaat tcactgctac tgagacttac 420

atcagtggta gaatcagtgt cttgaagatt caaactttcg aggttcaatt gctttctgtg 480

ttcgctcaag ctgcaaactt gcacttgtct ttgcttagag atgttgtgtt ctttggtcaa 540

agatggggtt tctccactac taccgtgaac aattactaca acgatttgac tgagggtatt 600

tctacttaca ctgattacgc tgttagatgg tacaacactg gtttggagag agtttggggt 660

ccagattcca gagattgggt cagatacaac cagttcagaa gggagttgac tttgactgtc 720

ttggacattg ttgctctctt ccctaactac gatagtcgtc gttaccctat tagaactgtt 780

tctcaactta ctagggaaat ctacactaac cctgttcttg agaacttcga tggtagtttc 840

cgtggtagtg ctcaagggat tgagcgttct attcgttctc ctcatcttat ggacattctt 900

aactctatta ctatctacac tgatgctcat cgtggttact attactggtc tggtcatcaa 960

attatggcta gtcctgttgg tttcagtggt cctgagttca ctttccctct ttacggtact 1020

atgggcaacg ctgcacctca acagaggatc gttgctcaac ttggtcaagg tgtttacagg 1080

actctttctt caacccttta caggcgtcct ttcaacattg ggatcaacaa ccagcagctt 1140

tctgttcttg atggaaccga gttcgcttac ggaacctctt caaaccttcc tagtgctgtt 1200

tacaggaagt ctggaaccgt tgacagtctt gatgagattc caccgcagaa caataacgtt 1260

ccacccaggc aaggcttcag tcataggctt tctcatgttt ctatgttccg ctctggattc 1320

agcaactctt cagtttctat tatcagggct ccaatgttct cgtggattca taggtctgcc 1380

gagttcaaca acattatcgc ttccgatagc attaaccaga ttccacttgt taagggattc 1440

cgtgtttggg gaggcacctc tgttattacc ggaccaggct tcaccggagg cgacattctt 1500

cgtcgtaaca ccttcggaga tttcgtttca cttcaagtga acattaactc accaatcacc 1560

cagcgctaca ggcttcgctt ccgctacgca tcatccaggg atgcaagggt gatcgtgctt 1620

accggagcag cctcaaccgg agtgggaggc caagtgagcg tgaacatgcc acttcagaag 1680

acgatggaga tcggcgagaa ccttacctca agaacctttc gttacaccga tttcagcaac 1740

ccattcagct ttcgtgcaaa cccagacatc atagggatct cagagcagcc actgtttgga 1800

gctggatcaa tctcatccgg agagctttac atcgacaaga tcgagatcat actcgcagat 1860

gcaaccttcg aggctgagag cgatctggag cgtgcacaga aggcagtgaa cgcactcttt 1920

acctctacca accagctcgg actcaagacc aacgtgaccg attaccacat cgaccaagtg 1980

agcaacctcg tgacctacct ctcagatgag ttctgcttgg atgagaaacg cgaactcagc 2040

gagaaggtga agcacgcaaa gcgtctctca gatgagcgta acctcctcca ggatagcaat 2100

ttcaaggaca tcaatcgtca gccagagcgt ggatggggag gctcaaccgg aatcaccatc 2160

cagggaggcg atgatgtgtt taaggagaat tacgtgacac tctccggaac attcgatgag 2220

tgctacccaa catacctcta tcagaagatc gacgagtcca agctcaaggc gttcacccgt 2280

tatcagctcc gtggctacat cgaggatagt caagacctgg aaatctacct catccgctac 2340

aatgcaaagc acgagacagt gaatgtgcca ggaacaggct ccctctggcc actctccgca 2400

cagtctccaa tcggcaagtg cggcgagcca aatcgctgcg cgccacacct ggagtggaat 2460

cccgacctgg actgctcctg ccgcgacggc gagaagtgcg cccaccactc ccaccacttt 2520

agcctggaca tcgacgtggg ctgtacagac ctgaatgagg atctgggcgt gtgggtgatc 2580

tttaagatca agacacagga cggccacgcc cgcctgggca atctggagtt tctggaggag 2640

aagcctctgg tgggcgaagc cctggcccgc gtgaagcgcg ccgagaagaa atggcgcgac 2700

aaacgcgaga aactggaatg ggaaacaaac atcgtgtaca aagaagccaa agaatccgtg 2760

gacgccctat ttgtgaactc ccagtatgac cagctacagg ccgacacaaa catcgcgatg 2820

atccacgctg cggacaagcg cgtgcactcc atacgcgaag cctatctacc cgaactatcc 2880

gtgatacccg gcgtcaatgc cgcgatcttt gaagaattgg aaggccgcat cttcacagcc 2940

tttagcctct atgacgcccg aaatgtcatc aagaatggcg actttaacaa tgggctatcc 3000

tgttggaatg tcaaagggca cgtggacgtc gaagagcaga acaatcagcg atccgtctta 3060

gtcgtacccg aatgggaagc cgaagtctcc caggaagtcc gagtctgtcc tggtagaggt 3120

tacatcttga gagtgactgc ttacaaggag ggttacggtg agggatgcgt gactattcac 3180

gagattgaga acaacactga tgagttgaag ttcagtaact gcgtggagga ggaaatctac 3240

cccaacaaca ctgtgacttg taacgattac accgtgaacc aggaggaata cggaggcgct 3300

tacacctcca gaaaccgtgg atacaatgag gctccctcgg tccccgctga ttatgcctcc 3360

gtctatgagg agaagtccta caccgatgga aggcgcgaga atccctgcga gttcaatcgc 3420

ggctatcgag actacactcc gctacccgtt ggctatgtca caaaggaact ggaatacttc 3480

ccggaaacag acaaagtctg gatcgaaatc ggcgaaacag aagggacgtt catagtcgat 3540

agcgtagaac ttctccttat ggaagaatga 3570

<210> 3
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC1100.

<400> 3
atggagattg tgaacaacca gaaccagtgc gttccttaca actgcttgaa caaccctgag 60

attgagattc ttgagggtgg tagaatttct gttggcaaca ctcctattga catctctttg 120

agtttgactc aattcttgtt gagtgagttc gttcctggtg ctggtttcgt cttgggtttg 180

attgatttga tttggggttt cgttggtcct agtcaatggg atgctttctt ggctcaagtt 240

gagcaattga ttaaccagag gatcgctgag gctgtgagga acactgctat tcaagagttg 300

gagggtatgg ctagagttta cagaacttac gctactgctt tcgctgagtg ggagaaggct 360

cctgatgacc ctgagttgag ggaggctttg agaactcaat tcactgctac tgagacttac 420

atcagtggta gaatcagtgt cttgaagatt caaactttcg aggttcaatt gctttctgtg 480

ttcgctcaag ctgcaaactt gcacttgtct ttgcttagag atgttgtgtt ctttggtcaa 540

agatggggtt tctccactac taccgtgaac aattactaca acgatttgac tgagggtatt 600

tctacttaca ctgattacgc tgttagatgg tacaacactg gtttggagag agtttggggt 660

ccagattcca gagattgggt cagatacaac cagttcagaa gggagttgac tttgactgtc 720

ttggacattg ttgctctctt ccctaactac gatagtcgtc gttaccctat tagaactgtt 780

tctcaactta ctagggaaat ctacactaac cctgttcttg agaacttcga tggtagtttc 840

cgtggtagtg ctcaagggat tgagcgttct attcgttctc ctcatcttat ggacattctt 900

aactctatta ctatctacac tgatgctcat cgtggttact attactggtc tggtcatcaa 960

attatggcta gtcctgttgg tttcagtggt cctgagttca ctttccctct ttacggtact 1020

atgggcaacg ctgcacctca acagaggatc gttgctcaac ttggtcaagg tgtttacagg 1080

actctttctt caacccttta caggcgtcct ttcaacattg ggatcaacaa ccagcagctt 1140

tctgttcttg atggaaccga gttcgcttac ggaacctctt caaaccttcc tagtgctgtt 1200

tacaggaagt ctggaaccgt tgacagtctt gatgagattc caccgcagaa caataacgtt 1260

ccacccaggc aaggcttcag tcataggctt tctcatgttt ctatgttccg ctctggattc 1320

agcaactctt cagtttctat tatcagggct ccaatgttct cgtggattca taggtctgcc 1380

gagttcaaca acattatcgc ttccgatagc attaaccaga ttccacttgt taagggattc 1440

cgtgtttggg gaggcacctc tgttattacc ggaccaggct tcaccggagg cgacattctt 1500

cgtcgtaaca ccttcggaga tttcgtttca cttcaagtga acattaactc accaatcacc 1560

cagcgctaca ggcttcgctt ccgctacgca tcatccaggg atgcaagggt gatcgtgctt 1620

accggagcag cctcaaccgg agtgggaggc caagtgagcg tgaacatgcc acttcagaag 1680

acgatggaga tcggcgagaa ccttacctca agaacctttc gttacaccga tttcagcaac 1740

ccattcagct ttcgtgcaaa cccagacatc atagggatct cagagcagcc actgtttgga 1800

gctggatcaa tctcatccgg agagctttac atcgacaaga tcgagatcat actcgcagat 1860

gcaaccttcg aggctgagag cgatctggag cgtgcacaga aggcagtgaa cgcactcttt 1920

acctctacca accagctcgg actcaagacc aacgtgaccg attaccacat cgaccaagtg 1980

agcaacctcg tgacctacct ctcagatgag ttctgcttgg atgagaaacg cgaactcagc 2040

gagaaggtga agcacgcaaa gcgtctctca gatgagcgta acctcctcca ggatagcaat 2100

ttcaaggaca tcaatcgtca gccagagcgt ggatggggag gctcaaccgg aatcaccatc 2160

cagggaggcg atgatgtgtt taaggagaat tacgtgacac tctccggaac attcgatgag 2220

tgctacccaa catacctcta tcagaagatc gacgagtcca agctcaaggc gttcacccgt 2280

tatcagctcc gtggctacat cgaggatagt caagacctgg aaatctacct catccgctac 2340

aatgcaaagc acgagacagt gaatgtacca ggaacaggct ccctctggcc actctccgca 2400

cagtctccaa tcggcaagtg cggcgagcca aatcgctgcg cgccacacct ggagtggaat 2460

cccgacctgg actgctcctg ccgcgacggc gagaagtgcg cccaccactc ccaccacttt 2520

agcctggaca tcgacgtggg ctgtacagac ctgaatgagg atctgggcgt gtgggtgatc 2580

tttaagatca agacacagga cggccacgcc cgcctgggca atctggagtt tctggaggag 2640

aagcctctgg tgggcgaagc cctggcccgc gtgaagcgcg ccgagaagaa atggcgcgac 2700

aaacgcgaga aactggaatg ggaaacaaac atcgtgtaca aagaagccaa agaatccgtg 2760

gacgccctat ttgtgaactc ccagtatgac cagctacagg ccgacacaaa catcgcgatg 2820

atccacgctg cggacaagcg cgtgcactcc atacgcgaag cctatctacc cgaactatcc 2880

gtgatacccg gcgtcaatgc cgcgatcttt gaagaattgg aaggccgcat cttcacagcc 2940

tttagcctct atgacgcccg aaatgtcatc aagaatggcg actttaacaa tgggctatcc 3000

tgttggaatg tcaaagggca cgtggacgtc gaagagcaga acaatcagcg atccgtctta 3060

gtcgtacccg aatgggaagc cgaagtctcc caggaagtcc gagtctgtcc tggtagaggt 3120

tacatcttga gagtgactgc ttacaaggag ggttacggtg agggatgcgt gactattcac 3180

gagattgaga acaacactga tgagttgaag ttcagtaact gcgtggagga ggaaatctac 3240

cccaacaaca ctgtgacttg taacgattac accgtgaacc aggaggaata cggaggcgct 3300

tacacctcca gaaaccgtgg atacaatgag gctccctcgg tccccgctga ttatgcctcc 3360

gtctatgagg agaagtccta caccgatgga aggcgcgaga atccctgcga gttcaatcgc 3420

ggctatcgag actacactcc gctacccgtt ggctatgtca caaaggaact ggaatacttc 3480

ccggaaacag acaaagtctg gatcgaaatc ggcgaaacag aagggacgtt catagtcgat 3540

agcgtagaac ttctccttat ggaagaatga 3570

<210> 4
<211> 1189
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1100.

<400> 4

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln

```
305                 310                 315                 320

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

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

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
                485                 490                 495

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

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

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
    530                 535                 540

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
545                 550                 555                 560

Thr Met Glu Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr
                565                 570                 575

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

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
        595                 600                 605

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
    610                 615                 620

Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe
625                 630                 635                 640
```

```
Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val Thr Asp Tyr His
                645                 650                 655

Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys
            660                 665                 670

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

Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn Phe Lys Asp Ile
    690                 695                 700

Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile
705                 710                 715                 720

Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly
                725                 730                 735

Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu
            740                 745                 750

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

Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His
    770                 775                 780

Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala
785                 790                 795                 800

Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His
                805                 810                 815

Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys
            820                 825                 830

Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys
        835                 840                 845

Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys
    850                 855                 860

Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu
865                 870                 875                 880

Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys
                885                 890                 895

Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val
            900                 905                 910

Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln
        915                 920                 925

Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala
    930                 935                 940

Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser
945                 950                 955                 960
```

Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg
965 970 975

Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn
980 985 990

Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val
995 1000 1005

Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu Val Val Pro
1010 1015 1020

Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly
1025 1030 1035

Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
1040 1045 1050

Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu
1055 1060 1065

Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr Pro Asn Asn
1070 1075 1080

Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu Glu Tyr Gly
1085 1090 1095

Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu Ala Pro Ser
1100 1105 1110

Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ser Tyr Thr
1115 1120 1125

Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg Gly Tyr Arg
1130 1135 1140

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

Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr
1160 1165 1170

Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu
1175 1180 1185

Glu

<210> 5
<211> 3672
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC860.

<400> 5
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccaacggta 60

tcgaatcctt ccacgcaaat gaatctatca ccagatgctc gtattgaaga tagcttgtgt 120

gtagccgagg tgaacaatat tgatccattt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccattttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

acttcaatta atcctgtaac attacagttt acgtctcgag acgtttatag aacagaatca 1200

aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt 1260

aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat 1320

cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa 1380

cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac 1440

actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaatacgatt 1500

ggaccaaata gaattaatca aatacctta gtgaaaggat ttagagtttg gggggcacc 1560

tctgtcatta caggaccagg atttacagga ggggatatcc ttcgaagaaa tacctttggt 1620

gattttgtat ctctacaagt caatattaat tcaccaatta cccaaagata ccgtttaaga 1680

tttcgttacg cttccagtag ggatgcacga gttatagtat taacaggagc ggcatccaca 1740

ggagtgggag gccaagttag tgtaaatatg cctcttcaga aaactatgga aataggggag 1800

aacttaacat ctagaacatt tagatatacc gattttagta atccttttc atttagagct 1860

aatccagata taattgggat aagtgaacaa cctctatttg gtgcaggttc tattagtagc 1920

ggtgaacttt atatagataa aattgaaatt attctagcag atgcaacatt tgaagcagaa 1980

tctgatttag aaagagcgca gaaggcggtg aatgcgctgt ttacgtctac aaaccaacta 2040

gggctaaaaa caaatgtaac ggattatcat attgatcaag tgtccaattt agttacgtat 2100

ttatcggatg aattttgtct ggatgaaaag cgagaattgt ccgagaaagt caaacatgcg 2160

aagcgactca gtgatgaacg caatttactc caagattcaa atttcaaaga cattaatagg 2220

caaccagaac gtgggtgggg cggaagtaca gggattacca tccaaggagg ggatgacgta 2280

tttaaagaaa attacgtcac actatcaggt acctttgatg agtgctatcc aacatatttg 2340

tatcaaaaaa tcgatgaatc aaaattaaaa gcctttaccc gttatcaatt aagagggtat 2400

atcgaagata gtcaagactt agaaatctat ttaattcgct acaatgcaaa acatgaaaca 2460

gtaaatgtgc caggtacggg ttccttatgg ccgctttcag cccaaagtcc aatcggaaag 2520

tgtggagagc cgaatcgatg cgcgccacac cttgaatgga atcctgactt agattgttcg 2580

tgtagggatg gagaaaagtg tgcccatcat tcgcatcatt tctccttaga cattgatgta 2640

ggatgtacag acttaaatga ggacctaggt gtatgggtga tctttaagat taagacgcaa 2700

gatgggcacg caagactagg gaatctagag tttctcgaag agaaaccatt agtaggagaa 2760

gcgctagctc gtgtgaaaag agcggagaaa aaatggagag acaaacgtga aaaattggaa 2820

tgggaaacaa atatcgttta taaagaggca aaagaatctg tagatgcttt atttgtaaac 2880

tctcaatatg atcaattaca agcggatacg aatattgcca tgattcatgc ggcagataaa 2940

cgtgttcata gcattcgaga agcttatctg cctgagctgt ctgtgattcc gggtgtcaat 3000

gcggctattt ttgaagaatt agaagggcgt attttcactg cattctccct atatgatgcg 3060

agaaatgtca ttaaaaatgg tgattttaat aatggcttat cctgctggaa cgtgaaaggg 3120

catgtagatg tagaagaaca aaacaaccaa cgttcggtcc ttgttgttcc ggaatgggaa 3180

gcagaagtgt cacaagaagt tcgtgtctgt ccgggtcgtg gctatatcct tcgtgtcaca 3240

gcgtacaagg agggatatgg agaaggttgc gtaaccattc atgagatcga gaacaataca 3300

gacgaactga agtttagcaa ctgcgtagaa gaggaaatct atccaaataa cacggtaacg 3360

tgtaatgatt atactgtaaa tcaagaagaa tacggaggtg cgtacacttc tcgtaatcga 3420

ggatataacg aagctccttc cgtaccagct gattatgcgt cagtctatga agaaaaatcg 3480

tatacagatg gacgaagaga gaatccttgt gaatttaaca gagggtatag ggattacacg 3540

ccactaccag ttggttatgt gacaaaagaa ttagaatact tcccagaaac cgataaggta 3600

tggattgaga ttggagaaac ggaaggaaca tttatcgtgg acagcgtgga attactcctt 3660

atggaggaat ag 3672

<210> 6

<211> 3672
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC860.

<400> 6
atgaccagca accggaagaa cgagaacgag atcatcaacg ccctgagcat cccgaccgtg 60

agcaacccta gcacccagat gaacctgagc cctgacgctc gcatcgagga ctccctctgc 120

gtggctgagg tgaacaacat cgacccgttc gtgtccgcct ccaccgtgca gaccggcatc 180

aacatcgcgg gccgcatcct cggcgtgctc ggcgtgccct ttgcgggcca gctcgcctcc 240

ttctactcct tcctcgtggg agagctgtgg ccctccggcc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catccgccag caagtcaccg agaacacccg caacaccgcc 360

atcgcccgcc tggagggcct gggccgtggc taccgctcct accagcaagc cctggagacc 420

tggctcgaca accgcaacga cgcccgctcc cgctccatca tcctggagcg ctacgtcgcc 480

ctggaactgg acatcaccac tgccatccca ctcttccgca tcaggaacga ggaggtgcct 540

ctgctgatgg tgtacgccca ggctgcgaac ctgcacctgc tgctgctgcg cgacgcaagc 600

ctgtttggct ccgagtgggg tatggcaagc tccgacgtca accagtacta ccaggagcag 660

atccgctaca ccgaggagta cagcaaccac tgcgtccagt ggtacaacac cggtctgaac 720

aatctcagag ggaccaacgc tgagagctgg ctgcgctaca accagttccg gcgggatctg 780

accctaggtg tcctggatct ggtcgctctg ttcccgagct acgataccag gacgtaccct 840

atcaacacct ctgctcagct taccagggag atctacactg atcctatcgg taggactaac 900

gctcctagtg gtttcgccag cactaactgg ttcaacaaca acgcgcctag tttctctgcc 960

atcgaggcgg cgatcttccg gcctcctcac ctcctcgact tcccggagca gcttactatc 1020

tactctgcgt cttcgcggtg gtcttcgact cagcacatga actactgggt tggtcaccgg 1080

cttaacttcc gcccgattgg aggaactctt aacaccagta cgcaaggtct tacgaacaac 1140

acttccatca acccggttac gttgcagttc acgtctcggg acgtttaccg gacggagtcg 1200

aatgctggga cgaacatcct gttcacgaca ccggtgaatg gtgttccgtg ggcacgtttc 1260

aacttcatca acccgcagaa catctacgag cgtggagcaa cgacatactc gcaaccatac 1320

caaggcgttg gcatccaact gtttgactcg gagacggaac tgccaccaga gacgacagaa 1380

cgtccgaatt acgagtcata ctcacacaga ctatcacaca ttggactcat tatcggaaac 1440

acactgagag caccagtgta ctcatggaca catcggtcag cagatcgtac gaacaccatc 1500

ggacccaatc ggatcaacca gatcccgctc gtgaagggct tccgcgtgtg gggcggcacc 1560

tccgtcatca ccggtccggg cttcaccggc ggcgacatcc tccgccgcaa caccttcggc 1620

gacttcgtgt cactccaagt gaacatcaac agcccgatca cccagcgcta tcgcctccgc 1680

ttccgctacg cctcctcccg cgacgctaga gtgatcgtgc tcaccggagc ggcgtccaca 1740

ggcgtaggcg gccaagtgtc tgtgaacatg ccgctccaga agactatgga gattggtgag 1800

aacctcacct ctcgcacctt ccgctacacc gacttctcca atccgttctc cttcagagcc 1860

aacccagaca tcatcggcat ctccgagcag cctctctttg gcgctggctc catctcctcc 1920

ggcgagctgt acatcgacaa gattgagatc atccttgccg acgccacctt cgaagctgag 1980

tccgatctcg agcgcgccca gaaggccgtg aacgccctct tcactagcac taaccagctc 2040

ggcctcaaga ctaacgtgac cgactaccac attgaccaag tgagcaacct agtgacctac 2100

cttagcgacg agttctgcct tgacgagaag cgtgagctga gcgagaaggt gaagcacgcc 2160

aagcgcctct ccgacgagcg caacctcctc caggactcca acttcaagga catcaaccgc 2220

cagcccgagc gcggctgggg cggtagcacc ggcatcacca tccagggcgg tgacgatgtg 2280

ttcaaggaga actacgtgac cctctccggc accttcgacg agtgctaccc gacctacctc 2340

taccagaaga tcgacgagtc caagctcaag gcgttcaccc gctaccagct tcgcggctac 2400

atcgaggact cccaggatct ggagatctac ctcatccgct acaacgccaa gcacgagacc 2460

gtgaacgtgc ccggcaccgg ctccctctgg ccgctctccg cccagagccc tatcggcaag 2520

tgcggcgagc ccaaccgctg cgcgcctcac ctggagtgga accctgacct cgactgctcc 2580

tgccgcgacg gcgagaagtg cgcccaccat agccaccact tctctctcga catcgacgtg 2640

ggctgcaccg acctcaacga ggatctgggc gtgtgggtga tcttcaagat caagacccag 2700

gacggccacg ccaggctggg caacctggag ttcctggagg agaagcctct ggtgggtgag 2760

gccctggcca gggtcaagag ggctgagaag aaatggaggg acaagaggga gaagctggag 2820

tgggagacca acatcgtgta caaggaggct aaggagtccg tggacgctct gttcgtcaac 2880

tctcagtacg atcagctcca ggctgacacc aacatcgcta tgatccacgc tgcggataag 2940

agggtccact ctatcaggga ggcttacctg cctgagcttt ctgtcatccc tggtgtcaac 3000

gcggcaatct tcgaggaact tgagggccgc atcttcactg cgttctcgct ttacgatgcg 3060

cggaacgtca ttaagaacgg tgacttcaac aatggtcttt cgtgctggaa cgtcaagggt 3120

catgtcgatg tcgaggaaca gaacaaccag cggtcggtcc ttgtcgttcc cgagtgggag 3180

gccgaggtct cgcaagaggt ccgggtctgc cctgggcgcg ggtacattct tcgtgtcact 3240

gcgtacaagg agggctacgg cgagggctgc gttactattc atgagattga gaacaatacg 3300

gatgagctta agtttagtaa ctgtgttgag gaggagatct acccgaacaa tacggttacg 3360

tgcaatgatt acacggtgaa ccaggaggaa tacggcggag catacacctc acgtaataga 3420

gggtacaatg aggcaccgtc agttccggca gattatgcct cagtttatga ggagaagtcc 3480

tacacggatg gaagacgcga gaatccatgt gagtttaata gaggataccg agactacaca 3540

ccactcccag ttggatacgt tacaaaggag ttggaatact tcccagaaac agataaagtt 3600

tggatagaga tcggagaaac agaaggaacc ttcatcgtgg acagtgtaga actgctgctg 3660

atggaagaat ga 3672

<210> 7
<211> 1223
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC860.

<400> 7

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Thr Val Ser Asn Pro Ser Thr Gln Met Asn Leu Ser Pro Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Val Ala Glu Val Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn
130 135 140

Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145 150 155 160

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

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
195 200 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr

```
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

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

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
    450                 455                 460

Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
                485                 490                 495

Thr Asn Thr Ile Gly Pro Asn Arg Ile Asn Gln Ile Pro Leu Val Lys
            500                 505                 510

Gly Phe Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe
        515                 520                 525

Thr Gly Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser
    530                 535                 540
```

```
Leu Gln Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg
545                 550                 555                 560

Phe Arg Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly
                565                 570                 575

Ala Ala Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu
            580                 585                 590

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

Tyr Thr Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile
    610                 615                 620

Ile Gly Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser
625                 630                 635                 640

Gly Glu Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr
                645                 650                 655

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

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

Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu
    690                 695                 700

Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala
705                 710                 715                 720

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

Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile
            740                 745                 750

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

Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile
    770                 775                 780

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

Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala
                805                 810                 815

Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu
            820                 825                 830

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

Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly
    850                 855                 860
```

```
Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val
865                 870                 875                 880

Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys
                885                 890                 895

Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu
            900                 905                 910

Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala
        915                 920                 925

Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn
    930                 935                 940

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

Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His
                965                 970                 975

Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu
            980                 985                 990

Leu Ser Val Ile Pro Gly Val Asn  Ala Ala Ile Phe Glu  Glu Leu Glu
        995                 1000                1005

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

Ile Lys  Asn Gly Asp Phe Asn  Asn Gly Leu Ser Cys  Trp Asn Val
    1025                1030                1035

Lys Gly  His Val Asp Val Glu  Glu Gln Asn Asn Gln  Arg Ser Val
    1040                1045                1050

Leu Val  Val Pro Glu Trp Glu  Ala Glu Val Ser Gln  Glu Val Arg
    1055                1060                1065

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

Glu Gly  Tyr Gly Glu Gly Cys  Val Thr Ile His Glu  Ile Glu Asn
    1085                1090                1095

Asn Thr  Asp Glu Leu Lys Phe  Ser Asn Cys Val Glu  Glu Glu Ile
    1100                1105                1110

Tyr Pro  Asn Asn Thr Val Thr  Cys Asn Asp Tyr Thr  Val Asn Gln
    1115                1120                1125

Glu Glu  Tyr Gly Gly Ala Tyr  Thr Ser Arg Asn Arg  Gly Tyr Asn
    1130                1135                1140

Glu Ala  Pro Ser Val Pro Ala  Asp Tyr Ala Ser Val  Tyr Glu Glu
    1145                1150                1155

Lys Ser  Tyr Thr Asp Gly Arg  Arg Glu Asn Pro Cys  Glu Phe Asn
    1160                1165                1170

Arg Gly  Tyr Arg Asp Tyr Thr  Pro Leu Pro Val Gly  Tyr Val Thr
```

1175 1180 1185

Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu
1190 1195 1200

Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu
1205 1210 1215

Leu Leu Met Glu Glu
1220

<210> 8
<211> 3564
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC867.

<400> 8
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta cacaaatacc attggtaaag gcgcataccc tccaatcggg taccactgta 1560

gtaaaagggc cagggtttac aggaggggat atcctccgtc gaacaagtgg aggaccattt 1620

gcttttagta atgttaatct agattttaac ttgtcacaaa ggtatcgtgc tagaattcgt 1680

tatgcctcta ctactaacct aagaatttac gtaacggttg caggtgaacg aatttttgct 1740

ggtcaatttg acaaaactat ggatgctggt gccccattaa cattccaatc ttttagttac 1800

gcaactatta atacagcttt tacattccca gaaagatcga gcagcttgac tgtaggtgcc 1860

gatacgttta gttcaggtaa tgaagtttat gtagatagat ttgaattaat cccagttact 1920

gcaaccttcg aggcagaatc tgatttagaa agagcacaaa aggcggtgaa tgagctgttt 1980

acttcttcca atcaaatcgg gttaaaaaca gatgtgacgg attatcatat tgatcaagta 2040

tccaatttag ttgagtgttt atctgatgaa ttttgtctgg atgaaaaaaa agaattgtcc 2100

gagaaagtca aacatgcgaa gcgacttagt gatgagcgga atttacttca agatccaaac 2160

tttagaggga tcaatagaca actagaccgt ggctggagag gaagtacgga tattaccatc 2220

caaggaggcg atgacgtatt caaagagaat tacgttacgc tattgggtac ctttgatgag 2280

tgctatccaa cgtatttata tcaaaaaata gatgagtcga aattaaaagc ctatacccgt 2340

taccaattaa gagggtatat cgaagatagt caagacttag aaatctattt aattcgctac 2400

aatgccaaac acgaaacagt aaatgtgcca ggtacgggtt ccttatggcc gctttcagcc 2460

ccaagtccaa tcggaaaatg tgcccatcat tcccatcatt tctccttgga cattgatgtt 2520

ggatgtacag acttaaatga ggacttaggt gtatgggtga tattcaagat taagacgcaa 2580

gatggccatg caagactagg aaatctagaa tttctcgaag agaaaccatt agtaggagaa 2640

gcactagctc gtgtgaaaag agcggagaaa aaatggagag acaaacgtga aaaattggaa 2700

tgggaaacaa atattgttta taaagaggca aaagaatctg tagatgcttt atttgtaaac 2760

tctcaatatg atagattaca agcggatacc aacatcgcga tgattcatgc ggcagataaa 2820

cgcgttcata gcattcgaga agcttatctg cctgagctgt ctgtgattcc gggtgtcaat 2880

gcggctattt ttgaagaatt agaagggcgt attttcactg cattctccct atatgatgcg 2940

agaaatgtca ttaaaaatgg tgattttaat aatggcttat cctgctggaa cgtgaaaggg 3000

catgtagatg tagaagaaca aaacaaccac cgttcggtcc ttgttgttcc ggaatgggaa 3060

gcagaagtgt cacaagaagt tcgtgtctgt ccgggtcgtg gctatatcct tcgtgtcaca 3120

gcgtacaagg agggatatgg agaaggttgc gtaaccattc atgagatcga gaacaataca 3180

gacgaactga agtttagcaa ctgtgtagaa gaggaagtat atccaaacaa cacggtaacg 3240

tgtaatgatt atactgcgac tcaagaagaa tatgagggta cgtacacttc tcgtaatcga 3300

ggatatgacg gagcctatga aagcaattct tctgtaccag ctgattatgc atcagcctat 3360

gaagaaaaag catatacaga tggacgaaga gacaatcctt gtgaatctaa cagaggatat 3420

ggggattaca caccactacc agctggctat gtgacaaaag aattagagta cttcccagaa 3480

accgataagg tatggattga gatcggagaa acggaaggaa cattcatcgt ggacagcgtg 3540

gaattacttc ttatggagga atag 3564

<210> 9
<211> 3564
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867.

<400> 9
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccttcg aagctgagtc ggacctggag cgtgcacaga aggcagtcaa cgagctgttc 1980

acctctagca accagatcgg cctcaagacc gacgtcacag actaccacat cgaccaagtg 2040

tccaacctgg tcgagtgcct tagcgacgag ttctgcctag acgagaagaa ggagctgtcg 2100

gagaaggtca aacacgccaa gcgtctgagc gatgagcgca acctgctcca agaccctaac 2160

ttccgtggca tcaacaggca gcttgaccgt ggctggcgcg gctcgacgga catcacgatc 2220

cagggtggcg acgacgtatt caaggagaat tacgtgacct tgcttgggac gtttgacgag 2280

tgctatccca cctacctcta ccagaagatt gatgaatcga aattgaaggc gtacacgaga 2340

taccagctcc gtggctacat cgaggacagc caggacttgg agatctacct catacgctac 2400

aacgctaaac atgagaccgt gaacgtccct gggacgggca gtctgtggcc actctctgct 2460

cctagcccta tcggcaagtg cgctcaccac tcgcaccact tcagccttga catcgacgtg 2520

ggatgtactg acctcaacga agacctgggc gtctgggtta tcttcaagat caagacccag 2580

gacggccacg cccgactcgg caacctggag ttcctggagg agaaaccact ggtgggcgag 2640

gcgctcgccc gcgtgaagcg tgccgagaag aagtggcggg acaagaggga gaagctagaa 2700

tgggagacga acatcgtgta caaggaggcc aaggaaagcg tcgatgccct gttcgtgaac 2760

tcacagtacg accgtctcca ggcggacacg aacatcgcca tgatccacgc ggctgacaag 2820

cgcgtccact ccatccgcga ggcgtactta ccggagctgt cggtgatccc aggcgtaaac 2880

gcggcgatct tcgaggagct agagggacgc atcttcacag cgttcagcct gtacgacgca 2940

cgcaacgtca tcaagaacgg cgatttcaac aacggactgt cctgctggaa cgtgaagggc 3000

cacgtcgatg tcgaggaaca gaacaaccac cgctctgtcc tggtggtccc agagtgggag 3060

gccgaggtct cccaggaggt ccgcgtgtgc cctgggcgtg gctacatcct ccgtgtgaca 3120

gcctacaagg agggctacgg tgagggctgc gtcaccattc acgagatcga gaacaacact 3180

gacgaactca agttctcgaa ttgcgtggag gaggaggtgt acccgaacaa tacggtgacg 3240

tgcaacgact acacggcaac ccaagaggag tacgagggca cctacaccag taggaaccgt 3300

ggctacgacg gtgcctacga gtcgaactcc agcgtccctg cggactacgc cagcgcgtac 3360

gaggagaagg cttacaccga cggacgccgg gacaacccat gcgagagcaa ccgtggctac 3420

ggcgactaca ctcctctccc ggccggatac gtcacaaagg agctggagta tttcccagag 3480

acggacaagg tgtggatcga aatcggagag acagagggaa ccttcatcgt ggacagcgtg 3540

gagctgctcc tcatggagga gtga 3564

<210> 10
<211> 1187
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC867.

<400> 10

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
```

```
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
                645                 650                 655

Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
            660                 665                 670

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
        675                 680                 685

Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys
    690                 695                 700

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
705                 710                 715                 720

Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr
                725                 730                 735

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
            740                 745                 750

Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
        755                 760                 765

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
    770                 775                 780
```

```
Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
785                 790                 795                 800

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
                805                 810                 815

Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His
            820                 825                 830

His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp
        835                 840                 845

Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala
    850                 855                 860

Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu
865                 870                 875                 880

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

Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu
            900                 905                 910

Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala
        915                 920                 925

Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser
    930                 935                 940

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

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

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly
            980                 985                 990

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

Asn His  Arg Ser Val Leu Val  Val Pro Glu Trp Glu  Ala Glu Val
    1010                1015                1020

Ser Gln  Glu Val Arg Val Cys  Pro Gly Arg Gly Tyr  Ile Leu Arg
    1025                1030                1035

Val Thr  Ala Tyr Lys Glu Gly  Tyr Gly Glu Gly Cys  Val Thr Ile
    1040                1045                1050

His Glu  Ile Glu Asn Asn Thr  Asp Glu Leu Lys Phe  Ser Asn Cys
    1055                1060                1065

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

Tyr Thr  Ala Thr Gln Glu Glu  Tyr Glu Gly Thr Tyr  Thr Ser Arg
    1085                1090                1095
```

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

Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly
1115 1120 1125

Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr
1130 1135 1140

Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe
1145 1150 1155

Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly
1160 1165 1170

Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180 1185

<210> 11
<211> 3642
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC867_20.

<400> 11
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta cacaaatacc attggtaaag gcgcataccc tccaatcggg taccactgta 1560

gtaaaagggc cagggtttac aggaggggat atcctccgtc gaacaagtgg aggaccattt 1620

gcttttagta atgttaatct agattttaac ttgtcacaaa ggtatcgtgc tagaattcgt 1680

tatgcctcta ctactaacct aagaatttac gtaacggttg caggtgaacg aatttttgct 1740

ggtcaatttg acaaaactat ggatgctggt gccccattaa cattccaatc ttttagttac 1800

gcaactatta atacagcttt tacattccca gaaagatcga gcagcttgac tgtaggtgcc 1860

gatacgttta gttcaggtaa tgaagtttat gtagatagat ttgaattaat cccagttact 1920

gcaacctttg aggcagaata tgatttagaa agagcgcaaa aggtggtgaa tgccctgttt 1980

acgtctacaa accaactagg gctaaaaaca gatgtgacgg attatcatat tgatcaggta 2040

tccaatctag ttgcgtgttt atcggatgaa ttttgtctgg atgaaaagag agaattgtcc 2100

gagaaagtta aacatgcaaa gcgactcagt gatgagcgga atttacttca agatccaaac 2160

ttcagaggga tcaataggca accagaccgt ggctggagag gaagtacgga tattactatc 2220

caaggaggag atgacgtatt caaagagaat tacgttacgc taccgggtac ctttgatgag 2280

tgctatccaa cgtatttata tcaaaaaata gatgagtcga aattaaaagc ctatacccgt 2340

tatcaattaa gagggtatat cgaagatagt caagacttag aaatctattt aattcgttac 2400

aatgcaaaac acgaaatagt aaatgtacca ggtacaggaa gtttatggcc tctttctgta 2460

gaaaatcaaa ttggaccttg tggagaaccg aatcgatgcg cgccacacct tgaatggaat 2520

cctgatttac actgttcctg cagagacggg gaaaaatgtg cacatcattc tcatcatttc 2580

tctttggaca ttgatgttgg atgtacagac ttaaatgagg acttaggtgt atgggtgata 2640

ttcaagatta agacgcaaga tggccacgca cgactaggga atctagagtt tctcgaagag 2700

aaaccattat taggagaagc actagctcgt gtgaaaagag cggagaaaaa atggagagac 2760

aaacgcgaaa cattacaatt ggaaacaact atcgtttata aagaggcaaa agaatctgta 2820

gatgctttat ttgtaaactc tcaatatgat agattacaag cggatacgaa catcgcgatg 2880

attcatgcgg cagataaacg cgttcataga attcgagaag cgtatctgcc ggagctgtct 2940

gtgattccgg gtgtcaatgc ggctattttt gaagaattag aagagcgtat tttcactgca 3000

ttttccctat atgatgcgag aaatattatt aaaaatggcg atttcaataa tggcttatta 3060

tgctggaacg tgaaagggca tgtagaggta gaagaacaaa acaatcaccg ttcagtcctg 3120

gttatcccag aatgggaggc agaagtgtca caagaggttc gtgtctgtcc aggtcgtggc 3180

tatatccttc gtgttacagc gtacaaagag ggatatggag aaggttgcgt aacgatccat 3240

gagatcgaga acaatacaga cgaactgaaa ttcaacaact gtgtagaaga ggaagtatat 3300

ccaaacaaca cggtaacgtg tattaattat actgcgactc aagaagaata tgagggtacg 3360

tacacttctc gtaatcgagg atatgacgaa gcctatggta ataacccttc cgtaccagct 3420

gattatgcgt cagtctatga agaaaaatcg tatacagata gacgaagaga gaatccttgt 3480

gaatctaaca gaggatatgg agattacaca ccactaccag ctggttatgt aacaaaggaa 3540

ttagagtact tcccagagac cgataaggta tggattgaga ttggagaaac agaaggaaca 3600

ttcatcgtgg acagcgtgga attactcctt atggaggaat ag 3642

<210> 12
<211> 3642
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_20.

<400> 12
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccttcg aggccgagta cgaccttgag cgcgcccaga aggtggtgaa cgccctcttc 1980

actagcacta accagctagg cctgaagact gacgtgaccg actaccacat cgaccaagtg 2040

agcaacctag tggcctgcct ctccgacgag ttctgcctcg acgagaagcg cgagctgtcc 2100

gagaaggtga agcacgccaa gcgcctctcc gacgagcgca acctgctcca ggaccccaac 2160

ttcaggggca tcaacaggca gcccgaccgc ggctggcgcg gctccaccga catcaccatc 2220

cagggcggtg acgacgtatt caaggagaac tacgttaccc tccccggcac cttcgacgag 2280

tgttacccca cctacctcta ccagaagatc gacgagtcca agctgaaggc ctacacccgc 2340

taccagctcc gcggctacat cgaggactcc caggacctgg aaatctacct catccgctac 2400

aacgccaagc acgagatcgt gaacgtgcct ggcaccggca gcctctggcc tctcagcgtg 2460

gagaaccaga tcggcccttg cggcgagcct aaccgctgcg cccctcacct cgagtggaac 2520

cctgacctcc actgctcgtg cagggacggc gagaagtgcg cccaccatag ccaccacttc 2580

tctctggaca tcgacgtggg ctgcaccgac ctgaacgagg acctgggcgt gtgggttatc 2640

ttcaagatca agacccagga cggtcacgcc aggctgggta acctggagtt ccttgaggaa 2700

aagcctctgc tgggtgaggc cctggccagg gtcaagaggg ctgagaagaa atggagggat 2760

aagagggaga ccctgcagct ggagaccact atcgtctaca aggaggctaa ggagtctgtc 2820

gatgctctgt tcgtcaactc tcagtacgat agactgcaag ctgataccaa catcgctatg 2880

atccacgctg cggataagcg ggtccaccgg atccgggagg cttaccttcc ggagctttct 2940

gtcatcccgg gtgtcaacgc tgcgatcttc gaggaacttg aggaacggat cttcactgcg 3000

tttagtcttt acgatgcgcg gaacatcatc aagaacgggg acttcaacaa tggtctgctg 3060

tgctggaacg tcaagggtca tgtcgaggtc gaggaacaaa acaatcatcg tagtgtcctt 3120

gtcattcctg agtgggaggc ggaggtctct caagaggtcc gtgtttgccc ggggcgtggg 3180

tacattcttc gtgttactgc gtacaaggag gggtacgggg aggggtgcgt tactattcat 3240

gagattgaga acaatactga tgagcttaag ttcaacaatt gtgttgagga ggaggtttac 3300

ccgaacaata ctgttacgtg catcaactac acggcaacgc aagaggaata cgaggggacg 3360

tacacctcgc gtaatagagg gtatgatgag gcgtacggaa acaacccgtc ggttccagca 3420

gattatgcct cggtttatga ggagaagtcg tacacggata gacgacgcga gaatccatgt 3480

gagtcaaatc gaggatacgg agattacaca ccattaccag caggatacgt tacaaaggag 3540

ttggaatact tcccggaaac agataaagtt tggattgaaa tcggagaaac agaaggaaca 3600

ttcatcgtcg actcagtaga attgttgttg atggaagaat ga 3642

<210> 13
<211> 1213
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_20.

<400> 13

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Ile Ala Ser

```
65                  70                  75                  80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400
```

```
Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val
                645                 650                 655

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val
            660                 665                 670

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser
        675                 680                 685

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
    690                 695                 700

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
705                 710                 715                 720
```

```
Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr
                725                 730                 735

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
            740                 745                 750

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
        755                 760                 765

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
    770                 775                 780

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
785                 790                 795                 800

Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp
                805                 810                 815

Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg
            820                 825                 830

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg
        835                 840                 845

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
    850                 855                 860

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
865                 870                 875                 880

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
                885                 890                 895

Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys
            900                 905                 910

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu
        915                 920                 925

Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
    930                 935                 940

Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met
945                 950                 955                 960

Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu
                965                 970                 975

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
            980                 985                 990

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

Ile Ile  Lys Asn Gly Asp Phe  Asn Asn Gly Leu Leu  Cys Trp Asn
    1010                1015                1020

Val Lys  Gly His Val Glu Val  Glu Glu Gln Asn Asn  His Arg Ser
    1025                1030                1035

Val Leu  Val Ile Pro Glu Trp  Glu Ala Glu Val Ser  Gln Glu Val
```

1040 1045 1050

Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr
1055 1060 1065

Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu
1070 1075 1080

Asn Asn Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu
1085 1090 1095

Val Tyr Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr
1100 1105 1110

Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr
1115 1120 1125

Asp Glu Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala
1130 1135 1140

Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn
1145 1150 1155

Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro
1160 1165 1170

Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp
1175 1180 1185

Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val
1190 1195 1200

Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1205 1210

<210> 14
<211> 3690
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC867_21.

<400> 14
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta cacaaatacc attggtaaag gcgcataccc tccaatcggg taccactgta 1560

gtaaaagggc cagggtttac aggaggggat atcctccgtc gaacaagtgg aggaccattt 1620

gcttttagta atgttaatct agattttaac ttgtcacaaa ggtatcgtgc tagaattcgt 1680

tatgcctcta ctactaacct aagaatttac gtaacggttg caggtgaacg aatttttgct 1740

ggtcaatttg acaaaactat ggatgctggt gccccattaa cattccaatc ttttagttac 1800

gcaactatta atacagcttt tacattccca gaaagatcga gcagcttgac tgtaggtgcc 1860

gatacgttta gttcaggtaa tgaagtttat gtagatagat ttgaattaat cccagttact 1920

gcaaccggaa cgacaaccta tgagtatgaa gagaagcaga atctagaaaa agcgcagaaa 1980

gcgttgaacg ctttgtttac ggatggcacg aatggctatc tacaaatgga tgccactgat 2040

tatgatatca atcaaactgc aaacttaata gaatgtgtat cagatgaatt gtatgcaaaa 2100

gaaaagatag ttttattaga tgaagtcaaa tatgcgaagc ggcttagcat atcacgtaac 2160

ctacttttga acgatgattt agaattttca gatggatttg gagaaaacgg atggacgaca 2220

agtgataata tttcaatcca ggcggataat ccccttttta aggggaatta tttaaaaatg 2280

tttggggcaa gagatattga tggaacccta tttccaactt atctctatca aaaaatagat 2340

gagtccaggt taaaaccata tacacgttat cgagtaagag ggtttgtggg aagtagtaaa 2400

aatctaaaat tagtggtaac acgctatgag aaagaaattg atgccattat gaatgttcca 2460

aatgatttgg cacatatgca gcttaaccct tcatgtggag attatcgctg tgaatcatcg 2520

tcccagtttt tggtgaacca agtgcatcct acaccaacag ctggatatgc tcttgatatg 2580

tatgcatgcc cgtcaagttc agataaaaaa catattatgt gtcacgatcg tcatccattt 2640

gattttcata ttgacaccgg agaattaaat ccaaacacaa acctgggtat tgatgtcttg 2700

tttaaaattt ctaatccaaa tggatacgct acattaggga atctagaagt cattgaagaa 2760

ggaccactaa cagatgaagc attggtacat gtaaaacaaa aggaaaagaa atggcgtcag 2820

cacatggaga aaaaacgaat ggaaacacaa caagcctatg atccagcaaa acaagctgta 2880

gatgcattat ttacaaatga acaagagtta gactatcata ctactttaga tcatattcag 2940

aacgccgatc agctggtaca ggcgattccc tatgtacacc atgcttggtt accggatgct 3000

ccaggtatga actatgatgt atatcaaggg ttaaacgcac gtatcatgca ggcgtacaat 3060

ttatatgatg cacgaaatgt cataataaat ggtgacttta cacaaggact acaaggatgg 3120

cacgcaacag gaaaagcagc ggtacaacaa atagatggag cttcagtatt agttctatca 3180

aactggagtg ccgaggtatc tcagaatctg catgcccaag atcatcatgg atatatgtta 3240

cgtgtgattg ccaaaaaaga aggtcctgga aaagggtatg taatgatgat ggattttaat 3300

ggaaagcagg aaacacttac gttcacttct tgtgaagaag gatatataac aaaaacaata 3360

gaggtattcc cggaaagtga tcgaatacga attgaaatgg gagaaacaga gggtacgttt 3420

tatgtagata gcatcgagtt gctttgtatg caaggatatg ctagcgataa taacccgcac 3480

acgggtaata tgtatgagca aagttataat ggaaattata atcaaaatac tagcgatgtg 3540

tatcaccaag gatatataaa caactataac caaaattcta gtagtatgta taatcaaaat 3600

tatattaaca atgatgacct gcattccggt tgcacatgta accaagggca taactctggc 3660

tgtacatgta atcaaggata taaccgttag 3690

<210> 15
<211> 3690
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_21.

<400> 15
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccggga ctaccaccta cgagtacgag gagaagcaga atctcgagaa ggctcagaag 1980

gctctgaacg ctctgttcac tgacgggacc aacggctacc tccagatgga cgccactgac 2040

tacgacatca accagacagc taacctgatt gagtgtgtga gtgacgaact gtacgctaag 2100

gagaagatcg tactcctgga cgaggtgaag tacgctaagc gcctgagcat tagccgtaac 2160

ctgctgctga acgacgatct ggagttcagc gacggctttg gcgagaacgg ctggaccacc 2220

agcgacaaca tctccatcca ggccgacaat ccactcttca aaggcaacta cctcaagatg 2280

ttcggagcca gggacatcga cggcaccctc tttccgacct acctctacca gaagatcgac 2340

gagtcccgcc tcaaacccta cacccgctac agggtgcgcg gcttcgtggg cagcagcaag 2400

aacctcaagc tcgtggtcac acggtatgag aaggagatcg acgccatcat gaacgtgccc 2460

aacgatctcg cccacatgca gctcaatcca tcctgcggcg actaccggtg cgagtccagc 2520

tcccagttcc tcgtgaacca ggtgcaccct actccgaccg ctggctatgc cctggacatg 2580

tacgcctgcc ctagttcctc cgacaagaag cacatcatgt gccacgaccg tcatccgttc 2640

gacttccaca tcgacaccgg cgaactgaac ccgaacacca acctgggcat cgacgtactg 2700

ttcaagattt ccaacccgaa cgggtacgcc accttgggca acctggaggt catcgaagaa 2760

ggcccgctga ccgacgaggc cctggtccac gtcaaacaga aggagaagaa gtggcggcag 2820

cacatggaga agaagcggat ggagactcaa caagcctacg acccggccaa gcaagctgtg 2880

gacgctctgt tcaccaacga gcaagagctt gactaccaca ctactcttga ccacatccag 2940

aatgctgacc agcttgtcca ggctattccg tacgtccacc acgcttggct accggacgct 3000

ccagggatga actacgatgt gtaccagggt ctgaacgcgc ggatcatgca agcgtacaac 3060

ctgtacgacg cgcgtaacgt catcatcaac ggtgacttca ctcagggtct tcaaggttgg 3120

cacgcgactg gcaaagcggc agtccagcag attgatggtg cgtctgttct tgtgttgagc 3180

aactggtctg cggaggtttc tcagaacctg cacgcacagg atcaccacgg ctacatgctg 3240

agggtgattg ctaagaagga gggccctggc aaaggctacg tcatgatgat ggacttcaac 3300

ggaaagcaag aaaccctgac cttcactagc tgtgaggagg gctacatcac taagaccatt 3360

gaggtctttc cggagtctga ccgcatccgg atcgagatgg gcgagaccga aggcacgttc 3420

tacgtggact ccatcgaact cctctgcatg caaggctacg cctccgacaa caacccacac 3480

acgggcaaca tgtacgagca gtcctacaac gggaactaca accagaacac ctccgatgtg 3540

taccatcagg gctacatcaa caactacaac cagaacagca gcagcatgta caaccagaac 3600

tacatcaaca acgatgactt gcactcgggt tgcacctgca accagggtca caacagtggg 3660

tgcacgtgca accagggata caaccgttga 3690

<210> 16
<211> 1229

<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_21.

<400> 16

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
165 170 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
195 200 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
210 215 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225 230 235 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
245 250 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
260 265 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
275 280 285

```
Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605
```

```
Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Gly Thr Thr Thr Tyr Glu Tyr Glu Glu Lys Gln Asn Leu Glu
                645                 650                 655

Lys Ala Gln Lys Ala Leu Asn Ala Leu Phe Thr Asp Gly Thr Asn Gly
            660                 665                 670

Tyr Leu Gln Met Asp Ala Thr Asp Tyr Asp Ile Asn Gln Thr Ala Asn
        675                 680                 685

Leu Ile Glu Cys Val Ser Asp Glu Leu Tyr Ala Lys Glu Lys Ile Val
    690                 695                 700

Leu Leu Asp Glu Val Lys Tyr Ala Lys Arg Leu Ser Ile Ser Arg Asn
705                 710                 715                 720

Leu Leu Leu Asn Asp Asp Leu Glu Phe Ser Asp Gly Phe Gly Glu Asn
                725                 730                 735

Gly Trp Thr Thr Ser Asp Asn Ile Ser Ile Gln Ala Asp Asn Pro Leu
            740                 745                 750

Phe Lys Gly Asn Tyr Leu Lys Met Phe Gly Ala Arg Asp Ile Asp Gly
        755                 760                 765

Thr Leu Phe Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Arg Leu
    770                 775                 780

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

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

Met Asn Val Pro Asn Asp Leu Ala His Met Gln Leu Asn Pro Ser Cys
            820                 825                 830

Gly Asp Tyr Arg Cys Glu Ser Ser Ser Gln Phe Leu Val Asn Gln Val
        835                 840                 845

His Pro Thr Pro Thr Ala Gly Tyr Ala Leu Asp Met Tyr Ala Cys Pro
    850                 855                 860

Ser Ser Ser Asp Lys Lys His Ile Met Cys His Asp Arg His Pro Phe
865                 870                 875                 880

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

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

Gly Asn Leu Glu Val Ile Glu Glu Gly Pro Leu Thr Asp Glu Ala Leu
        915                 920                 925

Val His Val Lys Gln Lys Glu Lys Lys Trp Arg Gln His Met Glu Lys
```

```
    930                 935                 940

Lys Arg Met Glu Thr Gln Gln Ala Tyr Asp Pro Ala Lys Gln Ala Val
945                 950                 955                 960

Asp Ala Leu Phe Thr Asn Glu Gln Glu Leu Asp Tyr His Thr Thr Leu
                965                 970                 975

Asp His Ile Gln Asn Ala Asp Gln Leu Val Gln Ala Ile Pro Tyr Val
            980                 985                 990

His His Ala Trp Leu Pro Asp Ala  Pro Gly Met Asn Tyr  Asp Val Tyr
        995                 1000                1005

Gln Gly  Leu Asn Ala Arg Ile  Met Gln Ala Tyr Asn  Leu Tyr Asp
    1010                1015                1020

Ala Arg  Asn Val Ile Ile Asn  Gly Asp Phe Thr Gln  Gly Leu Gln
    1025                1030                1035

Gly Trp  His Ala Thr Gly Lys  Ala Ala Val Gln Gln  Ile Asp Gly
    1040                1045                1050

Ala Ser  Val Leu Val Leu Ser  Asn Trp Ser Ala Glu  Val Ser Gln
    1055                1060                1065

Asn Leu  His Ala Gln Asp His  His Gly Tyr Met Leu  Arg Val Ile
    1070                1075                1080

Ala Lys  Lys Glu Gly Pro Gly  Lys Gly Tyr Val Met  Met Met Asp
    1085                1090                1095

Phe Asn  Gly Lys Gln Glu Thr  Leu Thr Phe Thr Ser  Cys Glu Glu
    1100                1105                1110

Gly Tyr  Ile Thr Lys Thr Ile  Glu Val Phe Pro Glu  Ser Asp Arg
    1115                1120                1125

Ile Arg  Ile Glu Met Gly Glu  Thr Glu Gly Thr Phe  Tyr Val Asp
    1130                1135                1140

Ser Ile  Glu Leu Leu Cys Met  Gln Gly Tyr Ala Ser  Asp Asn Asn
    1145                1150                1155

Pro His  Thr Gly Asn Met Tyr  Glu Gln Ser Tyr Asn  Gly Asn Tyr
    1160                1165                1170

Asn Gln  Asn Thr Ser Asp Val  Tyr His Gln Gly Tyr  Ile Asn Asn
    1175                1180                1185

Tyr Asn  Gln Asn Ser Ser Ser  Met Tyr Asn Gln Asn  Tyr Ile Asn
    1190                1195                1200

Asn Asp  Asp Leu His Ser Gly  Cys Thr Cys Asn Gln  Gly His Asn
    1205                1210                1215

Ser Gly  Cys Thr Cys Asn Gln  Gly Tyr Asn Arg
    1220                1225
```

<210> 17
<211> 3432

<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC867_22.

<400> 17
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta cacaaatacc attggtaaag gcgcataccc tccaatcggg taccactgta 1560

gtaaaagggc cagggtttac aggaggggat atcctccgtc gaacaagtgg aggaccattt 1620

gcttttagta atgttaatct agatttaac ttgtcacaaa ggtatcgtgc tagaattcgt 1680

tatgcctcta ctactaacct aagaatttac gtaacggttg caggtgaacg aatttttgct 1740

ggtcaatttg acaaaactat ggatgctggt gccccattaa cattccaatc ttttagttac 1800

gcaactatta atacagcttt tacattccca gaaagatcga gcagcttgac tgtaggtgcc 1860

gatacgttta gttcaggtaa tgaagtttat gtagatagat ttgaattaat cccagttact 1920

gcaaccaatc cgacgcgaga ggcggaagag gatctagaag cagcgaagaa agcggtggcg 1980

agcttgttta cacgtacaag ggacggatta caagtaaatg tgacagatta tcaagtcgat 2040

caagcggcaa atttagtgtc atgcttatca gatgaacaat atgggcatga caaaaagatg 2100

ttattggaag cggtaagagc ggcaaaacgc ctcagccgag aacgcaactt acttcaggat 2160

ccagatttta atacaatcaa tagtacagaa gaaaatggat ggaaagcaag taacggcgtt 2220

actattagcg agggcggtcc attctataaa ggccgtgcgc ttcagctagc aagcgcaaga 2280

gaaaattacc caacatacat ttatcaaaaa gtaaatgcat cagagttaaa gccgtataca 2340

cgttatagac tggatgggtt cgtgaagagt agtcaagatt tagaaattga tctcattcac 2400

catcataaag tccatctcgt gaaaaatgta ccagataatt tagtatccga tacttactcg 2460

gatggttctt gcagtggaat gaatcgatgt gaggaacaac agatggtaaa tgcgcaactg 2520

gaaacagaac atcatcatcc gatggattgc tgtgaagcgg ctcaaacaca tgagttttct 2580

tcctatatta atacaggcga tctaaattca agtgtagatc aaggcatttg ggttgtattg 2640

aaagttcgaa caaccgatgg ttatgcgacg ctaggaaatc ttgaattggt agaggtcgga 2700

ccgttatcgg gtgaatctct agaacgtgaa caaagggata atgcgaaatg gagtgcagag 2760

ctaggaagaa agcgtgcaga aacagatcgc gtgtatcaag atgccaaaca atccatcaat 2820

catttatttg tggattatca agatcaacaa ttaaatccag aaatagggat ggcagatatt 2880

attgacgctc aaaatcttgt cgcatcaatt tcagatgtgt atagcgatgc agtactgcaa 2940

atccctggaa ttaactatga gatttacaca gagctatcca atcgcttaca acaagcatcg 3000

tatctgtata cgtctcgaaa tgcggtgcaa aatggggact ttaacagcgg tctagatagt 3060

tggaatgcaa caggggggc tacggtacaa caggatggca atacgcattt cttagttctt 3120

tctcattggg atgcacaagt ttctcaacaa tttagagtgc agccgaattg taaatatgta 3180

ttacgtgtaa cagcagagaa agtaggcggc ggagacggat acgtgacaat ccgggatggt 3240

gctcatcata cagaaaagct tacatttaat gcatgtgatt atgatataaa tggcacgtac 3300

gtgactgata atacgtatct aacaaaagaa gtggtattct attcacatac agaacacatg 3360

tgggtagagg taagtgaaac agaaggtgca tttcatatag atagtattga attcgttgaa 3420

acagaaaagt ag 3432

<210> 18
<211> 3432
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_22.

<400> 18
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccaacc cgacgcggga agctgaggaa gacttggaag ccgccaagaa agcggtcgcc 1980

agcctgttta ctcggacgcg ggacgggctc caagtgaatg tgacggacta tcaagtggat 2040

caggccgcta acctcgtgtc atgcctgagc gacgagcagt acggtcacga caagaaaatg 2100

ctgctggagg ccgtccgggc cgccaagcgg ctgtccaggg agcgtaacct gctacaagat 2160

cccgacttta acacgatcaa cagcacagag gagaatggct ggaaggccag caacggagtt 2220

acgataagcg agggcggtcc gttctacaag ggtcgtgccc tccagctcgc ctctgcaagg 2280

gagaactatc caacctacat ctatcagaag gtgaacgcat ccgagcttaa gccctacaca 2340

cgctaccgcc tggacgggtt cgttaagtcc agtcaagacc tagagataga cctcatccac 2400

caccacaaag tgcatctggt caagaacgtt cccgataatc tcgtgagcga tacctactca 2460

gacggctcat gctctggcat gaacagatgt gaggagcaac agatggttaa tgctcaactc 2520

gaaaccgagc atcatcatcc tatggattgc tgcgaggccg cgcagaccca tgagttcagc 2580

tcttacatca acaccggaga cctcaacagt agcgtggatc agggaatttg ggtggtgctt 2640

aaagtgcgta caaccgacgg ctacgccacc ctcggcaacc ttgagcttgt cgaggtcgga 2700

ccacttagcg gcgagtccct ggaacgtgag cagcgggaca acgccaaatg gagcgcagag 2760

ctagggcgca aacgcgcgga gacggaccgg gtttatcagg acgcgaagca gtccatcaat 2820

cacctcttcg tggattatca ggaccagcag cttaatccag agatcggcat ggccgacatc 2880

atcgacgccc agaacctagt agcgtcgatt tccgatgtct attccgacgc cgtgcttcaa 2940

atacctggca tcaactacga gatctacaca gagttgtcca acaggctcca gcaagcgtca 3000

tacctctaca ccagccgcaa cgccgtccag aatggcgact tcaattccgg actagactcc 3060

tggaacgcca cgggcggagc tacggtgcaa caagacggca acacccactt cctcgtactt 3120

agccactggg acgctcaagt gagtcagcaa ttccgggttc agccgaactg caagtacgtc 3180

ctgcgcgtaa cggccgagaa ggttggaggc ggagacggct acgttaccat ccgcgacggc 3240

gctcaccaca ccgagaaact gacgttcaac gcttgtgact acgacatcaa cggcacttac 3300

gtgacggaca acacctacct gacgaaggag gtggtgttct attctcacac cgagcacatg 3360

tgggttgagg tcagcgagac cgagggagcc ttccacattg acagcatcga gttcgtggag 3420

actgagaagt ga 3432

<210> 19
<211> 1143
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_22.

<400> 19

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
165 170 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
195 200 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
210 215 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225 230 235 240

```
Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
```

```
                565                 570                 575

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Asn Pro Thr Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys
                645                 650                 655

Lys Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val
            660                 665                 670

Asn Val Thr Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys
        675                 680                 685

Leu Ser Asp Glu Gln Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala
    690                 695                 700

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

Pro Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala
                725                 730                 735

Ser Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe Tyr Lys Gly Arg
            740                 745                 750

Ala Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr
        755                 760                 765

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

Asp Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His
785                 790                 795                 800

His His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser
                805                 810                 815

Asp Thr Tyr Ser Asp Gly Ser Cys Ser Gly Met Asn Arg Cys Glu Glu
            820                 825                 830

Gln Gln Met Val Asn Ala Gln Leu Glu Thr Glu His His His Pro Met
        835                 840                 845

Asp Cys Cys Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn
    850                 855                 860

Thr Gly Asp Leu Asn Ser Ser Val Asp Gln Gly Ile Trp Val Val Leu
865                 870                 875                 880

Lys Val Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu
                885                 890                 895
```

Val Glu Val Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg
900 905 910

Asp Asn Ala Lys Trp Ser Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr
915 920 925

Asp Arg Val Tyr Gln Asp Ala Lys Gln Ser Ile Asn His Leu Phe Val
930 935 940

Asp Tyr Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly Met Ala Asp Ile
945 950 955 960

Ile Asp Ala Gln Asn Leu Val Ala Ser Ile Ser Asp Val Tyr Ser Asp
965 970 975

Ala Val Leu Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu
980 985 990

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

Val Gln Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala
1010 1015 1020

Thr Gly Gly Ala Thr Val Gln Gln Asp Gly Asn Thr His Phe Leu
1025 1030 1035

Val Leu Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val
1040 1045 1050

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

Gly Gly Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His
1070 1075 1080

Thr Glu Lys Leu Thr Phe Asn Ala Cys Asp Tyr Asp Ile Asn Gly
1085 1090 1095

Thr Tyr Val Thr Asp Asn Thr Tyr Leu Thr Lys Glu Val Val Phe
1100 1105 1110

Tyr Ser His Thr Glu His Met Trp Val Glu Val Ser Glu Thr Glu
1115 1120 1125

Gly Ala Phe His Ile Asp Ser Ile Glu Phe Val Glu Thr Glu Lys
1130 1135 1140

<210> 20
<211> 3696
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_23.

<400> 20
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccacgg cgaccttcga ggcggagtat gacttggagc gggctcagga ggccgtcaac 1980

gcgctgttca caaacaccaa tcctcgccgc ctcaagacgg gtgtgactga ttaccacatt 2040

gacgaggtct ccaacttggt cgcgtgtctg tccgatgagt tctgcctgga cgagaagcgg 2100

gaactgctgg agaaggtcaa gtacgccaag cgcctctccg acgaaaggaa cctcctccaa 2160

gatcccaact ttacttccat taacaagcag ccggacttca tctccaccaa cgagcagtcc 2220

aacttcacct caatccacga gcagtcggag cacgggtggt ggggcagcga gaacatcacc 2280

atccaagagg gcaacgacgt cttcaaggag aactacgtga tcctgcccgg caccttcaac 2340

gagtgttacc cgacctatct ctaccagaag attggcgaag cggaactcaa ggcttacacc 2400

cgttaccaac tgagtggcta cattgaggac tcacaagacc tggaaatcta cctgatccgc 2460

tacaacgcca agcacgagac cctcgacgtg cctggcacgg agtccgtctg gcccttgagc 2520

gtggagtctc ctatcggtcg ttgcggcgag cccaatcgct gcgctccgca ctttgagtgg 2580

aatcctgatt tggattgctc ctgccgagac ggtgagaaat gcgcccacca ctcgcaccac 2640

ttcagcctag acatcgacgt gggctgcatc gacctgcacg agaacttggg cgtctgggtc 2700

gtgttcaaga tcaagacaca ggagggccat gctcggcttg ggaacctgga gttcatcgag 2760

gagaagccac tgctgggtga agccttgtca cgggtgaaac gcgccgagaa gaagtggcgg 2820

gacaaacggg agaagctcca gttggagaca aagcgtgtgt acacagaggc caaggaggcc 2880

gtggatgcct tgttcgtgga cagtcagtac gacaggctgc aagcggacac caacatcggg 2940

atgatccacg cggctgataa gcttgttcac agaatccgcg aggcgtacct gtcagagctt 3000

agcgtgatcc caggcgtcaa cgccgaaatc ttcgaggaac tggagggccg cattatcacg 3060

gcaatctcac tttatgacgc gaggaatgtg gtcaagaacg gtgacttcaa caacggcttg 3120

gcgtgttgga acgttaaagg gcacgtggat gtacaacagt cacaccacag aagtgtcttg 3180

gtcatcccgg agtgggaggc ggaagtgagc caggccgtcc gggtctgccc tgggcgcggt 3240

tacatcctcc gcgtgacagc gtacaaggag ggctacggtg agggctgcgt gacgatccac 3300

gagattgaga acaacacgga cgagcttaag ttcaagaact gcgaggagga ggaagtgtac 3360

ccgacagaca ccggcacctg caacgactac accgcccacc aagggaccgc cgcctgcaac 3420

agccgcaacg cgggctatga agatgcgtac gaggttgata ccaccgcctc agtgaactac 3480

aaaccgactt atgaggagga gacatacacg gacgtcaggc gcgacaacca ttgtgagtac 3540

gaccgtggct acgtgaacta tccgccggtg ccagcgggct acatgacgaa ggagctagaa 3600

tacttccctg agacggacaa ggtgtggatt gaaatcggcg agaccgaggg caagtttatc 3660

gtggattctg tcgagctgct gctaatggag gagtag 3696

<210> 21
<211> 1231
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_23.

<400> 21

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
165 170 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
195 200 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
210 215 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225 230 235 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
245 250 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
260 265 270

```
Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
```

595 600 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
610 615 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625 630 635 640

Ala Thr Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln
645 650 655

Glu Ala Val Asn Ala Leu Phe Thr Asn Thr Asn Pro Arg Arg Leu Lys
660 665 670

Thr Gly Val Thr Asp Tyr His Ile Asp Glu Val Ser Asn Leu Val Ala
675 680 685

Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Leu Glu
690 695 700

Lys Val Lys Tyr Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
705 710 715 720

Asp Pro Asn Phe Thr Ser Ile Asn Lys Gln Pro Asp Phe Ile Ser Thr
725 730 735

Asn Glu Gln Ser Asn Phe Thr Ser Ile His Glu Gln Ser Glu His Gly
740 745 750

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

Lys Glu Asn Tyr Val Ile Leu Pro Gly Thr Phe Asn Glu Cys Tyr Pro
770 775 780

Thr Tyr Leu Tyr Gln Lys Ile Gly Glu Ala Glu Leu Lys Ala Tyr Thr
785 790 795 800

Arg Tyr Gln Leu Ser Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile
805 810 815

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Leu Asp Val Pro Gly
820 825 830

Thr Glu Ser Val Trp Pro Leu Ser Val Glu Ser Pro Ile Gly Arg Cys
835 840 845

Gly Glu Pro Asn Arg Cys Ala Pro His Phe Glu Trp Asn Pro Asp Leu
850 855 860

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
865 870 875 880

Phe Ser Leu Asp Ile Asp Val Gly Cys Ile Asp Leu His Glu Asn Leu
885 890 895

Gly Val Trp Val Val Phe Lys Ile Lys Thr Gln Glu Gly His Ala Arg
900 905 910

Leu Gly Asn Leu Glu Phe Ile Glu Glu Lys Pro Leu Leu Gly Glu Ala
915 920 925

```
Leu Ser Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu
    930                 935                 940

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

Val Asp Ala Leu Phe Val Asp Ser Gln Tyr Asp Arg Leu Gln Ala Asp
                965                 970                 975

Thr Asn Ile Gly Met Ile His Ala Ala Asp Lys Leu Val His Arg Ile
            980                 985                 990

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

Glu Ile  Phe Glu Glu Leu Glu  Gly Arg Ile Ile Thr  Ala Ile Ser
    1010                1015                1020

Leu Tyr  Asp Ala Arg Asn Val  Val Lys Asn Gly Asp  Phe Asn Asn
    1025                1030                1035

Gly Leu  Ala Cys Trp Asn Val  Lys Gly His Val Asp  Val Gln Gln
    1040                1045                1050

Ser His  His Arg Ser Val Leu  Val Ile Pro Glu Trp  Glu Ala Glu
    1055                1060                1065

Val Ser  Gln Ala Val Arg Val  Cys Pro Gly Arg Gly  Tyr Ile Leu
    1070                1075                1080

Arg Val  Thr Ala Tyr Lys Glu  Gly Tyr Gly Glu Gly  Cys Val Thr
    1085                1090                1095

Ile His  Glu Ile Glu Asn Asn  Thr Asp Glu Leu Lys  Phe Lys Asn
    1100                1105                1110

Cys Glu  Glu Glu Glu Val Tyr  Pro Thr Asp Thr Gly  Thr Cys Asn
    1115                1120                1125

Asp Tyr  Thr Ala His Gln Gly  Thr Ala Ala Cys Asn  Ser Arg Asn
    1130                1135                1140

Ala Gly  Tyr Glu Asp Ala Tyr  Glu Val Asp Thr Thr  Ala Ser Val
    1145                1150                1155

Asn Tyr  Lys Pro Thr Tyr Glu  Glu Glu Thr Tyr Thr  Asp Val Arg
    1160                1165                1170

Arg Asp  Asn His Cys Glu Tyr  Asp Arg Gly Tyr Val  Asn Tyr Pro
    1175                1180                1185

Pro Val  Pro Ala Gly Tyr Met  Thr Lys Glu Leu Glu  Tyr Phe Pro
    1190                1195                1200

Glu Thr  Asp Lys Val Trp Ile  Glu Ile Gly Glu Thr  Glu Gly Lys
    1205                1210                1215

Phe Ile  Val Asp Ser Val Glu  Leu Leu Leu Met Glu  Glu
    1220                1225                1230
```

<210> 22
<211> 3666
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_24.

<400> 22
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccaccg cgacgtttga agctgaatcc gacctcgagc gtgcgcgcaa ggcggtgaac 1980

gctctgttca cgagcaccaa ccctcgtggc ttgaagacgg atgtgacgga ctaccacatc 2040

gaccaagtct cgaacctcgt ggagtgcctg agcgacgagt tctgtcttga caagaagcgc 2100

gagctgctgg aggaggtgaa gtacgccaag cgcctctccg atgagcgcaa cctgctccaa 2160

gatcctacct tcacgtcgat ttccggccaa accgaccgtg gatggatcgg ctcgactggc 2220

atctccatcc agggcggcga cgacatcttc aaggagaact atgttcggct gccgggcacg 2280

gtggacgagt gttacccgac gtacctctac cagaagatag acgagagtca actcaagtcc 2340

tacacgcggt atcagttacg tggctacatt gaagactccc aggacttgga aatctatctc 2400

atacggtaca acgccaagca cgagacctta agcgtgccgg gaacggagtc gccctggcca 2460

agctctggcg tgtacccttc cggtaggtgc ggcgagccca accgctgtgc acctcgaatc 2520

gaatggaacc cggaccttga ctgctcttgc cggtacggcg agaagtgcgt ccatcattct 2580

caccacttca gcttggacat tgacgtcggc tgcaccgacc tcaatgaaga cctcggagtg 2640

tgggtcatct tcaagatcaa gacacaggac gggcacgcga aactaggaaa cctggagttc 2700

atcgaggaga agccactcct cggcaaggca ctttccaggg tcaagcgggc cgagaagaaa 2760

tggagggaca agtacgagaa actccagctc gaaacaaagc gggtgtacac ggaggcaaag 2820

gaatccgtgg acgccctgtt cgtggactct cagtacgaca agctccaggc gaacacaaac 2880

attggcatca tccacggtgc ggacaagcaa gtgcacagga tacgggagcc ttacctctcg 2940

gagctgccgg tgattccctc gatcaacgcg gcgatcttcg aggaactgga gggccacatc 3000

ttcaaggcgt attctctgta cgacgcgcgt aacgtcatca agaacggcga cttcaacaat 3060

gggctgtcct gctggaacgt taaaggccac gtcgatgtcc agcagaacca ccataggtca 3120

gtcctggtgc tgagcgagtg ggaggcggag gtgtcccaga aggtgcgcgt gtgcccggat 3180

cgcggctaca tcttgagggt gacagcctac aaggagggct acggcgaggg ctgtgtcacg 3240

atccatgagt tcgaggacaa cacggatgtc ctgaaattcc gtaacttcgt cgaggaggag 3300

gtctatccca acaacaccgt gacctgcaac gactacacga ccaatcagtc ggctgagggc 3360

agtaccgatg cctgcaacag ctacaaccgt ggttacgaag atggatacga gaaccgctac 3420

gagcccaatc cttcggctcc cgtgaattac actcccacgt acgaggaggg catgtacact 3480

gacactcagg gctacaacca ttgcgtcagc gaccgtggct accgcaacca cacgccgctc 3540

ccagcgggct acgtgacgct ggagctggaa tactttcccg agacagaaca agtgtggata 3600

gagatcggcg agaccgaggg cacattcatc gtgggctctg tggaattgct cctcatggag 3660

gagtaa 3666

<210> 23
<211> 1221
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_24.

<400> 23

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
165 170 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
195 200 205

```
Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Phe Ser Asn
```

```
    530                 535                 540

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Thr Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Arg
                645                 650                 655

Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Pro Arg Gly Leu Lys
            660                 665                 670

Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu
        675                 680                 685

Cys Leu Ser Asp Glu Phe Cys Leu Asp Lys Lys Arg Glu Leu Leu Glu
    690                 695                 700

Glu Val Lys Tyr Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
705                 710                 715                 720

Asp Pro Thr Phe Thr Ser Ile Ser Gly Gln Thr Asp Arg Gly Trp Ile
                725                 730                 735

Gly Ser Thr Gly Ile Ser Ile Gln Gly Gly Asp Asp Ile Phe Lys Glu
            740                 745                 750

Asn Tyr Val Arg Leu Pro Gly Thr Val Asp Glu Cys Tyr Pro Thr Tyr
        755                 760                 765

Leu Tyr Gln Lys Ile Asp Glu Ser Gln Leu Lys Ser Tyr Thr Arg Tyr
    770                 775                 780

Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu
785                 790                 795                 800

Ile Arg Tyr Asn Ala Lys His Glu Thr Leu Ser Val Pro Gly Thr Glu
                805                 810                 815

Ser Pro Trp Pro Ser Ser Gly Val Tyr Pro Ser Gly Arg Cys Gly Glu
            820                 825                 830

Pro Asn Arg Cys Ala Pro Arg Ile Glu Trp Asn Pro Asp Leu Asp Cys
        835                 840                 845

Ser Cys Arg Tyr Gly Glu Lys Cys Val His His Ser His His Phe Ser
    850                 855                 860
```

```
Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val
865                 870                 875                 880

Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Lys Leu Gly
                885                 890                 895

Asn Leu Glu Phe Ile Glu Glu Lys Pro Leu Leu Gly Lys Ala Leu Ser
            900                 905                 910

Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Tyr Glu Lys Leu
        915                 920                 925

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

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

Ile Gly Ile Ile His Gly Ala Asp Lys Gln Val His Arg Ile Arg Glu
                965                 970                 975

Pro Tyr Leu Ser Glu Leu Pro Val Ile Pro Ser Ile Asn Ala Ala Ile
            980                 985                 990

Phe Glu Glu Leu Glu Gly His Ile  Phe Lys Ala Tyr Ser  Leu Tyr Asp
        995                 1000                1005

Ala Arg  Asn Val Ile Lys Asn  Gly Asp Phe Asn Asn  Gly Leu Ser
    1010                1015                1020

Cys Trp  Asn Val Lys Gly His  Val Asp Val Gln Gln  Asn His His
    1025                1030                1035

Arg Ser  Val Leu Val Leu Ser  Glu Trp Glu Ala Glu  Val Ser Gln
    1040                1045                1050

Lys Val  Arg Val Cys Pro Asp  Arg Gly Tyr Ile Leu  Arg Val Thr
    1055                1060                1065

Ala Tyr  Lys Glu Gly Tyr Gly  Glu Gly Cys Val Thr  Ile His Glu
    1070                1075                1080

Phe Glu  Asp Asn Thr Asp Val  Leu Lys Phe Arg Asn  Phe Val Glu
    1085                1090                1095

Glu Glu  Val Tyr Pro Asn Asn  Thr Val Thr Cys Asn  Asp Tyr Thr
    1100                1105                1110

Thr Asn  Gln Ser Ala Glu Gly  Ser Thr Asp Ala Cys  Asn Ser Tyr
    1115                1120                1125

Asn Arg  Gly Tyr Glu Asp Gly  Tyr Glu Asn Arg Tyr  Glu Pro Asn
    1130                1135                1140

Pro Ser  Ala Pro Val Asn Tyr  Thr Pro Thr Tyr Glu  Glu Gly Met
    1145                1150                1155

Tyr Thr  Asp Thr Gln Gly Tyr  Asn His Cys Val Ser  Asp Arg Gly
    1160                1165                1170
```

Tyr Arg Asn His Thr Pro Leu Pro Ala Gly Tyr Val Thr Leu Glu
1175 1180 1185

Leu Glu Tyr Phe Pro Glu Thr Glu Gln Val Trp Ile Glu Ile Gly
1190 1195 1200

Glu Thr Glu Gly Thr Phe Ile Val Gly Ser Val Glu Leu Leu Leu
1205 1210 1215

Met Glu Glu
1220

<210> 24
<211> 3651
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC867_25.

<400> 24
atgaccagca accgaaagaa cgagaacgag atcatcaacg ccctgtccat accggccgtg 60

tcaaaccact ccgcccagat gaacctctcc accgacgcga ggatcgagga ctccctctgc 120

atcgccgagg gcaacaacat cgacccgttc gtgtctgcaa gcacggtcca gaccggcatc 180

aacatcgcgg gccgcatcct gggcgtgctc ggcgtgccct tcgcgggtca aatcgcctct 240

ttctactcat tcctcgtggg cgagctgtgg ccgcgcggac gtgacccgtg ggaaatcttc 300

ctggagcacg ttgagcagct catccggcag caagtgaccg agaacaccag ggacaccgca 360

ctggcacggc tccagggcct tggcaacagc ttccgcgcct accagcagtc gctggaggac 420

tggctggaga accgagacga cgccagaacc cgctcagttc tgtacacaca gtacatcgcc 480

ctagagctgg acttcctcaa cgctatgccg ctcttcgcca tccgtaacca ggaagtaccg 540

cttctgatgg tgtacgcaca agcagcgaac ctccatctgc tcctgctgcg agacgcatct 600

ctgttcggca gtgagttcgg gctgacgagc caggagatcc agcgctacta cgagcgccaa 660

gtggagaaga ctcgtgagta cagcgactac tgcgcgcgct ggtacaacac gggcttgaac 720

aaccttcgcg ggacaaacgc cgaatcctgg cttcgctaca accagttccg ccgcgacctc 780

acgctgggtg tgctggacct ggtcgcgctc ttcccgtcct acgacacacg ggtgtaccca 840

atgaacacga gcgcacagct cacccgtgag atctacacag atcccatcgg ccgcaccaac 900

gctcccagtg gcttcgcaag cacgaattgg ttcaacaata acgctccttc tttctctgcc 960

atcgaggccg ctgtcatcag accgccgcac ttactcgatt tcccggagca gctcactatc 1020

ttctctgtgt tgtcccggtg gtcgaacacg cagtacatga actactgggt gggccacagg 1080

ctagagagcc ggaccatccg tggcagtctc tcaacctcga cccacggcaa cacgaacacg 1140

agcatcaacc ctgtcactct ccagtttaca tctagggacg tttacaggac agagtcgttc 1200

gctggcatta acattctgtt gaccactccg gtgaacggcg tcccttgggc ccgcttcaac 1260

tggaggaatc ctctgaactc actgcgcggc agccttctct acactatcgg ctacaccggc 1320

gttgggacgc aactcttcga ctcggagacc gagctgccgc ccgagaccac cgagcggcct 1380

aactacgaga gttattcaca caggctctcc aacatccgct tgatttctgg gaacaccttg 1440

cgggctccgg tgtactcctg gacgcaccgc agcgccgaca gaactaatac catcagctcc 1500

gactcgatca cccagatccc gctggtgaag gctcacacgc ttcagtcggg caccacagtc 1560

gtcaagggcc ctggcttcac cggcggcgac atcctgcgtc gcacatctgg cggacccttc 1620

gccttcagca acgtgaactt ggacttcaat ttgtcacagc ggtatcgtgc cagaatccgg 1680

tacgccagca ctacgaacct gcgaatctat gttactgtgg cgggcgagcg gatcttcgcc 1740

gggcaattcg acaagacgat ggacgcggga gcacctctga cattccagtc attctcttac 1800

gccacgatca acacggcatt cacgtttccg gagcgttcca gtagcctgac cgtgggcgct 1860

gataccttca gtagcgggaa cgaggtgtac gttgaccgtt tcgagctgat cccggtcacc 1920

gccaccgatg ctacctttga agcagagtcc gacttggaac gtgcacagaa ggcagtgaac 1980

gcactcttca cctcaagcaa ccagatcgga ttgaagacag atgtgacaga ttaccacatc 2040

gaccaagtga gcaacttggt ggattgcttg tcagatgagt tctgcttgga tgagaagcgt 2100

gaactctccg agaaggtgaa gcacgcaaag cgtctctcag atgaacgtaa tctccttcaa 2160

gaccctaact ttcgtggtat caatcgtcag ccagatcgtg gatggcgtgg atcaacagac 2220

atcaccatcc agggaggcga tgatgtgttc aaggagaact acgtgaccct cccaggaacc 2280

gtggatgaat gctacccaac ctacctctac cagaagatcg acgagtcaaa gctcaaggct 2340

tacacccgtt atgaactccg tggctacatc gaagatagcc aggatctcga aatctatctc 2400

atccgttaca atgctaagca cgaaatcgtg aatgtgccag gaaccggctc actctggcca 2460

ctctcagcac agtcaccaat cggcaagtgc ggcgaaccca atcgctgcgc tcctcatctc 2520

gaatggaatc ccgatctcga ctgctcctgc cgagacggcg agaagtgtgc acatcactca 2580

caccacttca ccctcgacat cgacgtgggc tgcaccgacc tcaatgaaga cctgggcgtg 2640

tgggtgatct tcaagatcaa gacccaggac ggccacgcac gactgggcaa tctggagttt 2700

ctggaggaga agccactgct tggcgaggca ctggcacgag tgaaacgagc cgagaagaaa 2760

tggcgagaca aacgtgagaa gctgcaactg gagaccaaca tcgtgtacaa agaggccaaa 2820

gagtcagttg acgccctgtt tgtcaatagc cagtatgacc gactgcaagt tgacaccaac 2880

atcgccatga tccacgctgc ggacaagcgc gtccaccgca tccgcgaggc ttatctgccc 2940

gagctgagcg tcattcccgg cgtcaatgcc gcgatcttcg aggagttaga gggccgcatc 3000

ttcaccgcct acagcctcta tgacgcccgc aatgtcatta agaatggcga cttcaacaat 3060

ggcttactat gctggaatgt caaagggcac gttgacgtcg aggagcagaa caatcaccgc 3120

agcgtcttag tcatacccga gtgggaggcc gaagtcagcc aggaagtccg cgtctgtcca 3180

gggcgcgggt acatcctgcg ggtcaccgcc tacaaagagg gatacggcga gggttgtgtc 3240

accatacacg agatagagga caataccgac gaactcaagt tcagcaattg tgtcgaggag 3300

gaagtctatc ccaacaatac cgtaacctgc aacaactaca ccggaaccca ggaggagtat 3360

gaagggacgt acacctcgcg gaaccagggc tatgacgaag cctatgggaa caacccgtcg 3420

gtgcctgctg actatgcgtc ggtctatgag gagaaatcgt acacggacgg gcggcgggag 3480

aatccgtgtg agtcgaatcg cgggtatggt gactacacgc cgctaccggc gggctatgta 3540

acgaaagacc tggaatactt cccggagacg gacaaagtat ggatagagat aggcgagacg 3600

gagggaacgt tcatcgtgga ctcggtagag ctgctgctca tggaggagtg a 3651

<210> 25
<211> 1216
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC867_25.

<400> 25

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

```
Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480
```

```
Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

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

Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly
        515                 520                 525

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

Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

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

Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly Ala Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Asp Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln
                645                 650                 655

Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys
            660                 665                 670

Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Asp
        675                 680                 685

Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu
    690                 695                 700

Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
705                 710                 715                 720

Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg
                725                 730                 735

Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu
            740                 745                 750

Asn Tyr Val Thr Leu Pro Gly Thr Val Asp Glu Cys Tyr Pro Thr Tyr
        755                 760                 765

Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr
    770                 775                 780

Glu Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu
785                 790                 795                 800

Ile Arg Tyr Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly
```

```
                805                 810                 815

Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu
            820                 825                 830

Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys
        835                 840                 845

Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Thr
    850                 855                 860

Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val
865                 870                 875                 880

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

Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala
            900                 905                 910

Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu
        915                 920                 925

Gln Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp
    930                 935                 940

Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Val Asp Thr Asn
945                 950                 955                 960

Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu
                965                 970                 975

Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile
            980                 985                 990

Phe Glu Glu Leu Glu Gly Arg Ile  Phe Thr Ala Tyr Ser  Leu Tyr Asp
        995                 1000                1005

Ala Arg  Asn Val Ile Lys Asn  Gly Asp Phe Asn Asn  Gly Leu Leu
    1010                1015                1020

Cys Trp  Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
    1025                1030                1035

His Arg  Ser Val Leu Val Ile  Pro Glu Trp Glu Ala  Glu Val Ser
    1040                1045                1050

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1055                1060                1065

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1070                1075                1080

Glu Ile  Glu Asp Asn Thr Asp  Glu Leu Lys Phe Ser  Asn Cys Val
    1085                1090                1095

Glu Glu  Glu Val Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asn Tyr
    1100                1105                1110

Thr Gly  Thr Gln Glu Glu Tyr  Glu Gly Thr Tyr Thr  Ser Arg Asn
    1115                1120                1125
```

Gln Gly Tyr Asp Glu Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala
1130 1135 1140

Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg
1145 1150 1155

Arg Glu Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr
1160 1165 1170

Pro Leu Pro Ala Gly Tyr Val Thr Lys Asp Leu Glu Tyr Phe Pro
1175 1180 1185

Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr
1190 1195 1200

Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1205 1210 1215

<210> 26
<211> 3600
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC868.

<400> 26
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta atcaaatacc tttagtgaaa ggatttagag tttgggggggg cacctctgtc 1560

attacaggac caggatttac aggaggggat atccttcgaa gaaatacctt tggtgatttt 1620

gtatctctac aagtcaatat taattcacca attacccaaa gataccgttt aagatttcgt 1680

tacgcttcca gtagggatgc acgagttata gtattaacag gagcggcatc cacaggagtg 1740

ggaggccaag ttagtgtaaa tatgcctctt cagaaaacta tggaaatagg ggagaactta 1800

acatctagaa catttagata taccgatttt agtaatcctt tttcatttag agctaatcca 1860

gatataattg ggataagtga acaacctcta tttggtgcag gttctattag tagcggtgaa 1920

ctttatatag ataaaattga aattattcta gcagatgcaa catttgaagc agaatctgat 1980

ttagaaagag cacaaaaggc ggtgaatgag ctgtttactt cttccaatca aatcgggtta 2040

aaaacagatg tgacggatta tcatattgat caagtatcca atttagttga gtgtttatct 2100

gatgaatttt gtctggatga aaaaaaagaa ttgtccgaga aagtcaaaca tgcgaagcga 2160

cttagtgatg agcggaattt acttcaagat ccaaacttta gagggatcaa tagacaacta 2220

gaccgtggct ggagaggaag tacggatatt accatccaag gaggcgatga cgtattcaaa 2280

gagaattacg ttacgctatt gggtaccttt gatgagtgct atccaacgta tttatatcaa 2340

aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa 2400

gatagtcaag acttagaaat ctatttaatt cgctacaatg ccaaacacga aacagtaaat 2460

gtgccaggta cgggttcctt atggccgctt tcagccccaa gtccaatcgg aaaatgtgcc 2520

catcattccc atcatttctc cttggacatt gatgttggat gtacagactt aaatgaggac 2580

ttaggtgtat gggtgatatt caagattaag acgcaagatg gccatgcaag actaggaaat 2640

ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaaagagcg 2700

gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa 2760

gaggcaaaag aatctgtaga tgctttattt gtaaactctc aatatgatag attacaagcg 2820

gataccaaca tcgcgatgat tcatgcggca gataaacgcg ttcatagcat tcgagaagct 2880

tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatttttga agaattagaa 2940

gggcgtattt tcactgcatt ctccctatat gatgcgagaa atgtcattaa aaatggtgat 3000

tttaataatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac 3060

aaccaccgtt cggtccttgt tgttccggaa tgggaagcag aagtgtcaca agaagttcgt 3120

gtctgtccgg gtcgtggcta tatccttcgt gtcacagcgt acaaggaggg atatggagaa 3180

ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt 3240

gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa 3300

gaagaatatg agggtacgta cacttctcgt aatcgaggat atgacggagc ctatgaaagc 3360

aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga 3420

cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct 3480

ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggtatg gattgagatc 3540

ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag 3600

<210> 27
<211> 3600
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868.

<400> 27
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgttcgaggc cgagtctgac 1980

ctggagcggg ctcagaaggc tgtcaacgaa ctgttcacca gcagcaacca gattgggctc 2040

aagaccgacg tcacggacta tcacattgac caagtgtcca accttgtgga gtgcctgtcc 2100

gacgagttct gcctcgacga gaagaaggag ctgtccgaga aggtcaaaca cgcgaagcgt 2160

ctgagtgacg agcggaattt gctccaggac ccgaacttcc gtggcatcaa ccgccagctc 2220

gaccgtggtt ggcgcgggag tacagacatc accatccagg gaggcgacga tgtgttcaag 2280

gagaactatg tgacgctgct cgggactttc gacgaatgct acccgacgta tctctaccag 2340

aagatagacg agagtaaatt gaaggcgtac acccgctacc agcttcgcgg gtacatcgag 2400

gatagtcagg acctggaaat ctacctgatc cgatacaacg ccaagcacga gacagtgaac 2460

gtgccaggca cgggctcact ttggccattg agcgctccct ctccaatcgg aaagtgcgct 2520

caccactcgc accacttctc tctggacatc gacgtgggct gcaccgacct caacgaggac 2580

ctgggtgtct gggttatctt caagattaag acccaggacg gacatgcccg cctcggcaac 2640

ctggagttcc ttgaggagaa gcctctcgtg ggcgaggccc tcgctcgtgt gaagcgcgcc 2700

gagaagaaat ggcgagacaa gcgggagaag ctggagtggg agaccaacat cgtgtacaag 2760

gaggccaagg agtcagtgga cgcactcttc gtcaacagcc agtacgaccg cctccaggct 2820

gacaccaaca tcgccatgat ccacgcggct gacaagcggg tccacagcat ccgtgaggcg 2880

tacctgcccg agctgtcagt gatccctggt gtgaacgcgg cgatcttcga ggaactggag 2940

ggccgcatct tcacagcatt cagcctgtac gatgccagga atgttattaa gaacggtgac 3000

ttcaacaacg ggctgagttg ctggaacgtc aagggccatg tggacgtcga ggagcagaac 3060

aaccaccggt ccgtgctggt cgtgccggag tgggaggcag aggtgagcca ggaggtccgc 3120

gtctgccctg gtcgcggcta catcctccgt gtgactgcgt acaaggaagg ctacggtgaa 3180

ggctgcgtga ctatccacga gatcgagaac aacaccgacg agctcaagtt ctcgaactgt 3240

gtggaggagg aggtgtaccc gaacaacacc gttacttgca acgactacac tgccacgcaa 3300

gaggagtacg agggcactta cacttcccgg aatcgcggct atgatggcgc gtacgagtcc 3360

aacagcagcg tgcctgcgga ttatgcgtcc gcttacgagg agaaggcgta caccgacgga 3420

cggagggaca acccttgcga gtccaaccgt ggctacggtg actacactcc gctgcccgcc 3480

gggtacgtca ccaaggagct ggagtacttc ccggagaccg acaaagtctg gatcgagatc 3540

ggcgagacgg agggcacttt catcgtggac tcggtcgagc tgctactgat ggaggagtga 3600

<210> 28
<211> 1199
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC868.

<400> 28

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro

```
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415
```

```
Ala Arg Phe Asn Trp Arg Asn Pro Leu Asn Ser Leu Arg Gly Ser Leu
            420                 425                 430

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
                645                 650                 655

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

Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His
        675                 680                 685

Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys
    690                 695                 700

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

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

```
Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile
            740                 745                 750

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

Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu
    770                 775                 780

Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu
785                 790                 795                 800

Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His
                805                 810                 815

Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala
            820                 825                 830

Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu
        835                 840                 845

Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp
    850                 855                 860

Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn
865                 870                 875                 880

Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg
                885                 890                 895

Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu
            900                 905                 910

Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala
        915                 920                 925

Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile
    930                 935                 940

Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala
945                 950                 955                 960

Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe
                965                 970                 975

Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala
            980                 985                 990

Arg Asn Val Ile Lys Asn Gly Asp  Phe Asn Asn Gly Leu  Ser Cys Trp
        995                 1000                1005

Asn Val  Lys Gly His Val Asp  Val Glu Glu Gln Asn  Asn His Arg
    1010                1015                1020

Ser Val  Leu Val Val Pro Glu  Trp Glu Ala Glu Val  Ser Gln Glu
    1025                1030                1035

Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile Leu Arg  Val Thr Ala
    1040                1045                1050

Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val Thr Ile  His Glu Ile
```

1055 1060 1065

Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu
1070 1075 1080

Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala
1085 1090 1095

Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly
1100 1105 1110

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr
1115 1120 1125

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp
1130 1135 1140

Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu
1145 1150 1155

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr
1160 1165 1170

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile
1175 1180 1185

Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1190 1195

<210> 29
<211> 3600
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_9.

<400> 29
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

agcctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgcag 1200

gcgggcatta acatccttat gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgaagaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgttcgaggc cgagtctgac 1980

ctggagcggg ctcagaaggc tgtcaacgaa ctgttcacca gcagcaacca gattgggctc 2040

aagaccgacg tcacggacta tcacattgac caagtgtcca accttgtgga gtgcctgtcc 2100

gacgagttct gcctcgacga gaagaaggag ctgtccgaga aggtcaaaca cgcgaagcgt 2160

ctgagtgacg agcggaattt gctccaggac ccgaacttcc gtggcatcaa ccgccagctc 2220

gaccgtggtt ggcgcgggag tacagacatc accatccagg gaggcgacga tgtgttcaag 2280

gagaactatg tgacgctgct cgggactttc gacgaatgct acccgacgta tctctaccag 2340

aagatagacg agagtaaatt gaaggcgtac acccgctacc agcttcgcgg gtacatcgag 2400

gatagtcagg acctggaaat ctacctgatc cgatacaacg ccaagcacga gacagtgaac 2460

gtgccaggca cgggctcact ttggccattg agcgctccct ctccaatcgg aaagtgcgct 2520

caccactcgc accacttctc tctggacatc gacgtgggct gcaccgacct caacgaggac 2580

ctgggtgtct gggttatctt caagattaag acccaggacg gacatgcccg cctcggcaac 2640

ctggagttcc ttgaggagaa gcctctcgtg ggcgaggccc tcgctcgtgt gaagcgcgcc 2700

gagaagaaat ggcgagacaa gcgggagaag ctggagtggg agaccaacat cgtgtacaag 2760

gaggccaagg agtcagtgga cgcactcttc gtcaacagcc agtacgaccg cctccaggct 2820

gacaccaaca tcgccatgat ccacgcggct gacaagcggg tccacagcat ccgtgaggcg 2880

tacctgcccg agctgtcagt gatccctggt gtgaacgcgg cgatcttcga ggaactggag 2940

ggccgcatct tcacagcatt cagcctgtac gatgccagga atgttattaa gaacggtgac 3000

ttcaacaacg ggctgagttg ctggaacgtc aagggccatg tggacgtcga ggagcagaac 3060

aaccaccggt ccgtgctggt cgtgccggag tgggaggcag aggtgagcca ggaggtccgc 3120

gtctgccctg gtcgcggcta catcctccgt gtgactgcgt acaaggaagg ctacggtgaa 3180

ggctgcgtga ctatccacga gatcgagaac aacaccgacg agctcaagtt ctcgaactgt 3240

gtggaggagg aggtgtaccc gaacaacacc gttacttgca acgactacac tgccacgcaa 3300

gaggagtacg agggcactta cacttcccgg aatcgcggct atgatggcgc gtacgagtcc 3360

aacagcagcg tgcctgcgga ttatgcgtcc gcttacgagg agaaggcgta caccgacgga 3420

cggagggaca acccttgcga gtccaaccgt ggctacggtg actacactcc gctgcccgcc 3480

gggtacgtca ccaaggagct ggagtacttc ccggagaccg acaaagtctg gatcgagatc 3540

ggcgagacgg agggcacttt catcgtggac tcggtcgagc tgctactgat ggaggagtga 3600

<210> 30
<211> 1199
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_9.

<400> 30

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

```
Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Ser Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Gln
385                 390                 395                 400
```

```
Ala Gly Ile Asn Ile Leu Met Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

Ala Arg Phe Asn Trp Arg Asn Pro Lys Asn Ser Leu Arg Gly Ser Leu
            420                 425                 430

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
                645                 650                 655

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

Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His
        675                 680                 685

Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys
    690                 695                 700

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

Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile
```

```
                725                 730                 735

Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile
            740                 745                 750

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

Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu
    770                 775                 780

Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu
785                 790                 795                 800

Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His
                805                 810                 815

Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala
            820                 825                 830

Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu
        835                 840                 845

Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp
    850                 855                 860

Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn
865                 870                 875                 880

Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg
                885                 890                 895

Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu
            900                 905                 910

Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala
        915                 920                 925

Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile
    930                 935                 940

Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala
945                 950                 955                 960

Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe
                965                 970                 975

Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala
            980                 985                 990

Arg Asn Val Ile Lys Asn Gly Asp  Phe Asn Asn Gly Leu  Ser Cys Trp
        995                 1000                1005

Asn Val  Lys Gly His Val Asp  Val Glu Glu Gln Asn  Asn His Arg
    1010                1015                1020

Ser Val  Leu Val Val Pro Glu  Trp Glu Ala Glu Val  Ser Gln Glu
    1025                1030                1035

Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile Leu Arg  Val Thr Ala
    1040                1045                1050
```

Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile
1055 1060 1065

Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu
1070 1075 1080

Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala
1085 1090 1095

Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly
1100 1105 1110

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr
1115 1120 1125

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp
1130 1135 1140

Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu
1145 1150 1155

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr
1160 1165 1170

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile
1175 1180 1185

Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1190 1195

<210> 31
<211> 3678
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC868_10.

<400> 31
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

cttttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta atcaaatacc tttagtgaaa ggatttagag tttgggggggg cacctctgtc 1560

attacaggac caggatttac aggaggggat atccttcgaa gaaatacctt tggtgatttt 1620

gtatctctac aagtcaatat taattcacca attacccaaa gataccgttt aagatttcgt 1680

tacgcttcca gtagggatgc acgagttata gtattaacag gagcggcatc cacaggagtg 1740

ggaggccaag ttagtgtaaa tatgcctctt cagaaaacta tggaaatagg ggagaactta 1800

acatctagaa catttagata taccgatttt agtaatcctt tttcatttag agctaatcca 1860

gatataattg ggataagtga acaacctcta tttggtgcag gttctattag tagcggtgaa 1920

ctttatatag ataaaattga aattattcta gcagatgcaa catttgaggc agaatatgat 1980

ttagaaagag cgcaaaaggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta 2040

aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg 2100

gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga 2160

ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca 2220

gaccgtggct ggagaggaag tacggatatt actatccaag gaggagatga cgtattcaaa 2280

gagaattacg ttacgctacc gggtaccttt gatgagtgct atccaacgta tttatatcaa 2340

aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa 2400

gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaaat 2460

gtaccaggta caggaagttt atggcctctt tctgtagaaa atcaaattgg accttgtgga 2520

gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacactg ttcctgcaga 2580

gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttggatgt 2640

acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc 2700

cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta 2760

gctcgtgtga aaagagcgga gaaaaaatgg agagacaaac gcgaaacatt acaattggaa 2820

acaactatcg tttataaaga ggcaaaagaa tctgtagatg ctttatttgt aaactctcaa 2880

tatgatagat tacaagcgga tacgaacatc gcgatgattc atgcggcaga taaacgcgtt 2940

catagaattc gagaagcgta tctgccggag ctgtctgtga ttccgggtgt caatgcggct 3000

atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat 3060

attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta 3120

gaggtagaag aacaaaacaa tcaccgttca gtcctggtta tcccagaatg ggaggcagaa 3180

gtgtcacaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac 3240

aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagacgaa 3300

ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggt aacgtgtatt 3360

aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat 3420

gacgaagcct atggtaataa cccttccgta ccagctgatt atgcgtcagt ctatgaagaa 3480

aaatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat 3540

tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat 3600

aaggtatgga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta 3660

ctccttatgg aggaatag 3678

<210> 32
<211> 3678
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_10.

<400> 32
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgttcgaggc cgagtacgac 1980

cttgagcgcg cccagaaggt ggtgaacgcc ctcttcacta gcactaacca gctaggcctg 2040

aagactgacg tgaccgacta ccacatcgac caagtgagca acctagtggc ctgcctctcc 2100

gacgagttct gcctcgacga gaagcgcgag ctgtccgaga aggtgaagca cgccaagcgc 2160

ctctccgacg agcgcaacct gctccaggac cccaacttca ggggcatcaa caggcagccc 2220

gaccgcggct ggcgcggctc caccgacatc accatccagg gcggtgacga cgtattcaag 2280

gagaactacg ttaccctccc cggcaccttc gacgagtgtt accccaccta cctctaccag 2340

aagatcgacg agtccaagct gaaggcctac acccgctacc agctccgcgg ctacatcgag 2400

gactcccagg acctggaaat ctacctcatc cgctacaacg ccaagcacga gatcgtgaac 2460

gtgcctggca ccggcagcct ctggcctctc agcgtggaga accagatcgg cccttgcggc 2520

gagcctaacc gctgcgcccc tcacctcgag tggaaccctg acctccactg ctcgtgcagg 2580

gacggcgaga agtgcgccca ccatagccac cacttctctc tggacatcga cgtgggctgc 2640

accgacctga acgaggacct gggcgtgtgg gttatcttca agatcaagac ccaggacggt 2700

cacgccaggc tgggtaacct ggagttcctt gaggaaaagc ctctgctggg tgaggccctg 2760

gccagggtca agagggctga gaagaaatgg agggataaga gggagaccct gcagctggag 2820

accactatcg tctacaagga ggctaaggag tctgtcgatg ctctgttcgt caactctcag 2880

tacgatagac tgcaagctga taccaacatc gctatgatcc acgctgcgga taagcgggtc 2940

caccggatcc gggaggctta ccttccggag ctttctgtca tcccgggtgt caacgctgcg 3000

atcttcgagg aacttgagga acggatcttc actgcgttta gtctttacga tgcgcggaac 3060

atcatcaaga acggggactt caacaatggt ctgctgtgct ggaacgtcaa gggtcatgtc 3120

gaggtcgagg aacaaaacaa tcatcgtagt gtccttgtca ttcctgagtg ggaggcggag 3180

gtctctcaag aggtccgtgt ttgcccgggg cgtgggtaca ttcttcgtgt tactgcgtac 3240

aaggaggggt acggggaggg gtgcgttact attcatgaga ttgagaacaa tactgatgag 3300

cttaagttca acaattgtgt tgaggaggag gtttacccga acaatactgt tacgtgcatc 3360

aactacacgg caacgcaaga ggaatacgag gggacgtaca cctcgcgtaa tagagggtat 3420

gatgaggcgt acggaaacaa cccgtcggtt ccagcagatt atgcctcggt ttatgaggag 3480

aagtcgtaca cggatagacg acgcgagaat ccatgtgagt caaatcgagg atacggagat 3540

tacacaccat taccagcagg atacgttaca aaggagttgg aatacttccc ggaaacagat 3600

aaagtttgga ttgaaatcgg agaaacagaa ggaacattca tcgtcgactc agtagaattg 3660

ttgttgatgg aagaatga 3678

<210> 33
<211> 1225
<212> PRT
<213> Nhân tạo

<220>

<223> Trình tự axit amin của biến thể protein khảm TIC868_10.

<400> 33

```
Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1               5                  10                  15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
            20                  25                  30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
        35                  40                  45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300
```

```
Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
```

625 630 635 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
645 650 655

Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val Asn Ala Leu Phe
660 665 670

Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val Thr Asp Tyr His
675 680 685

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

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

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

Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile
740 745 750

Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly
755 760 765

Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu
770 775 780

Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu
785 790 795 800

Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His
805 810 815

Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Val
820 825 830

Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg Cys Ala Pro His
835 840 845

Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp Gly Glu Lys
850 855 860

Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys
865 870 875 880

Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys
885 890 895

Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu
900 905 910

Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys
915 920 925

Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu Thr Thr Ile Val
930 935 940

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

```
Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala
                965                 970                 975

Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser
            980                 985                 990

Val Ile Pro Gly Val Asn Ala Ala  Ile Phe Glu Glu Leu  Glu Glu Arg
        995                 1000                1005

Ile Phe  Thr Ala Phe Ser Leu  Tyr Asp Ala Arg Asn  Ile Ile Lys
    1010                1015                1020

Asn Gly  Asp Phe Asn Asn Gly  Leu Leu Cys Trp Asn  Val Lys Gly
    1025                1030                1035

His Val  Glu Val Glu Glu Gln  Asn Asn His Arg Ser  Val Leu Val
    1040                1045                1050

Ile Pro  Glu Trp Glu Ala Glu  Val Ser Gln Glu Val  Arg Val Cys
    1055                1060                1065

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

Tyr Gly  Glu Gly Cys Val Thr  Ile His Glu Ile Glu  Asn Asn Thr
    1085                1090                1095

Asp Glu  Leu Lys Phe Asn Asn  Cys Val Glu Glu Glu  Val Tyr Pro
    1100                1105                1110

Asn Asn  Thr Val Thr Cys Ile  Asn Tyr Thr Ala Thr  Gln Glu Glu
    1115                1120                1125

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

Tyr Gly  Asn Asn Pro Ser Val  Pro Ala Asp Tyr Ala  Ser Val Tyr
    1145                1150                1155

Glu Glu  Lys Ser Tyr Thr Asp  Arg Arg Arg Glu Asn  Pro Cys Glu
    1160                1165                1170

Ser Asn  Arg Gly Tyr Gly Asp  Tyr Thr Pro Leu Pro  Ala Gly Tyr
    1175                1180                1185

Val Thr  Lys Glu Leu Glu Tyr  Phe Pro Glu Thr Asp  Lys Val Trp
    1190                1195                1200

Ile Glu  Ile Gly Glu Thr Glu  Gly Thr Phe Ile Val  Asp Ser Val
    1205                1210                1215

Glu Leu  Leu Leu Met Glu Glu
    1220                1225
```

<210> 34
<211> 3726
<212> ADN
<213> Nhân tạo

<220>

<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC868_11.

<400> 34

atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta atcaaatacc tttagtgaaa ggatttagag tttgggggggg cacctctgtc 1560

attacaggac caggatttac aggaggggat atccttcgaa gaaatacctt tggtgatttt 1620

gtatctctac aagtcaatat taattcacca attacccaaa gataccgttt aagatttcgt 1680

tacgcttcca gtagggatgc acgagttata gtattaacag gagcggcatc cacaggagtg 1740

ggaggccaag ttagtgtaaa tatgcctctt cagaaaacta tggaaatagg ggagaactta 1800

acatctagaa catttagata taccgatttt agtaatcctt tttcatttag agctaatcca 1860

gatataattg ggataagtga acaacctcta tttggtgcag gttctattag tagcggtgaa 1920

ctttatatag ataaaattga aattattcta gcagatgcaa caggaacgac aacctatgag 1980

tatgaagaga agcagaatct agaaaaagcg cagaaagcgt tgaacgcttt gtttacggat 2040

ggcacgaatg gctatctaca aatggatgcc actgattatg atatcaatca aactgcaaac 2100

ttaatagaat gtgtatcaga tgaattgtat gcaaaagaaa agatagtttt attagatgaa 2160

gtcaaatatg cgaagcggct tagcatatca cgtaacctac ttttgaacga tgatttagaa 2220

ttttcagatg gatttggaga aaacggatgg acgacaagtg ataatatttc aatccaggcg 2280

gataatcccc tttttaaggg gaattattta aaaatgtttg gggcaagaga tattgatgga 2340

accctatttc caacttatct ctatcaaaaa atagatgagt ccaggttaaa accatataca 2400

cgttatcgag taagagggtt tgtgggaagt agtaaaaatc taaaattagt ggtaacacgc 2460

tatgagaaag aaattgatgc cattatgaat gttccaaatg atttggcaca tatgcagctt 2520

aacccttcat gtggagatta tcgctgtgaa tcatcgtccc agtttttggt gaaccaagtg 2580

catcctacac caacagctgg atatgctctt gatatgtatg catgcccgtc aagttcagat 2640

aaaaaacata ttatgtgtca cgatcgtcat ccatttgatt ttcatattga caccggagaa 2700

ttaaatccaa acacaaacct gggtattgat gtcttgttta aaatttctaa tccaaatgga 2760

tacgctacat tagggaatct agaagtcatt gaagaaggac cactaacaga tgaagcattg 2820

gtacatgtaa aacaaaagga aaagaaatgg cgtcagcaca tggagaaaaa acgaatggaa 2880

acacaacaag cctatgatcc agcaaaacaa gctgtagatg cattatttac aaatgaacaa 2940

gagttagact atcatactac tttagatcat attcagaacg ccgatcagct ggtacaggcg 3000

attccctatg tacaccatgc ttggttaccg gatgctccag gtatgaacta tgatgtatat 3060

caagggttaa acgcacgtat catgcaggcg tacaatttat atgatgcacg aaatgtcata 3120

ataaatggtg actttacaca aggactacaa ggatggcacg caacaggaaa agcagcggta 3180

caacaaatag atggagcttc agtattagtt ctatcaaact ggagtgccga ggtatctcag 3240

aatctgcatg cccaagatca tcatggatat atgttacgtg tgattgccaa aaaagaaggt 3300

cctggaaaag ggtatgtaat gatgatggat tttaatggaa agcaggaaac acttacgttc 3360

acttcttgtg aagaaggata tataacaaaa acaatagagg tattcccgga aagtgatcga 3420

atacgaattg aaatgggaga aacagagggt acgttttatg tagatagcat cgagttgctt 3480

tgtatgcaag gatatgctag cgataataac ccgcacacgg gtaatatgta tgagcaaagt 3540

tataatggaa attataatca aaatactagc gatgtgtatc accaaggata tataaacaac 3600

tataaccaaa attctagtag tatgtataat caaaattata ttaacaatga tgacctgcat 3660

tccggttgca catgtaacca agggcataac tctggctgta catgtaatca aggatataac 3720

cgttag 3726

<210> 35
<211> 3726
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_11.

<400> 35
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cggggactac cacctacgag 1980

tacgaggaga agcagaatct cgagaaggct cagaaggctc tgaacgctct gttcactgac 2040

gggaccaacg gctacctcca gatggacgcc actgactacg acatcaacca gacagctaac 2100

ctgattgagt gtgtgagtga cgaactgtac gctaaggaga agatcgtact cctggacgag 2160

gtgaagtacg ctaagcgcct gagcattagc cgtaacctgc tgctgaacga cgatctggag 2220

ttcagcgacg gctttggcga gaacggctgg accaccagcg acaacatctc catccaggcc 2280

gacaatccac tcttcaaagg caactacctc aagatgttcg gagccaggga catcgacggc 2340

accctctttc cgacctacct ctaccagaag atcgacgagt cccgcctcaa accctacacc 2400

cgctacaggg tgcgcggctt cgtgggcagc agcaagaacc tcaagctcgt ggtcacacgg 2460

tatgagaagg agatcgacgc catcatgaac gtgcccaacg atctcgccca catgcagctc 2520

aatccatcct gcggcgacta ccggtgcgag tccagctccc agttcctcgt gaaccaggtg 2580

cacctactc cgaccgctgg ctatgccctg gacatgtacg cctgccctag ttcctccgac 2640

aagaagcaca tcatgtgcca cgaccgtcat ccgttcgact tccacatcga caccggcgaa 2700

ctgaacccga acaccaacct gggcatcgac gtactgttca agatttccaa cccgaacggg 2760

tacgccacct tgggcaacct ggaggtcatc gaagaaggcc cgctgaccga cgaggccctg 2820

gtccacgtca aacagaagga gaagaagtgg cggcagcaca tggagaagaa gcggatggag 2880

actcaacaag cctacgaccc ggccaagcaa gctgtggacg ctctgttcac caacgagcaa 2940

gagcttgact accacactac tcttgaccac atccagaatg ctgaccagct tgtccaggct 3000

attccgtacg tccaccacgc ttggctaccg gacgctccag ggatgaacta cgatgtgtac 3060

cagggtctga acgcgcggat catgcaagcg tacaacctgt acgacgcgcg taacgtcatc 3120

atcaacggtg acttcactca gggtcttcaa ggttggcacg cgactggcaa agcggcagtc 3180

cagcagattg atggtgcgtc tgttcttgtg ttgagcaact ggtctgcgga ggtttctcag 3240

aacctgcacg cacaggatca ccacggctac atgctgaggg tgattgctaa gaaggagggc 3300

cctggcaaag gctacgtcat gatgatggac ttcaacggaa agcaagaaac cctgaccttc 3360

actagctgtg aggagggcta catcactaag accattgagg tctttccgga gtctgaccgc 3420

atccggatcg agatgggcga gaccgaaggc acgttctacg tggactccat cgaactcctc 3480

tgcatgcaag gctacgcctc cgacaacaac ccacacacgg gcaacatgta cgagcagtcc 3540

tacaacggga actacaacca gaacacctcc gatgtgtacc atcagggcta catcaacaac 3600

tacaaccaga acagcagcag catgtacaac cagaactaca tcaacaacga tgacttgcac 3660

tcgggttgca cctgcaacca gggtcacaac agtgggtgca cgtgcaacca gggatacaac 3720

cgttga 3726

<210> 36
<211> 1241
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_11.

<400> 36

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

```
Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480
```

```
Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Gly Thr
                645                 650                 655

Thr Thr Tyr Glu Tyr Glu Glu Lys Gln Asn Leu Glu Lys Ala Gln Lys
            660                 665                 670

Ala Leu Asn Ala Leu Phe Thr Asp Gly Thr Asn Gly Tyr Leu Gln Met
        675                 680                 685

Asp Ala Thr Asp Tyr Asp Ile Asn Gln Thr Ala Asn Leu Ile Glu Cys
    690                 695                 700

Val Ser Asp Glu Leu Tyr Ala Lys Glu Lys Ile Val Leu Leu Asp Glu
705                 710                 715                 720

Val Lys Tyr Ala Lys Arg Leu Ser Ile Ser Arg Asn Leu Leu Leu Asn
                725                 730                 735

Asp Asp Leu Glu Phe Ser Asp Gly Phe Gly Glu Asn Gly Trp Thr Thr
            740                 745                 750

Ser Asp Asn Ile Ser Ile Gln Ala Asp Asn Pro Leu Phe Lys Gly Asn
        755                 760                 765

Tyr Leu Lys Met Phe Gly Ala Arg Asp Ile Asp Gly Thr Leu Phe Pro
    770                 775                 780

Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Arg Leu Lys Pro Tyr Thr
785                 790                 795                 800

Arg Tyr Arg Val Arg Gly Phe Val Gly Ser Ser Lys Asn Leu Lys Leu
```

```
                805                 810                 815

Val Val Thr Arg Tyr Glu Lys Glu Ile Asp Ala Ile Met Asn Val Pro
            820                 825                 830

Asn Asp Leu Ala His Met Gln Leu Asn Pro Ser Cys Gly Asp Tyr Arg
        835                 840                 845

Cys Glu Ser Ser Ser Gln Phe Leu Val Asn Gln Val His Pro Thr Pro
    850                 855                 860

Thr Ala Gly Tyr Ala Leu Asp Met Tyr Ala Cys Pro Ser Ser Ser Asp
865                 870                 875                 880

Lys Lys His Ile Met Cys His Asp Arg His Pro Phe Asp Phe His Ile
                885                 890                 895

Asp Thr Gly Glu Leu Asn Pro Asn Thr Asn Leu Gly Ile Asp Val Leu
            900                 905                 910

Phe Lys Ile Ser Asn Pro Asn Gly Tyr Ala Thr Leu Gly Asn Leu Glu
        915                 920                 925

Val Ile Glu Glu Gly Pro Leu Thr Asp Glu Ala Leu Val His Val Lys
    930                 935                 940

Gln Lys Glu Lys Lys Trp Arg Gln His Met Glu Lys Lys Arg Met Glu
945                 950                 955                 960

Thr Gln Gln Ala Tyr Asp Pro Ala Lys Gln Ala Val Asp Ala Leu Phe
                965                 970                 975

Thr Asn Glu Gln Glu Leu Asp Tyr His Thr Thr Leu Asp His Ile Gln
            980                 985                 990

Asn Ala Asp Gln Leu Val Gln Ala  Ile Pro Tyr Val His  His Ala Trp
        995                 1000                1005

Leu Pro  Asp Ala Pro Gly Met  Asn Tyr Asp Val Tyr  Gln Gly Leu
    1010                1015                1020

Asn Ala  Arg Ile Met Gln Ala  Tyr Asn Leu Tyr Asp  Ala Arg Asn
    1025                1030                1035

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

Ala Thr  Gly Lys Ala Ala Val  Gln Gln Ile Asp Gly  Ala Ser Val
    1055                1060                1065

Leu Val  Leu Ser Asn Trp Ser  Ala Glu Val Ser Gln  Asn Leu His
    1070                1075                1080

Ala Gln  Asp His His Gly Tyr  Met Leu Arg Val Ile  Ala Lys Lys
    1085                1090                1095

Glu Gly  Pro Gly Lys Gly Tyr  Val Met Met Met Asp  Phe Asn Gly
    1100                1105                1110

Lys Gln  Glu Thr Leu Thr Phe  Thr Ser Cys Glu Glu  Gly Tyr Ile
    1115                1120                1125
```

Thr Lys Thr Ile Glu Val Phe Pro Glu Ser Asp Arg Ile Arg Ile
1130 1135 1140

Glu Met Gly Glu Thr Glu Gly Thr Phe Tyr Val Asp Ser Ile Glu
1145 1150 1155

Leu Leu Cys Met Gln Gly Tyr Ala Ser Asp Asn Asn Pro His Thr
1160 1165 1170

Gly Asn Met Tyr Glu Gln Ser Tyr Asn Gly Asn Tyr Asn Gln Asn
1175 1180 1185

Thr Ser Asp Val Tyr His Gln Gly Tyr Ile Asn Asn Tyr Asn Gln
1190 1195 1200

Asn Ser Ser Ser Met Tyr Asn Gln Asn Tyr Ile Asn Asn Asp Asp
1205 1210 1215

Leu His Ser Gly Cys Thr Cys Asn Gln Gly His Asn Ser Gly Cys
1220 1225 1230

Thr Cys Asn Gln Gly Tyr Asn Arg
1235 1240

<210> 37
<211> 3468
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC868_12.

<400> 37
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt 120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt 180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt 240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc 300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct 360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat 420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc 480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct 600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa 660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat 720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta 780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca 840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta atcaaatacc tttagtgaaa ggatttagag tttgggggggg cacctctgtc 1560

attacaggac caggatttac aggaggggat atccttcgaa gaaatacctt tggtgatttt 1620

gtatctctac aagtcaatat taattcacca attacccaaa gataccgttt aagatttcgt 1680

tacgcttcca gtagggatgc acgagttata gtattaacag gagcggcatc cacaggagtg 1740

ggaggccaag ttagtgtaaa tatgcctctt cagaaaacta tggaaatagg ggagaactta 1800

acatctagaa catttagata taccgatttt agtaatcctt tttcatttag agctaatcca 1860

gatataattg ggataagtga acaacctcta tttggtgcag gttctattag tagcggtgaa 1920

ctttatatag ataaaattga aattattcta gcagatgcaa caaatccgac gcgagaggcg 1980

gaagaggatc tagaagcagc gaagaaagcg gtggcgagct tgtttacacg tacaagggac 2040

ggattacaag taaatgtgac agattatcaa gtcgatcaag cggcaaattt agtgtcatgc 2100

ttatcagatg aacaatatgg gcatgacaaa aagatgttat tggaagcggt aagagcggca 2160

aaacgcctca gccgagaacg caacttactt caggatccag attttaatac aatcaatagt 2220

acagaagaaa atggatggaa agcaagtaac ggcgttacta ttagcgaggg cggtccattc 2280

tataaaggcc gtgcgcttca gctagcaagc gcaagagaaa attacccaac atacatttat 2340

caaaaagtaa atgcatcaga gttaaagccg tatacacgtt atagactgga tgggttcgtg 2400

aagagtagtc aagatttaga aattgatctc attcaccatc ataaagtcca tctcgtgaaa 2460

aatgtaccag ataatttagt atccgatact tactcggatg gttcttgcag tggaatgaat 2520

cgatgtgagg aacaacagat ggtaaatgcg caactggaaa cagaacatca tcatccgatg 2580

gattgctgtg aagcggctca aacacatgag ttttcttcct atattaatac aggcgatcta 2640

aattcaagtg tagatcaagg catttgggtt gtattgaaag ttcgaacaac cgatggttat 2700

gcgacgctag gaaatcttga attggtagag gtcggaccgt tatcgggtga atctctagaa 2760

cgtgaacaaa gggataatgc gaaatggagt gcagagctag gaagaaagcg tgcagaaaca 2820

gatcgcgtgt atcaagatgc caaacaatcc atcaatcatt tatttgtgga ttatcaagat 2880

caacaattaa atccagaaat agggatggca gatattattg acgctcaaaa tcttgtcgca 2940

tcaatttcag atgtgtatag cgatgcagta ctgcaaatcc ctggaattaa ctatgagatt 3000

tacacagagc tatccaatcg cttacaacaa gcatcgtatc tgtatacgtc tcgaaatgcg 3060

gtgcaaaatg gggactttaa cagcggtcta gatagttgga atgcaacagg gggggctacg 3120

gtacaacagg atggcaatac gcatttctta gttctttctc attgggatgc acaagtttct 3180

caacaattta gagtgcagcc gaattgtaaa tatgtattac gtgtaacagc agagaaagta 3240

ggcggcggag acggatacgt gacaatccgg gatggtgctc atcatacaga aaagcttaca 3300

tttaatgcat gtgattatga tataaatggc acgtacgtga ctgataatac gtatctaaca 3360

aaagaagtgg tattctattc acatacagaa cacatgtggg tagaggtaag tgaaacagaa 3420

ggtgcatttc atatagatag tattgaattc gttgaaacag aaaagtag 3468

<210> 38
<211> 3468
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_12.

<400> 38
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgaacccgac gcgggaagct 1980

gaggaagact tggaagccgc caagaaagcg gtcgccagcc tgtttactcg gacgcgggac 2040

gggctccaag tgaatgtgac ggactatcaa gtggatcagg ccgctaacct cgtgtcatgc 2100

ctgagcgacg agcagtacgg tcacgacaag aaaatgctgc tggaggccgt ccgggccgcc 2160

aagcggctgt ccagggagcg taacctgcta caagatcccg actttaacac gatcaacagc 2220

acagaggaga atggctggaa ggccagcaac ggagttacga taagcgaggg cggtccgttc 2280

tacaagggtc gtgccctcca gctcgcctct gcaagggaga actatccaac ctacatctat 2340

cagaaggtga acgcatccga gcttaagccc tacacacgct accgcctgga cgggttcgtt 2400

aagtccagtc aagacctaga gatagacctc atccaccacc acaaagtgca tctggtcaag 2460

aacgttcccg ataatctcgt gagcgatacc tactcagacg gctcatgctc tggcatgaac 2520

agatgtgagg agcaacagat ggttaatgct caactcgaaa ccgagcatca tcatcctatg 2580

gattgctgcg aggccgcgca gacccatgag ttcagctctt acatcaacac cggagacctc 2640

aacagtagcg tggatcaggg aatttgggtg gtgcttaaag tgcgtacaac cgacggctac 2700

gccaccctcg gcaaccttga gcttgtcgag gtcggaccac ttagcggcga gtccctggaa 2760

cgtgagcagc gggacaacgc caaatggagc gcagagctag ggcgcaaacg cgcggagacg 2820

gaccgggttt atcaggacgc gaagcagtcc atcaatcacc tcttcgtgga ttatcaggac 2880

cagcagctta atccagagat cggcatggcc gacatcatcg acgcccagaa cctagtagcg 2940

tcgatttccg atgtctattc cgacgccgtg cttcaaatac ctggcatcaa ctacgagatc 3000

tacacagagt tgtccaacag gctccagcaa gcgtcatacc tctacaccag ccgcaacgcc 3060

gtccagaatg gcgacttcaa ttccggacta gactcctgga acgccacggg cggagctacg 3120

gtgcaacaag acggcaacac ccacttcctc gtacttagcc actgggacgc tcaagtgagt 3180

cagcaattcc gggttcagcc gaactgcaag tacgtcctgc gcgtaacggc cgagaaggtt 3240

ggaggcggag acggctacgt taccatccgc gacggcgctc accacaccga gaaactgacg 3300

ttcaacgctt gtgactacga catcaacggc acttacgtga cggacaacac ctacctgacg 3360

aaggaggtgg tgttctattc tcacaccgag cacatgtggg ttgaggtcag cgagaccgag 3420

ggagccttcc acattgacag catcgagttc gtggagactg agaagtga 3468

<210> 39
<211> 1155
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_12.

<400> 39

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

```
Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

Ala Arg Phe Asn Trp Arg Asn Pro Leu Asn Ser Leu Arg Gly Ser Leu
```

```
            420                 425                 430

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr 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> 40
<211> 3732
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_13.

<400> 40
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgacggcgac cttcgaggcg 1980

gagtatgact tggagcgggc tcaggaggcc gtcaacgcgc tgttcacaaa caccaatcct 2040

cgccgcctca agacgggtgt gactgattac cacattgacg aggtctccaa cttggtcgcg 2100

tgtctgtccg atgagttctg cctggacgag aagcgggaac tgctggagaa ggtcaagtac 2160

gccaagcgcc tctccgacga aaggaacctc ctccaagatc ccaactttac ttccattaac 2220

aagcagccgg acttcatctc caccaacgag cagtccaact tcacctcaat ccacgagcag 2280

tcggagcacg ggtggtgggg cagcgagaac atcaccatcc aagagggcaa cgacgtcttc 2340

aaggagaact acgtgatcct gcccggcacc ttcaacgagt gttacccgac ctatctctac 2400

cagaagattg gcgaagcgga actcaaggct tacacccgtt accaactgag tggctacatt 2460

gaggactcac aagacctgga aatctacctg atccgctaca acgccaagca cgagaccctc 2520

gacgtgcctg gcacggagtc cgtctggccc ttgagcgtgg agtctcctat cggtcgttgc 2580

ggcgagccca atcgctgcgc tccgcacttt gagtggaatc ctgatttgga ttgctcctgc 2640

cgagacggtg agaaatgcgc ccaccactcg caccacttca gcctagacat cgacgtgggc 2700

tgcatcgacc tgcacgagaa cttgggcgtc tgggtcgtgt tcaagatcaa gacacaggag 2760

ggccatgctc ggcttgggaa cctggagttc atcgaggaga agccactgct gggtgaagcc 2820

ttgtcacggg tgaaacgcgc cgagaagaag tggcgggaca aacgggagaa gctccagttg 2880

gagacaaagc gtgtgtacac agaggccaag gaggccgtgg atgccttgtt cgtggacagt 2940

cagtacgaca ggctgcaagc ggacaccaac atcgggatga tccacgcggc tgataagctt 3000

gttcacagaa tccgcgaggc gtacctgtca gagcttagcg tgatcccagg cgtcaacgcc 3060

gaaatcttcg aggaactgga gggccgcatt atcacggcaa tctcacttta tgacgcgagg 3120

aatgtggtca agaacggtga cttcaacaac ggcttggcgt gttggaacgt taaagggcac 3180

gtggatgtac aacagtcaca ccacagaagt gtcttggtca tcccggagtg ggaggcggaa 3240

gtgagccagg ccgtccgggt ctgccctggg cgcggttaca tcctccgcgt gacagcgtac 3300

aaggagggct acggtgaggg ctgcgtgacg atccacgaga ttgagaacaa cacggacgag 3360

cttaagttca agaactgcga ggaggaggaa gtgtacccga cagacaccgg cacctgcaac 3420

gactacaccg cccaccaagg gaccgccgcc tgcaacagcc gcaacgcggg ctatgaagat 3480

gcgtacgagg ttgataccac cgcctcagtg aactacaaac cgacttatga ggaggagaca 3540

tacacggacg tcaggcgcga caaccattgt gagtacgacc gtggctacgt gaactatccg 3600

ccggtgccag cgggctacat gacgaaggag ctagaatact tccctgagac ggacaaggtg 3660

tggattgaaa tcggcgagac cgagggcaag tttatcgtgg attctgtcga gctgctgcta 3720

atggaggagt ag 3732

<210> 41
<211> 1243
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_13.

<400> 41

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val

```
           100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

Ala Arg Phe Asn Trp Arg Asn Pro Leu Asn Ser Leu Arg Gly Ser Leu
            420                 425                 430
```

```
Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Thr Ala
                645                 650                 655

Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Glu Ala Val Asn
            660                 665                 670

Ala Leu Phe Thr Asn Thr Asn Pro Arg Arg Leu Lys Thr Gly Val Thr
        675                 680                 685

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

Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Leu Glu Lys Val Lys Tyr
705                 710                 715                 720

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

Thr Ser Ile Asn Lys Gln Pro Asp Phe Ile Ser Thr Asn Glu Gln Ser
            740                 745                 750
```

```
Asn Phe Thr Ser Ile His Glu Gln Ser Glu His Gly Trp Trp Gly Ser
        755                 760                 765

Glu Asn Ile Thr Ile Gln Glu Gly Asn Asp Val Phe Lys Glu Asn Tyr
    770                 775                 780

Val Ile Leu Pro Gly Thr Phe Asn Glu Cys Tyr Pro Thr Tyr Leu Tyr
785                 790                 795                 800

Gln Lys Ile Gly Glu Ala Glu Leu Lys Ala Tyr Thr Arg Tyr Gln Leu
                805                 810                 815

Ser Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg
            820                 825                 830

Tyr Asn Ala Lys His Glu Thr Leu Asp Val Pro Gly Thr Glu Ser Val
        835                 840                 845

Trp Pro Leu Ser Val Glu Ser Pro Ile Gly Arg Cys Gly Glu Pro Asn
    850                 855                 860

Arg Cys Ala Pro His Phe Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys
865                 870                 875                 880

Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp
                885                 890                 895

Ile Asp Val Gly Cys Ile Asp Leu His Glu Asn Leu Gly Val Trp Val
            900                 905                 910

Val Phe Lys Ile Lys Thr Gln Glu Gly His Ala Arg Leu Gly Asn Leu
        915                 920                 925

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

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu
945                 950                 955                 960

Glu Thr Lys Arg Val Tyr Thr Glu Ala Lys Glu Ala Val Asp Ala Leu
                965                 970                 975

Phe Val Asp Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Gly
            980                 985                 990

Met Ile His Ala Ala Asp Lys Leu  Val His Arg Ile Arg  Glu Ala Tyr
        995                 1000                1005

Leu Ser  Glu Leu Ser Val Ile  Pro Gly Val Asn Ala  Glu Ile Phe
    1010                1015                1020

Glu Glu  Leu Glu Gly Arg Ile  Ile Thr Ala Ile Ser  Leu Tyr Asp
    1025                1030                1035

Ala Arg  Asn Val Val Lys Asn  Gly Asp Phe Asn Asn  Gly Leu Ala
    1040                1045                1050

Cys Trp  Asn Val Lys Gly His  Val Asp Val Gln Gln  Ser His His
    1055                1060                1065

Arg Ser  Val Leu Val Ile Pro  Glu Trp Glu Ala Glu  Val Ser Gln
```

1070 1075 1080

Ala Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr
1085 1090 1095

Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu
1100 1105 1110

Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Lys Asn Cys Glu Glu
1115 1120 1125

Glu Glu Val Tyr Pro Thr Asp Thr Gly Thr Cys Asn Asp Tyr Thr
1130 1135 1140

Ala His Gln Gly Thr Ala Ala Cys Asn Ser Arg Asn Ala Gly Tyr
1145 1150 1155

Glu Asp Ala Tyr Glu Val Asp Thr Thr Ala Ser Val Asn Tyr Lys
1160 1165 1170

Pro Thr Tyr Glu Glu Glu Thr Tyr Thr Asp Val Arg Arg Asp Asn
1175 1180 1185

His Cys Glu Tyr Asp Arg Gly Tyr Val Asn Tyr Pro Pro Val Pro
1190 1195 1200

Ala Gly Tyr Met Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp
1205 1210 1215

Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Lys Phe Ile Val
1220 1225 1230

Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1235 1240

<210> 42
<211> 3702
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_14.

<400> 42
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgaccgcgac gtttgaagct 1980

gaatccgacc tcgagcgtgc gcgcaaggcg gtgaacgctc tgttcacgag caccaaccct 2040

cgtggcttga agacggatgt gacggactac cacatcgacc aagtctcgaa cctcgtggag 2100

tgcctgagcg acgagttctg tcttgacaag aagcgcgagc tgctggagga ggtgaagtac 2160

gccaagcgcc tctccgatga gcgcaacctg ctccaagatc ctaccttcac gtcgatttcc 2220

ggccaaaccg accgtggatg gatcggctcg actggcatct ccatccaggg cggcgacgac 2280

atcttcaagg agaactatgt tcggctgccg ggcacggtgg acgagtgtta cccgacgtac 2340

ctctaccaga agatagacga gagtcaactc aagtcctaca cgcggtatca gttacgtggc 2400

tacattgaag actcccagga cttggaaatc tatctcatac ggtacaacgc caagcacgag 2460

accttaagcg tgccgggaac ggagtcgccc tggccaagct ctggcgtgta cccttccggt 2520

aggtgcggcg agcccaaccg ctgtgcacct cgaatcgaat ggaacccgga ccttgactgc 2580

tcttgccggt acggcgagaa gtgcgtccat cattctcacc acttcagctt ggacattgac 2640

gtcggctgca ccgacctcaa tgaagacctc ggagtgtggg tcatcttcaa gatcaagaca 2700

caggacgggc acgcgaaact aggaaacctg gagttcatcg aggagaagcc actcctcggc 2760

aaggcacttt ccagggtcaa gcgggccgag aagaaatgga gggacaagta cgagaaactc 2820

cagctcgaaa caaagcgggt gtacacggag gcaaaggaat ccgtggacgc cctgttcgtg 2880

gactctcagt acgacaagct ccaggcgaac acaaacattg gcatcatcca cggtgcggac 2940

aagcaagtgc acaggatacg ggagccttac ctctcggagc tgccggtgat tccctcgatc 3000

aacgcggcga tcttcgagga actggagggc cacatcttca aggcgtattc tctgtacgac 3060

gcgcgtaacg tcatcaagaa cggcgacttc aacaatgggc tgtcctgctg gaacgttaaa 3120

ggccacgtcg atgtccagca gaaccaccat aggtcagtcc tggtgctgag cgagtgggag 3180

gcggaggtgt cccagaaggt gcgcgtgtgc ccggatcgcg gctacatctt gagggtgaca 3240

gcctacaagg agggctacgg cgagggctgt gtcacgatcc atgagttcga ggacaacacg 3300

gatgtcctga aattccgtaa cttcgtcgag gaggaggtct atcccaacaa caccgtgacc 3360

tgcaacgact acacgaccaa tcagtcggct gagggcagta ccgatgcctg caacagctac 3420

aaccgtggtt acgaagatgg atacgagaac cgctacgagc ccaatccttc ggctcccgtg 3480

aattacactc ccacgtacga ggagggcatg tacactgaca ctcagggcta caaccattgc 3540

gtcagcgacc gtggctaccg caaccacacg ccgctcccag cgggctacgt gacgctggag 3600

ctggaatact ttcccgagac agaacaagtg tggatagaga tcggcgagac cgagggcaca 3660

ttcatcgtgg gctctgtgga attgctcctc atggaggagt aa 3702

<210> 43
<211> 1200
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_14.

<400> 43

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp

```
            20                  25                  30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
        35                  40                  45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350
```

```
Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Thr Ala
                645                 650                 655

Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Arg Lys Ala Val Asn
            660                 665                 670
```

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

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

Glu Phe Cys Leu Asp Lys Lys Arg Glu Leu Leu Glu Glu Val Lys Tyr
705                 710                 715                 720

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

Thr Ser Ile Ser Gly Gln Thr Asp Arg Gly Trp Ile Gly Ser Thr Gly
            740                 745                 750

Ile Ser Ile Gln Gly Gly Asp Asp Ile Phe Lys Glu Asn Tyr Val Arg
        755                 760                 765

Leu Pro Gly Thr Val Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys
    770                 775                 780

Ile Asp Glu Ser Gln Leu Lys Ser Tyr Thr Arg Tyr Gln Leu Arg Gly
785                 790                 795                 800

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn
                805                 810                 815

Ala Lys His Glu Thr Leu Ser Val Pro Gly Thr Glu Ser Pro Trp Pro
            820                 825                 830

Ser Ser Gly Val Tyr Pro Ser Gly Arg Cys Gly Glu Pro Asn Arg Cys
        835                 840                 845

Ala Pro Arg Ile Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Tyr
    850                 855                 860

Gly Glu Lys Cys Val His His Ser His His Phe Ser Leu Asp Ile Asp
865                 870                 875                 880

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
                885                 890                 895

Lys Ile Lys Thr Gln Asp Gly His Ala Lys Leu Gly Asn Leu Glu Phe
            900                 905                 910

Ile Glu Glu Lys Pro Leu Leu Gly Lys Ala Leu Ser Arg Val Lys Arg
        915                 920                 925

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

Lys Arg Val Tyr Thr Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
945                 950                 955                 960

Asp Ser Gln Tyr Asp Lys Leu Gln Ala Asn Thr Asn Ile Gly Ile Ile
                965                 970                 975

His Gly Ala Asp Lys Gln Val His Arg Ile Arg Glu Pro Tyr Leu Ser
            980                 985                 990

Glu Leu Pro Val Ile Pro Ser Ile  Asn Ala Ala Ile Phe  Glu Glu Leu
```

995 1000 1005

Glu Gly His Ile Phe Lys Ala Tyr Ser Leu Tyr Asp Ala Arg Asn
1010 1015 1020

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn
1025 1030 1035

Val Lys Gly His Val Asp Val Gln Gln Asn His His Arg Ser Val
1040 1045 1050

Leu Val Leu Ser Glu Trp Glu Ala Glu Val Ser Gln Lys Val Arg
1055 1060 1065

Val Cys Pro Asp Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys
1070 1075 1080

Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Phe Glu Asp
1085 1090 1095

Asn Thr Asp Val Leu Lys Phe Arg Asn Phe Val Glu Glu Glu Val
1100 1105 1110

Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Thr Asn Gln
1115 1120 1125

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

Tyr Glu Asp Gly Tyr Glu Asn Arg Tyr Glu Pro Asn Pro Ser Ala
1145 1150 1155

Pro Val Asn Tyr Thr Pro Thr Tyr Glu Glu Gly Met Tyr Thr Asp
1160 1165 1170

Thr Gln Gly Tyr Asn His Cys Val Ser Asp Arg Gly Tyr Arg Asn
1175 1180 1185

His Thr Pro Leu Pro Ala Gly Tyr Val Thr Leu Glu
1190 1195 1200

<210> 44
<211> 3687
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_15.

<400> 44
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagcagtc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctccgtgc tctcacgctg gtccaacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cggatgctac ctttgaagca 1980

gagtccgact tggaacgtgc acagaaggca gtgaacgcac tcttcacctc aagcaaccag 2040

atcggattga agacagatgt gacagattac cacatcgacc aagtgagcaa cttggtggat 2100

tgcttgtcag atgagttctg cttggatgag aagcgtgaac tctccgagaa ggtgaagcac 2160

gcaaagcgtc tctcagatga acgtaatctc cttcaagacc ctaactttcg tggtatcaat 2220

cgtcagccag atcgtggatg gcgtggatca acagacatca ccatccaggg aggcgatgat 2280

gtgttcaagg agaactacgt gaccctccca ggaaccgtgg atgaatgcta cccaacctac 2340

ctctaccaga agatcgacga gtcaaagctc aaggcttaca cccgttatga actccgtggc 2400

tacatcgaag atagccagga tctcgaaatc tatctcatcc gttacaatgc taagcacgaa 2460

atcgtgaatg tgccaggaac cggctcactc tggccactct cagcacagtc accaatcggc 2520

aagtgcggcg aacccaatcg ctgcgctcct catctcgaat ggaatcccga tctcgactgc 2580

tcctgccgag acggcgagaa gtgtgcacat cactcacacc acttcaccct cgacatcgac 2640

gtgggctgca ccgacctcaa tgaagacctg ggcgtgtggg tgatcttcaa gatcaagacc 2700

caggacggcc acgcacgact gggcaatctg gagtttctgg aggagaagcc actgcttggc 2760

gaggcactgg cacgagtgaa acgagccgag aagaaatggc gagacaaacg tgagaagctg 2820

caactggaga ccaacatcgt gtacaaagag gccaaagagt cagttgacgc cctgtttgtc 2880

aatagccagt atgaccgact gcaagttgac accaacatcg ccatgatcca cgctgcggac 2940

aagcgcgtcc accgcatccg cgaggcttat ctgcccgagc tgagcgtcat tcccggcgtc 3000

aatgccgcga tcttcgagga gttagagggc cgcatcttca ccgcctacag cctctatgac 3060

gcccgcaatg tcattaagaa tggcgacttc aacaatggct tactatgctg gaatgtcaaa 3120

gggcacgttg acgtcgagga gcagaacaat caccgcagcg tcttagtcat acccgagtgg 3180

gaggccgaag tcagccagga agtccgcgtc tgtccagggc gcgggtacat cctgcgggtc 3240

accgcctaca aagagggata cggcgagggt tgtgtcacca tacacgagat agaggacaat 3300

accgacgaac tcaagttcag caattgtgtc gaggaggaag tctatcccaa caataccgta 3360

acctgcaaca actacaccgg aacccaggag gagtatgaag ggacgtacac ctcgcggaac 3420

cagggctatg acgaagccta tgggaacaac ccgtcggtgc ctgctgacta tgcgtcggtc 3480

tatgaggaga aatcgtacac ggacgggcgg cgggagaatc cgtgtgagtc gaatcgcggg 3540

tatggtgact acacgccgct accggcgggc tatgtaacga aagacctgga atacttcccg 3600

gagacggaca aagtatggat agagataggc gagacggagg gaacgttcat cgtggactcg 3660

gtagagctgc tgctcatgga ggagtga 3687

<210> 45
<211> 1228
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_15.

<400> 45

```
Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1               5                  10                  15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
            20                  25                  30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
        35                  40                  45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
                85                  90                  95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
            100                 105                 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
        115                 120                 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320
```

```
Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

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

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640
```

```
Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Asp Ala
                645                 650                 655

Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn
            660                 665                 670

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

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Asp Cys Leu Ser Asp
    690                 695                 700

Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His
705                 710                 715                 720

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

Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp
            740                 745                 750

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

Leu Pro Gly Thr Val Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys
    770                 775                 780

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

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn
                805                 810                 815

Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro
            820                 825                 830

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys
        835                 840                 845

Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp
    850                 855                 860

Gly Glu Lys Cys Ala His His Ser His His Phe Thr Leu Asp Ile Asp
865                 870                 875                 880

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
                885                 890                 895

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
            900                 905                 910

Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys Arg
        915                 920                 925

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr
    930                 935                 940

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

Asn Ser Gln Tyr Asp Arg Leu Gln Val Asp Thr Asn Ile Ala Met Ile
```

965 970 975

His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro
980 985 990

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
995 1000 1005

Glu Gly Arg Ile Phe Thr Ala Tyr Ser Leu Tyr Asp Ala Arg Asn
1010 1015 1020

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn
1025 1030 1035

Val Lys Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser
1040 1045 1050

Val Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val
1055 1060 1065

Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr
1070 1075 1080

Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu
1085 1090 1095

Asp Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu
1100 1105 1110

Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asn Tyr Thr Gly Thr
1115 1120 1125

Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Gln Gly Tyr
1130 1135 1140

Asp Glu Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala
1145 1150 1155

Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn
1160 1165 1170

Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro
1175 1180 1185

Ala Gly Tyr Val Thr Lys Asp Leu Glu Tyr Phe Pro Glu Thr Asp
1190 1195 1200

Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val
1205 1210 1215

Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1220 1225

<210> 46
<211> 3600
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC868_29.

<400> 46
atgacgagca accggaagaa cgagaacgag atcatcaacg ccctctcgat ccctgctgtt 60

tcaaaccact ccgcgcagat gaacctgtcc accgacgcgc gcatcgagga ctccctctgc 120

atagccgagg gcaacaacat cgacccattc gtgtcggcca gcacggttca gaccggcatc 180

aacatcgcgg gccgtatcct cggcgtcctc ggtgtcccat tcgccggtca gatcgcgtcc 240

ttctactcgt tccttgtggg cgagctgtgg cctcgcggtc gtgacccgtg ggagatcttc 300

ctggagcatg tggagcagtt gatccggcag caagtcacgg agaacacccg cgatactgct 360

ctggccaggc tacagggcct gggaaactcc tttcgggcat accagtactc actggaggac 420

tggttggaga acagggatga cgcgcgaaca cgctcggtac tctacaccca gtacatcgct 480

ctcgaactcg acttcctgaa cgctatgccg ctgttcgcca tcaggaacca ggaagttcca 540

ctccttatgg tgtacgccca ggccgccaac ttacatctgc tcctgctgcg ggacgccagc 600

ctgttcggct ccgagttcgg actcacatct caagaaatcc agcgttacta cgagcgccaa 660

gtggagaaga cccgtgagta cagtgactac tgcgctcgat ggtacaacac agggctcaac 720

aacctgcgcg gcaccaacgc tgagtcatgg ctccgttaca accagttccg ccgcgacttg 780

actttgggtg tcctagacct ggtggcgcta ttcccgtctt acgacacacg ggtgtaccca 840

atgaacacta gcgcgcaact cacgcgggag atctacacag acccaatcgg ccggacgaac 900

gcaccctccg gtttcgcatc cacgaattgg ttcaacaaca acgcaccctc cttctcggca 960

atcgaggccg ccgtcatccg ccctcctcac ctgctcgact ttcccgagca gctcacgatc 1020

ttctcccagc tctcacgctg gtcccacaca cagtacatga actactgggt cgggcaccga 1080

ttggagagta ggacgatccg tggcagcttg agcaccagta cccacggcaa caccaacacc 1140

tccatcaacc cagttacgct acagttcacg agccgcgacg tttaccggac tgagtcgttc 1200

gcgggcatta acatccttct gacaacgccc gtcaacggcg tcccgtgggc ccggttcaac 1260

tggcgtaacc cgttgaactc cctgcgcggg tcattgctct acaccatcgg gtacacgggc 1320

gtcggcaccc agctcttcga cagtgaaact gagctgccgc ccgagaccac ggaacgcccg 1380

aactacgagt cctacagcca ccgcctgtcc aacatccggc tcatctctgg caacacgctg 1440

cgtgcgccgg tgtactcctg gacacaccgc agcgccgacc ggaccaacac gatctcttcc 1500

gactccatta accagatccc gctcgtgaag ggcttccgtg tgtggggtgg cacgagcgtc 1560

atcaccggtc cgggcttcac cggtggagac atactgcggc gcaacacttt cggcgacttc 1620

gtttcgttgc aagtgaacat caactcgccg atcacccagc gttaccgtct gaggttccgc 1680

tacgcttcaa gccgcgacgc gagggtcatt gtcctgaccg gagccgcgtc cacaggcgtg 1740

ggaggccaag tctcagtcaa catgcctctc cagaagacga tggagatagg cgagaacttg 1800

actagccgaa ccttccggta cactgatttc tcgaaccctt tctcattcag agcgaaccct 1860

gacatcattg ggatctccga gcaaccgctg ttcggtgctg gctccatcag ctctggcgaa 1920

ctgtacatcg acaagattga gatcatcctg gcggatgcga cgttcgaggc cgagtctgac 1980

ctggagcggg ctcagaaggc tgtcaacgaa ctgttcacca gcagcaacca gattgggctc 2040

aagaccgacg tcacggacta tcacattgac caagtgtcca accttgtgga gtgcctgtcc 2100

gacgagttct gcctcgacga gaagaaggag ctgtccgaga aggtcaaaca cgcgaagcgt 2160

ctgagtgacg agcggaattt gctccaggac ccgaacttcc gtggcatcaa ccgccagctc 2220

gaccgtggtt ggcgcgggag tacagacatc accatccagg gaggcgacga tgtgttcaag 2280

gagaactatg tgacgctgct cgggactttc gacgaatgct acccgacgta tctctaccag 2340

aagatagacg agagtaaatt gaaggcgtac acccgctacc agcttcgcgg gtacatcgag 2400

gatagtcagg acctggaaat ctacctgatc cgatacaacg ccaagcacga gacagtgaac 2460

gtgccaggca cgggctcact ttggccattg agcgctccct ctccaatcgg aaagtgcgct 2520

caccactcgc accacttctc tctggacatc gacgtgggct gcaccgacct caacgaggac 2580

ctgggtgtct gggttatctt caagattaag acccaggacg gacatgcccg cctcggcaac 2640

ctggagttcc ttgaggagaa gcctctcgtg ggcgaggccc tcgctcgtgt gaagcgcgcc 2700

gagaagaaat ggcgagacaa gcgggagaag ctggagtggg agaccaacat cgtgtacaag 2760

gaggccaagg agtcagtgga cgcactcttc gtcaacagcc agtacgaccg cctccaggct 2820

gacaccaaca tcgccatgat ccacgcggct gacaagcggg tccacagcat ccgtgaggcg 2880

tacctgcccg agctgtcagt gatccctggt gtgaacgcgg cgatcttcga ggaactggag 2940

ggccgcatct tcacagcatt cagcctgtac gatgccagga atgttattaa gaacggtgac 3000

ttcaacaacg ggctgagttg ctggaacgtc aagggccatg tggacgtcga ggagcagaac 3060

aaccaccggt ccgtgctggt cgtgccggag tgggaggcag aggtgagcca ggaggtccgc 3120

gtctgccctg gtcgcggcta catcctccgt gtgactgcgt acaaggaagg ctacggtgaa 3180

ggctgcgtga ctatccacga gatcgagaac aacaccgacg agctcaagtt ctcgaactgt 3240

gtggaggagg aggtgtaccc gaacaacacc gttacttgca acgactacac tgccacgcaa 3300

gaggagtacg agggcactta cacttcccgg aatcgcggct atgatggcgc gtacgagtcc 3360

aacagcagcg tgcctgcgga ttatgcgtcc gcttacgagg agaaggcgta caccgacgga 3420

cggagggaca acccttgcga gtccaaccgt ggctacggtg actacactcc gctgcccgcc 3480

gggtacgtca ccaaggagct ggagtacttc ccggagaccg acaaagtctg gatcgagatc 3540

ggcgagacgg agggcacttt catcgtggac tcggtcgagc tgctactgat ggaggagtga 3600

<210> 47
<211> 1199
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của biến thể protein khảm TIC868_29.

<400> 47

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

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

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

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

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

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145 150 155 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
165 170 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
180 185 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
195 200 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
210 215 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225 230 235 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
245 250 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
260 265 270

```
Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Gln Leu Ser Arg Trp Ser His Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

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

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Asn Gln Ile Pro Leu Val Lys Gly Phe
            500                 505                 510

Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln
    530                 535                 540

Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg
545                 550                 555                 560

Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala
                565                 570                 575

Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys
            580                 585                 590

Thr Met Glu Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr
```

```
        595                 600                 605

Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly
    610                 615                 620

Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu
625                 630                 635                 640

Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
                645                 650                 655

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

Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His
        675                 680                 685

Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys
    690                 695                 700

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

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

Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile
            740                 745                 750

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

Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu
    770                 775                 780

Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu
785                 790                 795                 800

Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His
                805                 810                 815

Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala
            820                 825                 830

Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu
        835                 840                 845

Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp
    850                 855                 860

Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn
865                 870                 875                 880

Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg
                885                 890                 895

Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu
            900                 905                 910

Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala
        915                 920                 925
```

```
Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile
    930                 935                 940

Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala
945                 950                 955                 960

Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe
                965                 970                 975

Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala
            980                 985                 990

Arg Asn Val Ile Lys Asn Gly Asp  Phe Asn Asn Gly Leu  Ser Cys Trp
        995                 1000                1005

Asn Val  Lys Gly His Val Asp  Val Glu Glu Gln Asn  Asn His Arg
    1010                1015                1020

Ser Val  Leu Val Val Pro Glu  Trp Glu Ala Glu Val  Ser Gln Glu
    1025                1030                1035

Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile Leu Arg  Val Thr Ala
    1040                1045                1050

Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val Thr Ile  His Glu Ile
    1055                1060                1065

Glu Asn  Asn Thr Asp Glu Leu  Lys Phe Ser Asn Cys  Val Glu Glu
    1070                1075                1080

Glu Val  Tyr Pro Asn Asn Thr  Val Thr Cys Asn Asp  Tyr Thr Ala
    1085                1090                1095

Thr Gln  Glu Glu Tyr Glu Gly  Thr Tyr Thr Ser Arg  Asn Arg Gly
    1100                1105                1110

Tyr Asp  Gly Ala Tyr Glu Ser  Asn Ser Ser Val Pro  Ala Asp Tyr
    1115                1120                1125

Ala Ser  Ala Tyr Glu Glu Lys  Ala Tyr Thr Asp Gly  Arg Arg Asp
    1130                1135                1140

Asn Pro  Cys Glu Ser Asn Arg  Gly Tyr Gly Asp Tyr  Thr Pro Leu
    1145                1150                1155

Pro Ala  Gly Tyr Val Thr Lys  Glu Leu Glu Tyr Phe  Pro Glu Thr
    1160                1165                1170

Asp Lys  Val Trp Ile Glu Ile  Gly Glu Thr Glu Gly  Thr Phe Ile
    1175                1180                1185

Val Asp  Ser Val Glu Leu Leu  Leu Met Glu Glu
    1190                1195
```

<210> 48
<211> 3432
<212> ADN
<213> Nhân tạo

<220>

<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC869.

<400> 48

atggagataa ataatcagaa gcaatgcata ccatataatt gcttaagtaa tcctgaggaa 60

gtacttttgg atggggagag gatattacct gatatcgatc cactcgaagt ttctttgtcg 120

cttttgcaat ttcttttgaa taactttgtt ccaggggggag gctttatttc aggattagtt 180

gataaaatat ggggggcttt gagaccatct gaatgggact tatttcttgc acagattgaa 240

cggttgattg atcaaagaat agaagcaaca gtaagagcaa aagcaatcac tgaattagaa 300

ggattaggga gaaattatca aatatacgct gaagcattta aagaatggga atcagatcct 360

gataacgaag cggctaaaag tagagtaatt gatcgctttc gtatacttga tggtctaatt 420

gaagcaaata tcccttcatt tcggataatt ggatttgaag tgccactttt atcggtttat 480

gttcaagcag ctaatctaca tctcgctcta ttgagagatt ctgttatttt tggagagaga 540

tggggattga cgacaaaaaa tgtcaatgat atctataata gacaaattag agaaattcat 600

gaatatagca atcattgcgt agatacgtat aacacagaac tagaacgtct agggtttaga 660

tctatagcgc agtggagaat atataatcag tttagaagag aactaacact aactgtatta 720

gatattgtcg ctcttttccc gaactatgac agtagactgt atccgatcca aacttttct 780

caattgacaa gagaaattgt tacatcccca gtaagcgaat tttattatgg tgttattaat 840

agtggtaata taattggtac tcttactgaa cagcagataa ggcgaccaca tcttatggac 900

ttctttaact ccatgatcat gtatacatca gataatagac gggaacatta ttggtcagga 960

cttgaaatga cggcttattt tacaggattt gcaggagctc aagtgtcatt ccctttagtc 1020

gggactagag gggagtcagc tccaccatta actgttagaa gtgttaatga tggaatttat 1080

agaatattat cggcaccgtt ttattcagcg ccttttctag gcaccattgt attgggaagt 1140

cgtggagaaa aatttgattt tgcgcttaat aatatttcac ctccgccatc tacaatatac 1200

agacatcctg gaacagtaga ttcactagtc agtataccgc cacaggataa tagcgtacca 1260

ccgcacaggg gatctagtca tcgattaagt catgttacaa tgcgcgcaag ttcccctata 1320

ttccattgga cgcatcgcag cgcaaccact acaaatacaa ttaatccaaa tgctattatc 1380

caaataccac tagtaaaagc atttaacctt cattcaggtg ccactgttgt tagaggacca 1440

gggtttacag gtggagatct cttacgaaga acgaatactg gtacatttgc agacataaga 1500

gtcaatgttc cttcatcact attttcccaa agatatcgcg taaggattcg ttatgcttct 1560

actaccgatt tacaattttt cacgagaatt aatggaactt ctgttaatca aggtaatttc 1620

tcaaaaacga tggatagagg ggataaactg aaatctgaaa actttagaac tgccggattt 1680

agtactcctt ttagattttc aaattttcaa agtacattca cgttgggtac tcaggctttt 1740

tcaaatcagg aagtttatat agatagaatt gaatttgtcc cggcagaagt aacattcgag 1800

gcagaatctg atttagaaag agcacaaaag gcggtgaatg agctgtttac ttcttccaat 1860

caaatcgggt taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1920

gagtgtttat ctgatgaatt ttgtctggat gaaaaaaaag aattgtccga gaaagtcaaa 1980

catgcgaagc gacttagtga tgagcggaat ttacttcaag atccaaactt tagagggatc 2040

aatagacaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggcgat 2100

gacgtattca aagagaatta cgttacgcta ttgggtacct ttgatgagtg ctatccaacg 2160

tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgtta ccaattaaga 2220

gggtatatcg aagatagtca agacttagaa atctatttaa ttcgctacaa tgccaaacac 2280

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagcccc aagtccaatc 2340

ggaaaatgtg cccatcattc ccatcatttc tccttggaca ttgatgttgg atgtacagac 2400

ttaaatgagg acttaggtgt atgggtgata ttcaagatta agacgcaaga tggccatgca 2460

agactaggaa atctagaatt tctcgaagag aaaccattag taggagaagc actagctcgt 2520

gtgaaaagag cggagaaaaa atggagagac aaacgtgaaa aattggaatg ggaaacaaat 2580

attgtttata aagaggcaaa agaatctgta gatgctttat ttgtaaactc tcaatatgat 2640

agattacaag cggataccaa catcgcgatg attcatgcgg cagataaacg cgttcatagc 2700

attcgagaag cttatctgcc tgagctgtct gtgattccgg gtgtcaatgc ggctattttt 2760

gaagaattag aagggcgtat tttcactgca ttctccctat atgatgcgag aaatgtcatt 2820

aaaaatggtg attttaataa tggcttatcc tgctggaacg tgaaagggca tgtagatgta 2880

gaagaacaaa acaaccaccg ttcggtcctt gttgttccgg aatgggaagc agaagtgtca 2940

caagaagttc gtgtctgtcc gggtcgtggc tatatccttc gtgtcacagc gtacaaggag 3000

ggatatggag aaggttgcgt aaccattcat gagatcgaga acaatacaga cgaactgaag 3060

tttagcaact gtgtagaaga ggaagtatat ccaaacaaca cggtaacgtg taatgattat 3120

actgcgactc aagaagaata tgagggtacg tacacttctc gtaatcgagg atatgacgga 3180

gcctatgaaa gcaattcttc tgtaccagct gattatgcat cagcctatga agaaaaagca 3240

tatacagatg gacgaagaga caatccttgt gaatctaaca gaggatatgg ggattacaca 3300

ccactaccag ctggctatgt gacaaaagaa ttagagtact tcccagaaac cgataaggta 3360

tggattgaga tcggagaaac ggaaggaaca ttcatcgtgg acagcgtgga attacttctt 3420

atggaggaat ag 3432

<210> 49
<211> 3432

<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC869.

<400> 49
atggagataa acaaccagaa gcagtgcatt ccgtacaact gcctcagcaa cccggaggag 60

gtgctgctgg acggcgagcg tatcctccca gacatcgacc cactggaggt cagcctgagc 120

ctcctccagt tcctcctcaa taacttcgtg ccaggcggcg gcttcatctc cggcctggtg 180

gacaagatct ggggcgcact ccggccaagt gagtgggatc tgttcctggc ccaaatcgag 240

cgcctgatcg accagaggat cgaggcgacg gtccgcgcca aggcgataac cgagctggag 300

ggcctcggtc gcaactacca gatctacgca gaggcgttca aggagtggga gagcgacccg 360

gacaacgagg cggccaagtc tcgggtgatt gaccgcttcc gcatcctcga cggcctcatc 420

gaagccaaca tcccttcctt ccggatcata ggcttcgaag tcccgctcct cagcgtgtac 480

gtgcaagcgg ccaatctcca cctcgcgttg ctccgtgaca gcgtcatctt tggcgagaga 540

tggggcctga cgacgaagaa cgtgaacgac atctacaaca ggcagatccg agagattcac 600

gagtacagca accactgcgt ggacacatac aacacggagc tggagcggct cggcttccgc 660

tcaatcgctc agtggcggat ctacaaccag ttccgccgcg agctgaccct caccgtgctc 720

gacatcgtcg cattgtttcc caattacgac tcacgcctct acccaatcca gactttcagc 780

cagctcacac gcgagattgt gaccagcccg gtgtcagagt tctactacgg cgtcatcaac 840

tcaggcaaca tcatcgggac actgactgaa cagcagatca gacgtccgca cttgatggac 900

ttcttcaact ccatgattat gtacacatca gacaacagga gagagcacta ctggtccggg 960

ttggagatga ctgcttactt caccggcttc gccggtgccc aagtgagctt cccactggtc 1020

ggaactcgtg gcgagtcagc tcctccgcta actgtgcgat ctgtcaacga cgggatctac 1080

agaatactgt cggctccctt ctacagtgcg ccgttcctcg gcaccatcgt cctcggctca 1140

cgtggtgaga agttcgactt cgcactgaac aacattagcc cgccgcctag tacaatctac 1200

aggcaccctg gcaccgtgga ctcactggtt tcgatcccgc cacaagacaa cagtgtgccg 1260

ccacatcgtg gttctagcca caggctctcc catgtgacca tgcgcgcctc ttcaccgatc 1320

tttcactgga cccatcggtc cgctacaacc acaaacacca tcaaccctaa cgccatcatc 1380

caaatcccgc tggtgaaggc gtttaacctc cacagcggcg caactgtcgt gcgcggccct 1440

ggattcaccg gtggtgacct gctccgtcgg accaatactg gcacgttcgc agacatccga 1500

gtgaacgtcc cgtcctcgct gttcagtcag cgctaccgtg tccgcattcg gtacgcttcc 1560

accacggatc tccagttctt tactcgcatc aatgggacga gcgtcaacca gggcaacttc 1620

agcaagacga tggaccgtgg agataagctc aagtccgaga acttccgcac ggctggcttc 1680

tcgacaccgt tcagattcag caacttccag agcactttca cgctgggcac acaggcgttc 1740

tccaaccagg aggtgtacat cgaccgcatc gagttcgtgc ctgctgaggt taccttcgag 1800

gcggaaagcg acctcgaaag ggcccagaag gccgtcaacg agctgttcac ctccagcaac 1860

cagatcggtc tcaagaccga cgtcactgac tatcacattg accaagtcag caacctggtg 1920

gagtgcctca gtgatgagtt ctgcctggat gagaagaagg agcttagcga gaaggtcaag 1980

cacgcaaagc gcttgagcga cgagcgcaac cttctccagg acccgaattt ccgtggtatc 2040

aatagacagc ttgaccgtgg gtggcgcggt agtaccgaca taaccatcca gggtggcgac 2100

gatgtgttca aggagaatta tgttacgctg ctcggtacgt tcgacgagtg ctatcccacg 2160

tacttgtacc agaagattga cgagagcaag ctcaaggcgt acacccgtta ccagctccgt 2220

ggctacatcg aggacagcca ggatctggaa atctaccttа tccgatacaa tgctaagcac 2280

gagacagtca acgtgcccgg aacagggtcg ctctggccgc tcagtgctcc gtcgcccatt 2340

ggcaagtgcg cgcaccattc gcatcacttc tcacttgaca ttgacgtggg ctgcaccgac 2400

ctgaacgagg atctgggtgt ctgggtcatc ttcaagatca agacccaaga cggccacgcg 2460

cgcctcggga acctggagtt cctggaggag aagcctttgg taggtgaagc cctggcccgc 2520

gtcaagcgcg cggagaagaa gtggcgcgac aagagggaga agctggaatg ggagaccaac 2580

atcgtgtaca aggaggcgaa ggagtcggtg gacgcactat tcgtgaactc ccagtacgac 2640

cgtctccagg ccgacaccaa catcgccatg atccacgccg ctgacaaacg agttcattcc 2700

attcgtgaag cctatcttcc cgagctgtct gtcataccgg gcgtcaacgc ggccatcttc 2760

gaggagttag agggtcggat ctttacagct ttctcactgt acgatgcccg caacgtcatc 2820

aagaacggcg acttcaacaa cggtctctcc tgttggaacg tgaagggcca cgtggatgtc 2880

gaggagcaga acaaccaccg ctctgtgctt gtggtgcccg agtgggaggc cgaggtgagc 2940

caggaggtcc gcgtctgtcc gggtcgcggc tacatcctgc gggtcaccgc ctacaaggag 3000

ggctacggcg aaggctgcgt tactattcac gagattgaga acaataccga cgaactcaag 3060

ttctccaact gtgtcgagga ggaggtgtac ccgaacaaca ccgtgacgtg caacgactac 3120

accgcgacac aggaggaata cgagggcacc tacaccagcc gcaaccgagg ctacgacgga 3180

gcgtacgaga gcaactcgtc cgtgcccgct gattacgcga gtgcgtacga ggagaaggct 3240

tacaccgacg gacggcgcga caatccctgc gagagtaacc gtggatacgg agattacacg 3300

ccgctacccg ctggctacgt cactaaggaa ctggagtact tcccagagac ggacaaggtg 3360

tggatcgaaa tcggcgagac agagggcacg ttcatcgtgg actccgtgga gctgctgctg 3420

atggaggagt ga 3432

<210> 50
<211> 1143
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC869.

<400> 50

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

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

Asp Pro Leu Glu Val Ser Leu Ser Leu Leu Gln Phe Leu Leu Asn Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Ile Ser Gly Leu Val Asp Lys Ile Trp
50 55 60

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

Arg Leu Ile Asp Gln Arg Ile Glu Ala Thr Val Arg Ala Lys Ala Ile
85 90 95

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

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

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Ile Glu Ala Asn Ile
130 135 140

Pro Ser Phe Arg Ile Ile Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145 150 155 160

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

Phe Gly Glu Arg Trp Gly Leu Thr Thr Lys Asn Val Asn Asp Ile Tyr
180 185 190

Asn Arg Gln Ile Arg Glu Ile His Glu Tyr Ser Asn His Cys Val Asp
195 200 205

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

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

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

Gln Thr Phe Ser Gln Leu Thr Arg Glu Ile Val Thr Ser Pro Val Ser

```
            260                 265                 270

Glu Phe Tyr Tyr Gly Val Ile Asn Ser Gly Asn Ile Ile Gly Thr Leu
        275                 280                 285

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

Met Ile Met Tyr Thr Ser Asp Asn Arg Arg Glu His Tyr Trp Ser Gly
305                 310                 315                 320

Leu Glu Met Thr Ala Tyr Phe Thr Gly Phe Ala Gly Ala Gln Val Ser
                325                 330                 335

Phe Pro Leu Val Gly Thr Arg Gly Glu Ser Ala Pro Pro Leu Thr Val
            340                 345                 350

Arg Ser Val Asn Asp Gly Ile Tyr Arg Ile Leu Ser Ala Pro Phe Tyr
        355                 360                 365

Ser Ala Pro Phe Leu Gly Thr Ile Val Leu Gly Ser Arg Gly Glu Lys
    370                 375                 380

Phe Asp Phe Ala Leu Asn Asn Ile Ser Pro Pro Pro Ser Thr Ile Tyr
385                 390                 395                 400

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

Asn Ser Val Pro Pro His Arg Gly Ser Ser His Arg Leu Ser His Val
            420                 425                 430

Thr Met Arg Ala Ser Ser Pro Ile Phe His Trp Thr His Arg Ser Ala
        435                 440                 445

Thr Thr Thr Asn Thr Ile Asn Pro Asn Ala Ile Ile Gln Ile Pro Leu
    450                 455                 460

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

Gly Phe Thr Gly Gly Asp Leu Leu Arg Arg Thr Asn Thr Gly Thr Phe
                485                 490                 495

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

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

Arg Ile Asn Gly Thr Ser Val Asn Gln Gly Asn Phe Ser Lys Thr Met
    530                 535                 540

Asp Arg Gly Asp Lys Leu Lys Ser Glu Asn Phe Arg Thr Ala Gly Phe
545                 550                 555                 560

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

Thr Gln Ala Phe Ser Asn Gln Glu Val Tyr Ile Asp Arg Ile Glu Phe
            580                 585                 590
```

```
Val Pro Ala Glu Val Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala
        595                 600                 605

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

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

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

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

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

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

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

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

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

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

Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala
    770                 775                 780

His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp
785                 790                 795                 800

Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln
                805                 810                 815

Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro
            820                 825                 830

Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp
        835                 840                 845

Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys
    850                 855                 860

Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp
865                 870                 875                 880

Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys
                885                 890                 895

Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile
            900                 905                 910
```

Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe
915 920 925

Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp
930 935 940

Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val
945 950 955 960

Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu
965 970 975

Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile
980 985 990

Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr
995 1000 1005

Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn
1010 1015 1020

Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn
1025 1030 1035

Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser
1040 1045 1050

Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val
1055 1060 1065

Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
1070 1075 1080

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp
1085 1090 1095

Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr
1100 1105 1110

Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu
1115 1120 1125

Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1130 1135 1140

<210> 51
<211> 3513
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tái tổ hợp được sử dụng để biểu hiện ở tế bào vi khuẩn ghi mã TIC836.

<400> 51
atggagaata atattcaaaa tcaatgcgta ccttacaatt gtttaaataa tcctgaagta 60

gaaatattaa atgaagaaag aagtactggc agattaccgt tagatatatc cttatcgctt 120

acacgtttcc ttttgagtga atttgttcca ggtgtgggag ttgcgtttgg attatttgat 180

ttaatatggg gttttataac tccttctgat tggagcttat ttcttttaca gattgaacaa 240

ttgattgagc aaagaataga aacattggaa aggaaccggg caattactac attacgaggg 300

ttagcagata gctatgaaat ttatattgaa gcactaagag agtgggaagc aaatcctaat 360

aatgcacaat taagggaaga tgtgcgtatt cgatttgcta atacagacga cgctttaata 420

acagcaataa ataattttac acttacaagt tttgaaatcc ctcttttatc ggtctatgtt 480

caagcggcga atttacattt atcactatta agagacgctg tatcgtttgg gcagggttgg 540

ggactggata tagctactgt taataatcat tataatagat taataaatct tattcataga 600

tatacgaaac attgtttgga cacatacaat caaggattag aaaacttaag aggtactaat 660

actcgacaat gggcaagatt caatcagttt aggagagatt taacacttac tgtattagat 720

atcgttgctc tttttccgaa ctacgatgtt agaacatatc caattcaaac gtcatcccaa 780

ttaacaaggg aaatttatac aagttcagta attgaggatt ctccagtttc tgctaatata 840

cctaatggtt ttaatagggc ggaatttgga gttagaccgc cccatcttat ggactttatg 900

aattctttgt ttgtaactgc agagactgtt agaagtcaaa ctgtgtgggg aggacactta 960

gttagttcac gaaatacggc tggtaaccgt ataaatttcc ctagttacgg ggtcttcaat 1020

cctggtggcg ccatttggat tgcagatgag gatccacgtc ctttttatcg gacattatca 1080

gatcctgttt ttgtccgagg aggatttggg aatcctcatt atgtactggg gcttagggga 1140

gtagcatttc aacaaactgg tacgaaccac acccgaacat ttagaaatag tgggaccata 1200

gattctctag atgaaatccc acctcaggat aatagtgggg caccttggaa tgattatagt 1260

catgtattaa atcatgttac atttgtacga tggccaggtg agatttcagg aagtgattca 1320

tggagagctc caatgttttc ttggacgcac cgtagtgcaa cccctacaaa tacaattgat 1380

ccggagagga ttacacaaat acctttaaca aaatctacta atcttggctc tggaacttct 1440

gtcgttaaag gaccaggatt tacaggagga gatattcttc gaagaacttc acctggccag 1500

atttcaacct taagagtaaa tattactgca ccattatcac aaagatatcg ggtaagaatt 1560

cgctacgctt ctaccacaaa tttacaattc catacatcaa ttgacggaag acctattaat 1620

caggggaatt tttcagcaac tatgagtagt gggagtaatt tacagtccgg aagctttagg 1680

actgtaggtt ttactactcc gtttaacttt tcaaatggat caagtgtatt tacgttaagt 1740

gctcatgtct tcaattcagg caatgaagtt tatatagatc gaattgaatt tgttccggca 1800

gaagtaacct ttgaggcaga atatgattta gaaagagcgc agaaggcggt gaatgcgctg 1860

tttacgtcta caaaccaact agggctaaaa acaaatgtaa cggattatca tattgatcaa 1920

gtgtccaatt tagttacgta tttatcggat gaattttgtc tggatgaaaa gcgagaattg 1980

tccgagaaag tcaaacatgc gaagcgactc agtgatgaac gcaatttact ccaagattca 2040

aatttcaaag acattaatag gcaaccagaa cgtgggtggg gcggaagtac agggattacc 2100

atccaaggag gggatgacgt atttaaagaa aattacgtca cactatcagg tacctttgat 2160

gagtgctatc caacatattt gtatcaaaaa atcgatgaat caaaattaaa agcctttacc 2220

cgttatcaat taagagggta tatcgaagat agtcaagact tagaaatcta tttaattcgc 2280

tacaatgcaa aacatgaaac agtaaatgtg ccaggtacgg gttccttatg gccgctttca 2340

gcccaaagtc caatcggaaa gtgtggagag ccgaatcgat gcgcgccaca ccttgaatgg 2400

aatcctgact tagattgttc gtgtagggat ggagaaaagt gtgcccatca ttcgcatcat 2460

ttctccttag acattgatgt aggatgtaca gacttaaatg aggacctagg tgtatgggtg 2520

atctttaaga ttaagacgca agatgggcac gcaagactag ggaatctaga gtttctcgaa 2580

gaaaaaccat tagtaggaga agcgctagct cgtgtgaaaa gagcggagaa aaaatggaga 2640

gacaaacgtg aaaaattgga atgggaaaca aatatcgttt ataaagaggc aaaagaatct 2700

gtagatgctt tatttgtaaa ctctcaatat gatcaattac aagcggatac gaatattgcc 2760

atgattcatg cggcagataa acgtgttcat agcattcgag aagcttatct gcctgagctg 2820

tctgtgattc cgggtgtcaa tgcggctatt tttgaagaat tagaagggcg tattttcact 2880

gcattctccc tatatgatgc gagaaatgtc attaaaaatg gtgattttaa taatggctta 2940

tcctgctgga acgtgaaagg gcatgtagat gtagaagaac aaaacaacca acgttcggtc 3000

cttgttgttc cggaatggga agcagaagtg tcacaagaag ttcgtgtctg tccgggtcgt 3060

ggctatatcc ttcgtgtcac agcgtacaag gagggatatg gagaaggttg cgtaaccatt 3120

catgagatcg agaacaatac agacgaactg aagtttagca actgcgtaga ggaggaaatc 3180

tatccaaata acacggtaac gtgtaatgat tatactgtaa atcaagaaga atacggaggt 3240

gcgtacactt ctcgtaatcg aggatataac gaagctcctt ccgtaccagc tgattatgcg 3300

tcagtctatg aagaaaaatc gtatacagat ggacgtagag agaatccttg tgaatttaac 3360

agagggtata gggattacac gccactacca gttggttatg tgacaaaaga attagaatac 3420

ttcccagaaa ccgataaggt atggattgag attggagaaa cggaaggaac atttatcgtg 3480

gacagcgtgg aattactcct tatggaggaa taa 3513

<210> 52
<211> 3513
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự nucleotit tổng hợp được thiết kế để biểu hiện ở tế bào thực vật ghi mã TIC836.

<400> 52

atggagaaca acatccagaa ccagtgcgtg ccctacaact gcctgaacaa ccctgaggtt 60

gagatcctga acgaggagcg tagcaccggt aggctcccgc tagacatctc cctgagcctg 120

acccgcttcc tccttagtga gttcgtgccc ggcgtgggcg tggccttcgg cctcttcgac 180

ctcatctggg gcttcatcac tccttccgac tggtccctct tcctccttca gattgagcaa 240

ctgatcgagc agcgcatcga gacccttgag cgcaaccgcg ccatcaccac tctcagaggt 300

ctcgccgact cctacgaaat ctacatcgag gcactccgtg agtgggaggc caacccgaac 360

aatgcccagc tccgcgagga cgtgaggatc agattcgcca acaccgacga tgccctcatc 420

accgccatca acaatttcac cctcacctcc ttcgagatcc ctcttctgtc tgtgtacgtt 480

caagctgcta accttcacct ttccctcctg cgcgacgccg tgagcttcgg ccagggctgg 540

ggcctcgaca tcgccaccgt gaacaatcac tacaaccgcc tcatcaacct catccaccgc 600

tacaccaagc actgccttga cacctacaac cagggccttg agaacctccg tggcaccaac 660

acccgccagt gggcccgctt caaccagttc cgcagagacc tcaccctcac cgtgctcgac 720

atcgtggcac tcttcccaaa ctacgacgtg cgtacctacc ctatccagac ctccagccag 780

ctcaccaggg aaatctacac ctccagcgtg atcgaggact ctcctgtgtc cgccaacatc 840

cctaacggct tcaaccgcgc cgagttcggc gtgcgccctc ctcacctcat ggacttcatg 900

aactccctct tcgtcactgc cgagaccgtg cgctcccaga ccgtgtgggg cggtcacctc 960

gtgtccagcc gtaacaccgc tggcaacagg atcaacttcc cgtcctacgg cgtgttcaac 1020

ccaggcggtg ccatctggat cgccgatgaa gaccctcgtc ctttctaccg taccctgtcc 1080

gaccctgtgt tcgtgcgtgg cggtttcggc aaccctcact acgtgctggg cctgcgtggc 1140

gtggccttcc agcaaaccgg caccaaccac accaggacgt tccgtaactc cggcaccatc 1200

gacagtcttg acgagatccc tccgcaagac aactccggtg caccttggaa cgactactcc 1260

cacgtgctga accacgtgac cttcgtgagg tggcctggcg aaatctccgg ctccgactcc 1320

tggagggctc ctatgttcag ttggacccac aggagcgcta cgcctaccaa caccatcgac 1380

cctgagcgta tcactcagat ccctctgact aagagcacta acctgggcag cggcactagc 1440

gtggtcaagg gccctggctt cactggcggt gacatcctga ggcggactag ccctggccag 1500

atcagcactc tgagggtgaa catcactgct ccgctgagcc agcgttacag ggtcagaatc 1560

cgttacgctt ctactactaa ccttcagttc cacactagca tcgacggccg tccgatcaac 1620

cagggcaact tctctgctac tatgagttct ggcagtaacc tccagtctgg tagtttccgg 1680

actgtcggtt tcactacgcc gttcaacttc tccaacggta gttctgtctt cactctgtct 1740

gctcacgtgt tcaactctgg caacgaggtg tacatcgacc ggatcgagtt cgtccctgct 1800

gaggtgacgt tcgaggccga gtacgacctg gagcgggctc agaaggctgt caacgctctg 1860

ttcacttcta ctaaccagct tggtttgaag actaacgtga ccgactacca cattgatcaa 1920

gtcagtaacc tggtcacgta cctgtctgac gagttctgtc ttgacgagaa gcgggagctg 1980

tctgagaagg tcaagcacgc taagcggctg tctgacgagc ggaacctgct tcaagacagt 2040

aacttcaagg acattaaccg ccagcctgag cgtggttggg gagggtccac gggtattacg 2100

attcaaggag gtgacgatgt ctttaaggag aactatgtga cgctttcggg tacgtttgat 2160

gagtgctatc caacgtacct ttaccagaag attgacgagt cgaagctgaa ggctttcact 2220

cgttaccagc ttcgtggtta cattgaggac tcgcaagacc tcgaaatcta cctcattcgt 2280

tacaacgcta agcacgagac tgtcaacgtc cctggtacgg gtagtctttg gccgctttct 2340

gctcagtcgc cgattggcaa gtgtggcgag ccgaaccgtt gcgctcctca cttggagtgg 2400

aacccggatc tcgattgctc gtgccgtgac ggtgagaagt gcgcgcacca tagtcatcac 2460

tttagccttg acattgatgt cggttgcacg gatcttaacg aggatcttgg agtctgggtg 2520

attttcaaga tcaaaactca ggatgggcac gcgcgtcttg ggaatcttga gttcctggag 2580

gagaagccac ttgtcggtga ggcgcttgcg cgtgtcaagc gtgcggagaa gaaatggcgt 2640

gataagcgtg agaagttgga gtgggagacg aacatcgtgt acaaggaggc gaaggagtcg 2700

gtcgatgcgt tgtttgtcaa tagtcaatac gatcaattgc aagcggatac gaacatcgca 2760

atgattcatg cggcagataa gcgtgtccat tcgattcgtg aggcgtactt gccagagttg 2820

tcggtcatcc caggagttaa tgcggcaatc tttgaggaat tggagggcag aatcttcacg 2880

gcgttctcgt tgtacgatgc aagaaatgtt attaagaatg gagatttcaa caatgggttg 2940

tcatgctgga atgttaaggg tcacgttgat gttgaagaac agaacaacca gagatcagtg 3000

ttggttgtac cagagtggga ggcagaggtt tcacaagagg tgagagtttg cccaggcaga 3060

ggctacatct tgagagttac agcatacaaa gagggatacg gcgagggatg tgttacaatc 3120

cacgaaatcg agaacaatac cgatgagcta aagttctcaa attgtgttga ggaggagatc 3180

tacccgaaca acacggttac ttgtaatgat tacacagtga accaggagga gtatggtggt 3240

gcatacacat caagaaatag aggctacaat gaagcaccat cagttccagc agattatgcc 3300

tcagtttatg aggagaagtc atacacagat ggacgacgtg agaatccatg tgagttcaat 3360

cgaggatacc gagattacac accactacca gttggatacg ttacaaagga actagaatac 3420

ttcccagaaa cagataaagt atggatagag atcggagaaa cagaaggaac attcatcgtt 3480

gattcagtag aactactact tatggaagaa tga 3513

<210> 53
<211> 1170
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC836.

<400> 53

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

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

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
        35                  40                  45

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

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

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

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

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
        115                 120                 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
    130                 135                 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145                 150                 155                 160

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

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
            180                 185                 190

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

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
    210                 215                 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225                 230                 235                 240

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

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

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

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
```

```
    290                 295                 300

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

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

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

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

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

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

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

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

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

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

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

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

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

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

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

Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu
                645                 650                 655

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

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

Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly
    690                 695                 700

Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp
705                 710                 715                 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
                725                 730                 735

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

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

Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro
    770                 775                 780

Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp
785                 790                 795                 800

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

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

Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp
        835                 840                 845

Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu
    850                 855                 860

Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg
865                 870                 875                 880

Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu
                885                 890                 895

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

Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg
        915                 920                 925

Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro
    930                 935                 940
```

Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr
945 950 955 960

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

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

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

Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile
1010 1015 1020

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

Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser
1040 1045 1050

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

Asn Asp Tyr Thr Val Asn Gln Glu Glu Tyr Gly Gly Ala Tyr Thr
1070 1075 1080

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

Tyr Ala Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg
1100 1105 1110

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

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

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

Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1160 1165 1170

<210> 54
<211> 3588
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC713

<400> 54
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atacgattgg accaaataga attacacaaa taccattggt aaaagcactg 1440

aatcttcatt caggtgttac tgttgttgga gggccaggat ttacaggtgg ggatatcctt 1500

cgtagaacaa atacgggtac atttggagat atacgattaa atattaatgt gccattatcc 1560

caaagatatc gcgtaaggat tcgttatgct tctactacag atttacaatt tttcacgaga 1620

attaatggaa ccactgttaa tattggtaat ttctcaagaa ctatgaatag gggggataat 1680

ttagaatata gaagttttag aactgcagga tttagtactc cttttaattt tttaaatgcc 1740

caaagcacat tcacattggg tgctcagagt ttttcaaatc aggaagttta tatagataga 1800

gtcgaatttg ttccagcaga ggtaacattt gaggcagaat atgatttaga aagagcacaa 1860

aaggcggtga atgctctgtt tacttctaca aatccaagaa gattgaaaac agatgtgaca 1920

gattatcata ttgaccaagt gtccaatatg gtggcatgtt tatcagatga attttgcttg 1980

gatgagaagc gagaattatt tgagaaagtg aaatatgcga agcgactcag tgatgaaaga 2040

aacttactcc aagatccaaa cttcacattc atcagtgggc aattaagttt cgcatccatc 2100

gatggacaat caaacttccc ctctattaat gagctatctg aacatggatg gtggggaagt 2160

gcgaatgtta ccattcagga agggaatgac gtatttaaag agaattacgt cacactaccg 2220

ggtactttta atgagtgtta tccaaattat ttatatcaaa aaataggaga gtcagaatta 2280

aaagcttata cgcgctatca attaagaggg tatattgaag atagtcaaga tctagagatt 2340

tatttaattc gttacaatgc aaagcatgaa acattggatg ttccaggtac cgattcccta 2400

tggccgcttt cagttgaaag cccaatcgga aggtgcggag aaccaaatcg atgcgcacca 2460

cattttgaat ggaatcctga tctagattgt tcctgcagag atggagaaag atgtgcgcat 2520

cattcccatc atttcacttt ggatattgat gttgggtgca cagacttgca tgagaaccta 2580

ggcgtgtggg tggtattcaa gattaagacg caggaaggtt atgcaagatt aggaaatctg 2640

gaatttatcg aagagaaacc attaattgga gaagcactgt ctcgtgtgaa gagagcggaa 2700

aaaaaatgga gagacaaacg ggaaaaacta caattggaaa caaaacgagt atatacagag 2760

gcaaaagaag ctgtggatgc tttattcgta gattctcaat atgatcaatt acaagcggat 2820

acaaacattg gcatgattca tgcggcagat aaacttgttc atcgaattcg agaggcgtat 2880

ctttcagaat tacctgttat cccaggtgta aatgcggaaa tttttgaaga attagaaggt 2940

cacattatca ctgcaatgtc cttatacgat gcgagaaatg tcgttaaaaa tggtgatttt 3000

aataatggat taacatgttg gaatgtaaaa gggcatgtag atgtacaaca gagccatcat 3060

cgttctgacc ttgttatccc agaatgggaa gcagaagtgt cacaagcagt tcgcgtctgt 3120

ccggggcgtg gctatatcct tcgtgtcaca gcgtacaaag agggatatgg agagggctgc 3180

gtaacgatcc atgaaatcga gaacaataca gacgaactaa aatttaaaaa ctgtgaagaa 3240

gaggaagtgt atccaacgga tacaggaacg tgtaatgatt atactgcaca ccaaggtaca 3300

gcagcatgta attcccgtaa tgctggatat gaggatgcat atgaagttga tactacagca 3360

tctgttaatt acaaaccgac ttatgaagaa gaaacgtata cagatgtacg aagagataat 3420

cattgtgaat atgacagagg gtatgtgaat tatccaccag taccagctgg ttatgtgaca 3480

aaagaattag aatacttccc agaaacagat acagtatgga ttgagattgg agaaacggaa 3540

ggaaagttta ttgtagatag cgtggaacta ctcctcatgg aagaatag 3588

<210> 55
<211> 1195
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC713.

<400> 55

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln
305                 310                 315                 320
```

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Thr Ile Gly Pro Asn Arg Ile Thr Gln Ile Pro Leu Val Lys Ala Leu
465                 470                 475                 480

Asn Leu His Ser Gly Val Thr Val Val Gly Gly Pro Gly Phe Thr Gly
                485                 490                 495

Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Thr Phe Gly Asp Ile Arg
            500                 505                 510

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

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

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

Leu Glu Tyr Arg Ser Phe Arg Thr Ala Gly Phe Ser Thr Pro Phe Asn
                565                 570                 575

Phe Leu Asn Ala Gln Ser Thr Phe Thr Leu Gly Ala Gln Ser Phe Ser
            580                 585                 590

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

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

Ala Leu Phe Thr Ser Thr Asn Pro Arg Arg Leu Lys Thr Asp Val Thr
625                 630                 635                 640

Asp Tyr His Ile Asp Gln Val Ser Asn Met Val Ala Cys Leu Ser Asp
```

645 650 655

Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Phe Glu Lys Val Lys Tyr
660 665 670

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

Thr Phe Ile Ser Gly Gln Leu Ser Phe Ala Ser Ile Asp Gly Gln Ser
690 695 700

Asn Phe Pro Ser Ile Asn Glu Leu Ser Glu His Gly Trp Trp Gly Ser
705 710 715 720

Ala Asn Val Thr Ile Gln Glu Gly Asn Asp Val Phe Lys Glu Asn Tyr
725 730 735

Val Thr Leu Pro Gly Thr Phe Asn Glu Cys Tyr Pro Asn Tyr Leu Tyr
740 745 750

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

Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg
770 775 780

Tyr Asn Ala Lys His Glu Thr Leu Asp Val Pro Gly Thr Asp Ser Leu
785 790 795 800

Trp Pro Leu Ser Val Glu Ser Pro Ile Gly Arg Cys Gly Glu Pro Asn
805 810 815

Arg Cys Ala Pro His Phe Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys
820 825 830

Arg Asp Gly Glu Arg Cys Ala His His Ser His His Phe Thr Leu Asp
835 840 845

Ile Asp Val Gly Cys Thr Asp Leu His Glu Asn Leu Gly Val Trp Val
850 855 860

Val Phe Lys Ile Lys Thr Gln Glu Gly Tyr Ala Arg Leu Gly Asn Leu
865 870 875 880

Glu Phe Ile Glu Glu Lys Pro Leu Ile Gly Glu Ala Leu Ser Arg Val
885 890 895

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu
900 905 910

Glu Thr Lys Arg Val Tyr Thr Glu Ala Lys Glu Ala Val Asp Ala Leu
915 920 925

Phe Val Asp Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Gly
930 935 940

Met Ile His Ala Ala Asp Lys Leu Val His Arg Ile Arg Glu Ala Tyr
945 950 955 960

Leu Ser Glu Leu Pro Val Ile Pro Gly Val Asn Ala Glu Ile Phe Glu
965 970 975

Glu Leu Glu Gly His Ile Ile Thr Ala Met Ser Leu Tyr Asp Ala Arg
980 985 990

Asn Val Val Lys Asn Gly Asp Phe Asn Asn Gly Leu Thr Cys Trp Asn
995 1000 1005

Val Lys Gly His Val Asp Val Gln Gln Ser His His Arg Ser Asp
1010 1015 1020

Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Ala Val Arg
1025 1030 1035

Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys
1040 1045 1050

Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn
1055 1060 1065

Asn Thr Asp Glu Leu Lys Phe Lys Asn Cys Glu Glu Glu Glu Val
1070 1075 1080

Tyr Pro Thr Asp Thr Gly Thr Cys Asn Asp Tyr Thr Ala His Gln
1085 1090 1095

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

Tyr Glu Val Asp Thr Thr Ala Ser Val Asn Tyr Lys Pro Thr Tyr
1115 1120 1125

Glu Glu Glu Thr Tyr Thr Asp Val Arg Arg Asp Asn His Cys Glu
1130 1135 1140

Tyr Asp Arg Gly Tyr Val Asn Tyr Pro Pro Val Pro Ala Gly Tyr
1145 1150 1155

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Thr Val Trp
1160 1165 1170

Ile Glu Ile Gly Glu Thr Glu Gly Lys Phe Ile Val Asp Ser Val
1175 1180 1185

Glu Leu Leu Leu Met Glu Glu
1190 1195

<210> 56
<211> 3531
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC843.

<400> 56
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atacgattgg accaaataga attacacaaa taccattggt aaaagcactg 1440

aatcttcatt caggtgttac tgttgttgga gggccaggat ttacaggtgg ggatatcctt 1500

cgtagaacaa atacgggtac atttggagat atacgattaa atattaatgt gccattatcc 1560

caaagatatc gcgtaaggat tcgttatgct tctactacag atttacaatt tttcacgaga 1620

attaatggaa ccactgttaa tattggtaat ttctcaagaa ctatgaatag gggggataat 1680

ttagaatata gaagttttag aactgcagga tttagtactc cttttaattt tttaaatgcc 1740

caaagcacat tcacattggg tgctcagagt ttttcaaatc aggaagttta tatagataga 1800

gtcgaatttg ttccagcaga ggtaacattt gaggcagaat atgatttaga aagagcgcag 1860

aaggcggtga atgcgctgtt tacgtctaca aaccaactag ggctaaaaac aaatgtaacg 1920

gattatcata ttgatcaagt gtccaattta gttacgtatt tatcggatga attttgtctg 1980

gatgaaaagc gagaattgtc cgagaaagtc aaacatgcga agcgactcag tgatgaacgc 2040

aatttactcc aagattcaaa tttcaaagac attaataggc aaccagaacg tgggtggggc 2100

ggaagtacag ggattaccat ccaaggaggg gatgacgtat ttaaagaaaa ttacgtcaca 2160

ctatcaggta cctttgatga gtgctatcca acatatttgt atcaaaaaat cgatgaatca 2220

aaattaaaag cctttacccg ttatcaatta agagggtata tcgaagatag tcaagactta 2280

gaaatctatt taattcgcta caatgcaaaa catgaaacag taaatgtgcc aggtacgggt 2340

tccttatggc cgctttcagc ccaaagtcca atcggaaagt gtggagagcc gaatcgatgc 2400

gcgccacacc ttgaatggaa tcctgactta gattgttcgt gtagggatgg agaaaagtgt 2460

gcccatcatt cgcatcattt ctccttagac attgatgtag gatgtacaga cttaaatgag 2520

gacctaggtg tatgggtgat ctttaagatt aagacgcaag atgggcacgc aagactaggg 2580

aatctagagt ttctcgaaga gaaaccatta gtaggagaag cgctagctcg tgtgaaaaga 2640

gcggagaaaa aatggagaga caaacgtgaa aaattggaat gggaaacaaa tatcgtttat 2700

aaagaggcaa aagaatctgt agatgcttta tttgtaaact ctcaatatga tcaattacaa 2760

gcggatacga atattgccat gattcatgcg gcagataaac gtgttcatag cattcgagaa 2820

gcttatctgc ctgagctgtc tgtgattccg ggtgtcaatg cggctatttt tgaagaatta 2880

gaagggcgta ttttcactgc attctcccta tatgatgcga gaaatgtcat taaaaatggt 2940

gattttaata atggcttatc ctgctggaac gtgaaagggc atgtagatgt agaagaacaa 3000

aacaaccaac gttcggtcct tgttgttccg gaatgggaag cagaagtgtc acaagaagtt 3060

cgtgtctgtc cgggtcgtgg ctatatcctt cgtgtcacag cgtacaagga gggatatgga 3120

gaaggttgcg taaccattca tgagatcgag aacaatacag acgaactgaa gtttagcaac 3180

tgcgtagaag aggaaatcta tccaaataac acggtaacgt gtaatgatta tactgtaaat 3240

caagaagaat acggaggtgc gtacacttct cgtaatcgag gatataacga agctccttcc 3300

gtaccagctg attatgcgtc agtctatgaa gaaaaatcgt atacagatgg acgaagagag 3360

aatccttgtg aatttaacag agggtatagg gattacacgc cactaccagt tggttatgtg 3420

acaaaagaat tagaatactt cccagaaacc gataaggtat ggattgagat tggagaaacg 3480

gaaggaacat ttatcgtgga cagcgtggaa ttactcctta tggaggaata g 3531

<210> 57
<211> 1176
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC843.

<400> 57

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
```

```
                325                 330                 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Thr Ile Gly Pro Asn Arg Ile Thr Gln Ile Pro Leu Val Lys Ala Leu
465                 470                 475                 480

Asn Leu His Ser Gly Val Thr Val Val Gly Gly Pro Gly Phe Thr Gly
                485                 490                 495

Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Thr Phe Gly Asp Ile Arg
            500                 505                 510

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

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

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

Leu Glu Tyr Arg Ser Phe Arg Thr Ala Gly Phe Ser Thr Pro Phe Asn
                565                 570                 575

Phe Leu Asn Ala Gln Ser Thr Phe Thr Leu Gly Ala Gln Ser Phe Ser
            580                 585                 590

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

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

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

Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp
                645                 650                 655
```

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

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

Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly
    690                 695                 700

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

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

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

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

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

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

Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp
                805                 810                 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
            820                 825                 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
        835                 840                 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
    850                 855                 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
865                 870                 875                 880

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Met Glu Glu
1175

<210> 58
<211> 3501
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC862.

<400> 58
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatatta 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attactcaaa tccctgcagt gaagggaaac 1440

tttcttttta atggttctgt aatttcagga ccaggattta ctggtgggga cttagttaga 1500

tgtacgaatg gatctggatt aactttatat gttacaccgg caccggactt gacgtattct 1560

aaaacatata aaattcgaat tcgttatgct tctacatctc aggtgagatt tggaattgac 1620

ttaggcagtt acactcatag tatttcgtat ttcgataaaa cgatggataa aggaaataca 1680

ttaacgtata attcatttaa tttatcaagt gtcagcagac caattgaaat atcaggaggg 1740

aataaaatcg gggtatccgt cggaggtatt ggctctgggg atgaagttta tatagacaaa 1800

atcgaattta ttccaatgga aagagcgcag aaggcggtga atgcgctgtt tacgtctaca 1860

aaccaactag ggctaaaaac aaatgtaacg gattatcata ttgatcaagt gtccaattta 1920

gttacgtatt tatcggatga attttgtctg gatgaaaagc gagaattgtc cgagaaagtc 1980

aaacatgcga agcgactcag tgatgaacgc aatttactcc aagattcaaa tttcaaagac 2040

attaataggc aaccagaacg tgggtggggc ggaagtacag ggattaccat ccaaggaggg 2100

gatgacgtat ttaaagaaaa ttacgtcaca ctatcaggta cctttgatga gtgctatcca 2160

acatatttgt atcaaaaaat cgatgaatca aaattaaaag cctttacccg ttatcaatta 2220

agagggtata tcgaagatag tcaagactta gaaatctatt taattcgcta caatgcaaaa 2280

catgaaacag taaatgtgcc aggtacgggt tccttatggc cgctttcagc ccaaagtcca 2340

atcggaaagt gtggagagcc gaatcgatgc gcgccacacc ttgaatggaa tcctgactta 2400

gattgttcgt gtagggatgg agaaaagtgt gcccatcatt cgcatcattt ctccttagac 2460

attgatgtag gatgtacaga cttaaatgag gacctaggtg tatgggtgat ctttaagatt 2520

aagacgcaag atgggcacgc aagactaggg aatctagagt ttctcgaaga aaaaccatta 2580

gtaggagaag cgctagctcg tgtgaaaaga gcggagaaaa aatggagaga caaacgtgaa 2640

aaattggaat gggaaacaaa tatcgtttat aaagaggcaa aagaatctgt agatgcttta 2700

tttgtaaact ctcaatatga tcaattacaa gcggatacga atattgccat gattcatgcg 2760

gcagataaac gtgttcatag cattcgagaa gcttatctgc ctgagctgtc tgtgattccg 2820

ggtgtcaatg cggctatttt tgaagaatta gaagggcgta ttttcactgc attctcccta 2880

tatgatgcga gaaatgtcat taaaaatggt gattttaata atggcttatc ctgctggaac 2940

gtgaaagggc atgtagatgt agaagaacaa aacaaccaac gttcggtcct tgttgttccg 3000

gaatgggaag cagaagtgtc acaagaagtt cgtgtctgtc cgggtcgtgg ctatatcctt 3060

cgtgtcacag cgtacaagga gggatatgga gaaggttgcg taaccattca tgagatcgag 3120

aacaatacag acgaactgaa gtttagcaac tgcgtagagg aggaaatcta tccaaataac 3180

acggtaacgt gtaatgatta tactgtaaat caagaagaat acggaggtgc gtacacttct 3240

cgtaatcgag gatataacga agctccttcc gtaccagctg attatgcgtc agtctatgaa 3300

gaaaaatcgt atacagatgg acgtagagag aatccttgtg aatttaacag agggtatagg 3360

gattacacgc cactaccagt tggttatgtg acaaaagaat tagaatactt cccagaaacc 3420

gataaggtat ggattgagat tggagaaacg gaaggaacat ttatcgtgga cagcgtggaa 3480

ttactcctta tggaggaata a 3501

<210> 59
<211> 1166
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC862.

<400> 59

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

Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly

```
            20                  25                  30

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350
```

```
Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

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

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

Asp Leu Val Arg Cys Thr Asn Gly Ser Gly Leu Thr Leu Tyr Val Thr
            500                 505                 510

Pro Ala Pro Asp Leu Thr Tyr Ser Lys Thr Tyr Lys Ile Arg Ile Arg
        515                 520                 525

Tyr Ala Ser Thr Ser Gln Val Arg Phe Gly Ile Asp Leu Gly Ser Tyr
    530                 535                 540

Thr His Ser Ile Ser Tyr Phe Asp Lys Thr Met Asp Lys Gly Asn Thr
545                 550                 555                 560

Leu Thr Tyr Asn Ser Phe Asn Leu Ser Ser Val Ser Arg Pro Ile Glu
                565                 570                 575

Ile Ser Gly Gly Asn Lys Ile Gly Val Ser Val Gly Gly Ile Gly Ser
            580                 585                 590

Gly Asp Glu Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Met Glu Arg
        595                 600                 605

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

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

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

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

```
Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
        675                 680                 685

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
    690                 695                 700

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

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

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

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

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
    770                 775                 780

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

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

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

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

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

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

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

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

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

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

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

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

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

Gln Arg Ser Val Leu Val Val Pro  Glu Trp Glu Ala Glu  Val Ser Gln
```

995 1000 1005

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

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

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

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

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

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

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

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

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

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

Val Glu Leu Leu Leu Met Glu Glu
1160 1165

<210> 60
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099.

<400> 60
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 61
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099.

<400> 61

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
```

```
    370                 375                 380

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700
```

```
Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

Gln Arg  Ser Val Leu Val Val  Pro Glu Trp Glu Ala  Glu Val Ser
    1010                1015                1020
```

Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val
1025 1030 1035

Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His
1040 1045 1050

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

Glu Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr
1070 1075 1080

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 62
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 T507E.

<400> 62
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcga atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 63
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 T507E.

<400> 63

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
435 440 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465 470 475 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Glu Phe Gly Thr Leu Arg
500 505 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
530 535 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545 550 555 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
565 570 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
580 585 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
610 615 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
645 650 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
660 665 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
690 695 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe

```
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035
```

Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His
1040 1045 1050

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

Glu Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr
1070 1075 1080

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 64
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 R522K.

<400> 64
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaaaatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 65
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 R522K.

<400> 65

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

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

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

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

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile

```
65                  70                  75                  80

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720
```

```
Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
```

```
    1040                1045                1050

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

Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr
    1070                1075                1080

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

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

Val Tyr  Glu Glu Lys Ser Tyr  Thr Asp Gly Arg Arg  Glu Asn Pro
    1115                1120                1125

Cys Glu  Phe Asn Arg Gly Tyr  Arg Asp Tyr Thr Pro  Leu Pro Val
    1130                1135                1140

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

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

Ser Val  Glu Leu Leu Leu Met  Glu Glu
    1175                1180
```

<210> 66
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 K490S.

<400> 66

```
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa       60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg      120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta      180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt      240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta      300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat      360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc      420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta      480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa      540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt      600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga      660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta      720
```

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttatca ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 67
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 K490S.

<400> 67

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
```

```
                405                 410                 415

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735
```

```
Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr
            740                 745                 750

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050
```

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

Glu Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr
1070 1075 1080

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 68
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 T562R.

<400> 68
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctacgctacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 69
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 T526R.

<400> 69

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Arg Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr
```

```
            740                 745                 750

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050

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

Glu Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr
1070 1075 1080

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 70
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 S553R.

<400> 70
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggcgta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 71
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 S553R.

<400> 71

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

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

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

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

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

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

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu

```
            100                 105                 110

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

```
Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Arg Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr
            740                 745                 750
```

```
Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile
        755                 760                 765

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050

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

Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr
```

1070 1075 1080

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 72
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 G498D.

<400> 72
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg ggatatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 73
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 G498D.

<400> 73

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
```

```
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Asp Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile
        755                 760                 765
```

```
Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050

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

Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr
    1070                1075                1080
```

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

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 74
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 K490A.

<400> 74
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttagca ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg aatcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 75
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 K490A.

<400> 75

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445
```

```
Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
    450                 455                 460

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Glu Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
```

```
    770                 775                 780

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050

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

Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr
    1070                1075                1080

Thr Val  Asn Gln Glu Glu Tyr  Gly Gly Ala Tyr Thr  Ser Arg Asn
    1085                1090                1095
```

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

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 76
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1099 E564A.

<400> 76
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attacacaat taccattggt aaaggcatct 1440

gcacctgttt cgggtactac ggtcttaaaa ggtccaggat ttacaggagg gggtatactc 1500

cgaagaacaa ctaatggcac atttggaacg ttaagagtaa cagttaattc accattaaca 1560

caaagatatc gcgtaagagt tcgttttgct tcatcaggaa atttcagcat aaggatactg 1620

cgtggaaata cctctatagc ttatcaaaga tttgggagta caatgaacag aggacaggaa 1680

ctaacttacg catcatttgt cacaagtgag ttcactacta atcagagcga tctgcctttt 1740

acatttacac aagctcaaga aaatttaaca atccttgcag aaggtgttag caccggtagt 1800

gaatatttta tagatagaat tgaaatcatc cctgtgaacc cggcacgaga agcagaagag 1860

gatttagaaa gagcgcagaa ggcggtgaat gcgctgttta cgtctacaaa ccaactaggg 1920

ctaaaaacaa atgtaacgga ttatcatatt gatcaagtgt ccaatttagt tacgtattta 1980

tcggatgaat tttgtctgga tgaaaagcga gaattgtccg agaaagtcaa acatgcgaag 2040

cgactcagtg atgaacgcaa tttactccaa gattcaaatt tcaaagacat taataggcaa 2100

ccagaacgtg ggtggggcgg aagtacaggg attaccatcc aaggagggga tgacgtattt 2160

aaagaaaatt acgtcacact atcaggtacc tttgatgagt gctatccaac atatttgtat 2220

caaaaaatcg atgaatcaaa attaaaagcc tttacccgtt atcaattaag agggtatatc 2280

gaagatagtc aagacttaga aatctattta attcgctaca atgcaaaaca tgaaacagta 2340

aatgtgccag gtacgggttc cttatggccg ctttcagccc aaagtccaat cggaaagtgt 2400

ggagagccga atcgatgcgc gccacacctt gaatggaatc ctgacttaga ttgttcgtgt 2460

agggatggag aaaagtgtgc ccatcattcg catcatttct ccttagacat tgatgtagga 2520

tgtacagact taaatgagga cctaggtgta tgggtgatct ttaagattaa gacgcaagat 2580

gggcacgcaa gactagggaa tctagagttt ctcgaagaaa aaccattagt aggagaagcg 2640

ctagctcgtg tgaaaagagc ggagaaaaaa tggagagaca aacgtgaaaa attggaatgg 2700

gaaacaaata tcgtttataa agaggcaaaa gaatctgtag atgctttatt tgtaaactct 2760

caatatgatc aattacaagc ggatacgaat attgccatga ttcatgcggc agataaacgt 2820

gttcatagca ttcgagaagc ttatctgcct gagctgtctg tgattccggg tgtcaatgcg 2880

gctatttttg aagaattaga agggcgtatt ttcactgcat tctccctata tgatgcgaga 2940

aatgtcatta aaaatggtga ttttaataat ggcttatcct gctggaacgt gaaagggcat 3000

gtagatgtag aagaacaaaa caaccaacgt tcggtccttg ttgttccgga atgggaagca 3060

gaagtgtcac aagaagttcg tgtctgtccg ggtcgtggct atatccttcg tgtcacagcg 3120

tacaaggagg gatatggaga aggttgcgta accattcatg agatcgagaa caatacagac 3180

gaactgaagt ttagcaactg cgtagaggag gaaatctatc caaataacac ggtaacgtgt 3240

aatgattata ctgtaaatca agaagaatac ggaggtgcgt acacttctcg taatcgagga 3300

tataacgaag ctccttccgt accagctgat tatgcgtcag tctatgaaga aaaatcgtat 3360

acagatggac gtagagagaa tccttgtgaa tttaacagag ggtataggga ttacacgcca 3420

ctaccagttg gttatgtgac aaaagaatta gaatacttcc cagaaaccga taaggtatgg 3480

attgagattg gagaaacgga aggaacattt atcgtggaca gcgtggaatt actccttatg 3540

gaggaataa 3549

<210> 77
<211> 1182
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1099 E564A.

<400> 77

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

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

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

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

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

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

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

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

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

```
    130                 135                 140

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
    450                 455                 460
```

```
Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Leu Val Lys Ala Ser
465                 470                 475                 480

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

Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg
            500                 505                 510

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

Phe Ala Ser Ser Gly Asn Phe Ser Ile Arg Ile Leu Arg Gly Asn Thr
    530                 535                 540

Ser Ile Ala Tyr Gln Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu
545                 550                 555                 560

Leu Thr Tyr Ala Ser Phe Val Thr Ser Glu Phe Thr Thr Asn Gln Ser
                565                 570                 575

Asp Leu Pro Phe Thr Phe Thr Gln Ala Gln Glu Asn Leu Thr Ile Leu
            580                 585                 590

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

Ile Ile Pro Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Arg
    610                 615                 620

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

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
                645                 650                 655

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
            660                 665                 670

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

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
    690                 695                 700

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
705                 710                 715                 720

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
                725                 730                 735

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

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

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
    770                 775                 780
```

```
Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
785                 790                 795                 800

Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
                805                 810                 815

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
            820                 825                 830

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
        835                 840                 845

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
    850                 855                 860

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
865                 870                 875                 880

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

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

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
        915                 920                 925

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
    930                 935                 940

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

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

Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu
            980                 985                 990

Ser Cys Trp Asn Val Lys Gly His  Val Asp Val Glu Glu  Gln Asn Asn
        995                 1000                1005

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

Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val
    1025                1030                1035

Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His
    1040                1045                1050

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

Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr
    1070                1075                1080

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

Arg Gly  Tyr Asn Glu Ala Pro  Ser Val Pro Ala Asp  Tyr Ala Ser
```

1100 1105 1110

Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro
1115 1120 1125

Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val
1130 1135 1140

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

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

Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180

<210> 78
<211> 3534
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1103.

<400> 78
atggataaca atccgaacat caatgaatgc attccttata attgtttaag taaccctgaa 60

gtagaagtat taggtggaga aagaatagaa actggttaca ccccaatcga tatttccttg 120

tcgctaacgc aatttctttt gagtgaattt gttcccggtg ctggatttgt gttaggacta 180

gttgatataa tatggggaat ttttggtccc tctcaatggg acgcatttct tgtacaaatt 240

gaacagttaa ttaaccaaag aatagaagaa ttcgctagga accaagccat ttctagatta 300

gaaggactaa gcaatcttta tcaaatttac gcagaatctt ttagagagtg ggaagcagat 360

cctactaatc cagcattaag agaagagatg cgtattcaat tcaatgacat gaacagtgcc 420

cttacaaccg ctattcctct tttggcagtt caaaattatc aagttcctct tttatcagta 480

tatgttcaag ctgcaaattt acatttatca gttttgagag atgtttcagt gtttggacaa 540

aggtggggat ttgatgccgc gactatcaat agtcgttata atgatttaac taggcttatt 600

ggcaactata cagattatgc tgtacgctgg tacaatacgg gattagaacg tgtatgggga 660

ccggattcta gagattgggt aaggtataat caatttagaa gagaattaac actaactgta 720

ttagatatcg ttgctctgtt cccgaattat gatagtagaa gatatccaat tcgaacagtt 780

tcccaattaa caagagaaat ttatacaaac ccagtattag aaaattttga tggtagtttt 840

cgaggctcgg ctcagggcat agaaagaagt attaggagtc cacatttgat ggatatactt 900

aacagtataa ccatctatac ggatgctcat aggggttatt attattggtc agggcatcaa 960

ataatggctt ctcctgtcgg tttttcgggg ccagaattca cgtttccgct atatggaacc 1020

atgggaaatg cagctccaca acaacgtatt gttgctcaac taggtcaggg cgtgtataga 1080

acattatcgt ccactttata tagaagacct tttaatatag ggataaataa tcaacaacta 1140

tctgttcttg acgggacaga atttgcttat ggaacctcct caaatttgcc atccgctgta 1200

tacagaaaaa gcggaacggt agattcgctg gatgaaatac cgccacagaa taacaacgtg 1260

ccacctaggc aaggatttag tcatcgatta agccatgttt caatgtttcg ttcaggcttt 1320

agtaatagta gtgtaagtat aataagagct cctatgttct cttggataca tcgtagtgct 1380

gaatttaata atataattgc atcggatagt attactcaac tacccatggt aaaggcaaat 1440

gctttacttt caggtacttc cgtgataaag ggtcctggat ctacaggggg agatatatta 1500

agaagaacaa gtgttggcag agtaggaaat tttaaggtta acgtaaatgg accattaaca 1560

caacgatatc ttgtaagaat tcgttatgct tctacaactg atctagactt ttacgtctat 1620

cgtggaggta ccactgtaag taattataga tttaacaaaa caatgaataa gggagcttct 1680

ttaacttacg acatttttaa atttgcaagt ttttctaccc ctttcacatt tacaaaaaca 1740

caagatgaat taggaatatc tatacaaaac tttagtagcg gtgaagaagt gtatatagac 1800

aggattgaag tcattccagt aggtacgaca tatgaagcgg aaacggattt agaaagagcg 1860

cagaaggcgg tgaatgcgct gtttacgtct acaaaccaac tagggctaaa aacaaatgta 1920

acggattatc atattgatca agtgtccaat ttagttacgt atttatcgga tgaattttgt 1980

ctggatgaaa agcgagaatt gtccgagaaa gtcaaacatg cgaagcgact cagtgatgaa 2040

cgcaatttac tccaagattc aaatttcaaa gacattaata ggcaaccaga acgtgggtgg 2100

ggcggaagta cagggattac catccaagga ggggatgacg tatttaaaga aaattacgtc 2160

acactatcag gtacctttga tgagtgctat ccaacatatt tgtatcaaaa aatcgatgaa 2220

tcaaaattaa aagcctttac ccgttatcaa ttaagagggt atatcgaaga tagtcaagac 2280

ttagaaatct atttaattcg ctacaatgca aaacatgaaa cagtaaatgt gccaggtacg 2340

ggttccttat ggccgctttc agcccaaagt ccaatcggaa agtgtggaga gccgaatcga 2400

tgcgcgccac accttgaatg gaatcctgac ttagattgtt cgtgtaggga tggagaaaag 2460

tgtgcccatc attcgcatca tttctcctta gacattgatg taggatgtac agacttaaat 2520

gaggacctag gtgtatgggt gatctttaag attaagacgc aagatgggca cgcaagacta 2580

gggaatctag agtttctcga agaaaaacca ttagtaggag aagcgctagc tcgtgtgaaa 2640

agagcggaga aaaaatggag agacaaacgt gaaaaattgg aatgggaaac aaatatcgtt 2700

tataaagagg caaaagaatc tgtagatgct ttatttgtaa actctcaata tgatcaatta 2760

caagcggata cgaatattgc catgattcat gcggcagata aacgtgttca tagcattcga 2820

gaagcttatc tgcctgagct gtctgtgatt ccgggtgtca atgcggctat ttttgaagaa 2880

ttagaagggc gtattttcac tgcattctcc ctatatgatg cgagaaatgt cattaaaaat 2940

ggtgatttta ataatggctt atcctgctgg aacgtgaaag ggcatgtaga tgtagaagaa 3000

caaaacaacc aacgttcggt ccttgttgtt ccggaatggg aagcagaagt gtcacaagaa 3060

gttcgtgtct gtccgggtcg tggctatatc cttcgtgtca cagcgtacaa ggagggatat 3120

ggagaaggtt gcgtaaccat tcatgagatc gagaacaata cagacgaact gaagtttagc 3180

aactgcgtag aggaggaaat ctatccaaat aacacggtaa cgtgtaatga ttatactgta 3240

aatcaagaag aatacggagg tgcgtacact tctcgtaatc gaggatataa cgaagctcct 3300

tccgtaccag ctgattatgc gtcagtctat gaagaaaaat cgtatacaga tggacgtaga 3360

gagaatcctt gtgaatttaa cagagggtat agggattaca cgccactacc agttggttat 3420

gtgacaaaag aattagaata cttcccagaa accgataagg tatggattga gattggagaa 3480

acggaaggaa catttatcgt ggacagcgtg gaattactcc ttatggagga ataa 3534

<210> 79
<211> 1177
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1103.

<400> 79

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
            340                 345                 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
        355                 360                 365

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

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

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

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

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
        435                 440                 445

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

Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro Met Val Lys Ala Asn
465                 470                 475                 480
```

```
Ala Leu Leu Ser Gly Thr Ser Val Ile Lys Gly Pro Gly Ser Thr Gly
                485                 490                 495

Gly Asp Ile Leu Arg Arg Thr Ser Val Gly Arg Val Gly Asn Phe Lys
            500                 505                 510

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

Tyr Ala Ser Thr Thr Asp Leu Asp Phe Tyr Val Tyr Arg Gly Gly Thr
    530                 535                 540

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

Leu Thr Tyr Asp Ile Phe Lys Phe Ala Ser Phe Ser Thr Pro Phe Thr
                565                 570                 575

Phe Thr Lys Thr Gln Asp Glu Leu Gly Ile Ser Ile Gln Asn Phe Ser
            580                 585                 590

Ser Gly Glu Glu Val Tyr Ile Asp Arg Ile Glu Val Ile Pro Val Gly
        595                 600                 605

Thr Thr Tyr Glu Ala Glu Thr Asp Leu Glu Arg Ala Gln Lys Ala Val
    610                 615                 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
625                 630                 635                 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
                645                 650                 655

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
            660                 665                 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
        675                 680                 685

Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
    690                 695                 700

Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705                 710                 715                 720

Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
                725                 730                 735

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
            740                 745                 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
        755                 760                 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
    770                 775                 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785                 790                 795                 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
```

```
                805                 810                 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
            820                 825                 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
        835                 840                 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
    850                 855                 860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
865                 870                 875                 880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
                885                 890                 895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
            900                 905                 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
        915                 920                 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
    930                 935                 940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
945                 950                 955                 960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
                965                 970                 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
            980                 985                 990

Lys Gly His Val Asp Val Glu Glu  Gln Asn Asn Gln Arg  Ser Val Leu
        995                 1000                1005

Val Val  Pro Glu Trp Glu Ala  Glu Val Ser Gln Glu  Val Arg Val
    1010                1015                1020

Cys Pro  Gly Arg Gly Tyr Ile  Leu Arg Val Thr Ala  Tyr Lys Glu
    1025                1030                1035

Gly Tyr  Gly Glu Gly Cys Val  Thr Ile His Glu Ile  Glu Asn Asn
    1040                1045                1050

Thr Asp  Glu Leu Lys Phe Ser  Asn Cys Val Glu Glu  Glu Ile Tyr
    1055                1060                1065

Pro Asn  Asn Thr Val Thr Cys  Asn Asp Tyr Thr Val  Asn Gln Glu
    1070                1075                1080

Glu Tyr  Gly Gly Ala Tyr Thr  Ser Arg Asn Arg Gly  Tyr Asn Glu
    1085                1090                1095

Ala Pro  Ser Val Pro Ala Asp  Tyr Ala Ser Val Tyr  Glu Glu Lys
    1100                1105                1110

Ser Tyr  Thr Asp Gly Arg Arg  Glu Asn Pro Cys Glu  Phe Asn Arg
    1115                1120                1125
```

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
1160 1165 1170

Leu Met Glu Glu
1175

<210> 80
<211> 3597
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC1101.

<400> 80
atgaaaaaca gtatcaaatt atcagaactt tggtatttca atgaaagaaa atggaggtat 60

tttatggaga tagtgaataa tcagaatcaa tgcgtgcctt ataattgttt gaataatccc 120

gaaatcgaaa tattagaagg cggaagaata tcagttggta ataccccaat tgatatttct 180

ctttcgctta ctcagtttct tttgagtgaa tttgtcccag gtgcggggtt tgtattagga 240

ttaattgatt taatatgggg atttgtaggt ccttcccaat gggacgcatt tcttgctcaa 300

gtggaacagt taattaacca aagaatagca gaagctgtaa gaaatacagc aattcaggaa 360

ttagagggaa tggcacgggt ttatagaacc tatgctactg cttttgctga gtgggaaaaa 420

gctcctgatg acccagagct aagagaagca ctacgtacac aatttacagc aactgagact 480

tatataagtg gaagaatatc cgttttaaaa attcaaactt ttgaagtaca gctgttatca 540

gtgtttgccc aagctgcaaa tttacattta tctttattaa gagacgttgt gtttttgggg 600

caaagatggg gtttttcaac gacaaccgta aataattact acaatgattt aacagaaggg 660

attagtacct atacagatta tgctgtacgc tggtacaata cgggattaga acgtgtatgg 720

ggaccggatt ctagagattg ggtaaggtat aatcaattta gaagagaatt aacactaact 780

gtattagata tcgttgctct gttcccgaat tatgatagta gaagatatcc aattcgaaca 840

gtttcccaat taacaagaga aatttataca aacccagtat tagaaaattt tgatggtagt 900

tttcgaggct cggctcaggg catagaaaga agtattagga gtccacattt gatggatata 960

cttaacagta taaccatcta tacggatgct cataggggtt attattattg gtcagggcat 1020

caaataatgg cttctcctgt cggtttttcg gggccagaat tcacgtttcc gctatatgga 1080

accatgggaa atgcagctcc acaacaacgt attgttgctc aactaggtca gggcgtgtat 1140

agaacattat cgtccacttt atatagaaga ccttttaata tagggataaa taatcaacaa 1200

ctatctgttc ttgacgggac agaatttgct tatggaacct cctcaaattt gccatccgct 1260

gtatacagaa aaagcggaac ggtagattcg ctggatgaaa taccgccaca gaataacaac 1320

gtgccaccta ggcaaggatt tagtcatcga ttaagccatg tttcaatgtt tcgttcaggc 1380

tttagtaata gtagtgtaag tataataaga gctcctatgt tctcttggat acatcgtagt 1440

gctgaattta ataatataat tgcatcggat agtattactc aactacccat ggtaaaggca 1500

aatgctttac tttcaggtac ttccgtgata aagggtcctg gatctacagg gggagatata 1560

ttaagaagaa caagtgttgg cagagtagga aattttaagg ttaacgtaaa tggaccatta 1620

acacaacgat atcttgtaag aattcgttat gcttctacaa ctgatctaga cttttacgtc 1680

tatcgtggag gtaccactgt aagtaattat agatttaaca aaacaatgaa taagggagct 1740

tctttaactt acgacatttt taaatttgca agtttttcta ccccttcac atttacaaaa 1800

acacaagatg aattaggaat atctatacaa aactttagta gcggtgaaga agtgtatata 1860

gacaggattg aagtcattcc agtaggtacg acatatgaag cggaaacgga tttagaaaga 1920

gcgcagaagg cggtgaatgc gctgtttacg tctacaaacc aactagggct aaaaacaaat 1980

gtaacggatt atcatattga tcaagtgtcc aatttagtta cgtatttatc ggatgaattt 2040

tgtctggatg aaaagcgaga attgtccgag aaagtcaaac atgcgaagcg actcagtgat 2100

gaacgcaatt tactccaaga ttcaaatttc aaagacatta ataggcaacc agaacgtggg 2160

tggggcggaa gtacagggat taccatccaa ggaggggatg acgtatttaa agaaaattac 2220

gtcacactat caggtacctt tgatgagtgc tatccaacat atttgtatca aaaaatcgat 2280

gaatcaaaat taaaagcctt tacccgttat caattaagag ggtatatcga agatagtcaa 2340

gacttagaaa tctatttaat tcgctacaat gcaaaacatg aaacagtaaa tgtgccaggt 2400

acgggttcct tatggccgct ttcagcccaa agtccaatcg gaaagtgtgg agagccgaat 2460

cgatgcgcgc cacaccttga atggaatcct gacttagatt gttcgtgtag ggatggagaa 2520

aagtgtgccc atcattcgca tcatttctcc ttagacattg atgtaggatg tacagactta 2580

aatgaggacc taggtgtatg ggtgatcttt aagattaaga cgcaagatgg gcacgcaaga 2640

ctagggaatc tagagtttct cgaagagaaa ccattagtag gagaagcgct agctcgtgtg 2700

aaaagagcgg agaaaaaatg gagagacaaa cgtgaaaaat tggaatggga aacaaatatc 2760

gtttataaag aggcaaaaga atctgtagat gctttatttg taaactctca atatgatcaa 2820

ttacaagcgg atacgaatat tgccatgatt catgcggcag ataaacgtgt tcatagcatt 2880

cgagaagctt atctgcctga gctgtctgtg attccgggtg tcaatgcggc tatttttgaa 2940

gaattagaag ggcgtatttt cactgcattc tccctatatg atgcgagaaa tgtcattaaa 3000

aatggtgatt ttaataatgg cttatcctgc tggaacgtga aagggcatgt agatgtagaa 3060

gaacaaaaca accaacgttc ggtccttgtt gttccggaat gggaagcaga agtgtcacaa 3120

gaagttcgtg tctgtccggg tcgtggctat atccttcgtg tcacagcgta caaggaggga 3180

tatggagaag gttgcgtaac cattcatgag atcgagaaca atacagacga actgaagttt 3240

agcaactgcg tagaagagga aatctatcca aataacacgg taacgtgtaa tgattatact 3300

gtaaatcaag aagaatacgg aggtgcgtac acttctcgta atcgaggata taacgaagct 3360

ccttccgtac cagctgatta tgcgtcagtc tatgaagaaa aatcgtatac agatggacga 3420

agagagaatc cttgtgaatt taacagaggg tatagggatt acacgccact accagttggt 3480

tatgtgacaa aagaattaga atacttccca gaaaccgata aggtatggat tgagattgga 3540

gaaacggaag gaacatttat cgtggacagc gtggaattac tccttatgga ggaatag 3597

<210> 81
<211> 1198
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC1101.

<400> 81

Met Lys Asn Ser Ile Lys Leu Ser Glu Leu Trp Tyr Phe Asn Glu Arg
1 5 10 15

Lys Trp Arg Tyr Phe Met Glu Ile Val Asn Asn Gln Asn Gln Cys Val
20 25 30

Pro Tyr Asn Cys Leu Asn Asn Pro Glu Ile Glu Ile Leu Glu Gly Gly
35 40 45

Arg Ile Ser Val Gly Asn Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr
50 55 60

Gln Phe Leu Leu Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly
65 70 75 80

Leu Ile Asp Leu Ile Trp Gly Phe Val Gly Pro Ser Gln Trp Asp Ala
85 90 95

Phe Leu Ala Gln Val Glu Gln Leu Ile Asn Gln Arg Ile Ala Glu Ala
100 105 110

Val Arg Asn Thr Ala Ile Gln Glu Leu Glu Gly Met Ala Arg Val Tyr
115 120 125

Arg Thr Tyr Ala Thr Ala Phe Ala Glu Trp Glu Lys Ala Pro Asp Asp
130 135 140

Pro Glu Leu Arg Glu Ala Leu Arg Thr Gln Phe Thr Ala Thr Glu Thr
145 150 155 160

Tyr Ile Ser Gly Arg Ile Ser Val Leu Lys Ile Gln Thr Phe Glu Val

```
                 165                 170                 175

Gln Leu Leu Ser Val Phe Ala Gln Ala Ala Asn Leu His Leu Ser Leu
            180                 185                 190

Leu Arg Asp Val Val Phe Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr
        195                 200                 205

Thr Val Asn Asn Tyr Tyr Asn Asp Leu Thr Glu Gly Ile Ser Thr Tyr
    210                 215                 220

Thr Asp Tyr Ala Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp
225                 230                 235                 240

Gly Pro Asp Ser Arg Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu
                245                 250                 255

Leu Thr Leu Thr Val Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp
            260                 265                 270

Ser Arg Arg Tyr Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile
        275                 280                 285

Tyr Thr Asn Pro Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser
    290                 295                 300

Ala Gln Gly Ile Glu Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile
305                 310                 315                 320

Leu Asn Ser Ile Thr Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr
                325                 330                 335

Trp Ser Gly His Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro
            340                 345                 350

Glu Phe Thr Phe Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln
        355                 360                 365

Gln Arg Ile Val Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser
    370                 375                 380

Ser Thr Leu Tyr Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln
385                 390                 395                 400

Leu Ser Val Leu Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn
                405                 410                 415

Leu Pro Ser Ala Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp
            420                 425                 430

Glu Ile Pro Pro Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser
        435                 440                 445

His Arg Leu Ser His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser
    450                 455                 460

Ser Val Ser Ile Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser
465                 470                 475                 480

Ala Glu Phe Asn Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Leu Pro
                485                 490                 495
```

```
Met Val Lys Ala Asn Ala Leu Leu Ser Gly Thr Ser Val Ile Lys Gly
            500                 505                 510

Pro Gly Ser Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Val Gly Arg
        515                 520                 525

Val Gly Asn Phe Lys Val Asn Val Asn Gly Pro Leu Thr Gln Arg Tyr
    530                 535                 540

Leu Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu Asp Phe Tyr Val
545                 550                 555                 560

Tyr Arg Gly Gly Thr Thr Val Ser Asn Tyr Arg Phe Asn Lys Thr Met
                565                 570                 575

Asn Lys Gly Ala Ser Leu Thr Tyr Asp Ile Phe Lys Phe Ala Ser Phe
            580                 585                 590

Ser Thr Pro Phe Thr Phe Thr Lys Thr Gln Asp Glu Leu Gly Ile Ser
        595                 600                 605

Ile Gln Asn Phe Ser Ser Gly Glu Glu Val Tyr Ile Asp Arg Ile Glu
    610                 615                 620

Val Ile Pro Val Gly Thr Thr Tyr Glu Ala Glu Thr Asp Leu Glu Arg
625                 630                 635                 640

Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly
                645                 650                 655

Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu
            660                 665                 670

Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu
        675                 680                 685

Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu
    690                 695                 700

Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly
705                 710                 715                 720

Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe
                725                 730                 735

Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro
            740                 745                 750

Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr
        755                 760                 765

Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile
    770                 775                 780

Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly
785                 790                 795                 800

Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys
                805                 810                 815
```

```
Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu
            820                 825                 830

Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His Ser His His
        835                 840                 845

Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu
    850                 855                 860

Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg
865                 870                 875                 880

Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala
                885                 890                 895

Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu
            900                 905                 910

Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser
        915                 920                 925

Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp
    930                 935                 940

Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser Ile
945                 950                 955                 960

Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala
                965                 970                 975

Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu
            980                 985                 990

Tyr Asp Ala Arg Asn Val Ile Lys  Asn Gly Asp Phe Asn  Asn Gly Leu
        995                 1000                1005

Ser Cys  Trp Asn Val Lys Gly  His Val Asp Val Glu  Glu Gln Asn
    1010                1015                1020

Asn Gln  Arg Ser Val Leu Val  Val Pro Glu Trp Glu  Ala Glu Val
    1025                1030                1035

Ser Gln  Glu Val Arg Val Cys  Pro Gly Arg Gly Tyr  Ile Leu Arg
    1040                1045                1050

Val Thr  Ala Tyr Lys Glu Gly  Tyr Gly Glu Gly Cys  Val Thr Ile
    1055                1060                1065

His Glu  Ile Glu Asn Asn Thr  Asp Glu Leu Lys Phe  Ser Asn Cys
    1070                1075                1080

Val Glu  Glu Glu Ile Tyr Pro  Asn Asn Thr Val Thr  Cys Asn Asp
    1085                1090                1095

Tyr Thr  Val Asn Gln Glu Glu  Tyr Gly Gly Ala Tyr  Thr Ser Arg
    1100                1105                1110

Asn Arg  Gly Tyr Asn Glu Ala  Pro Ser Val Pro Ala  Asp Tyr Ala
    1115                1120                1125

Ser Val  Tyr Glu Glu Lys Ser  Tyr Thr Asp Gly Arg  Arg Glu Asn
```

1130 1135 1140

Pro Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro Leu Pro
1145 1150 1155

Val Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp
1160 1165 1170

Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val
1175 1180 1185

Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1190 1195

<210> 82
<211> 3573
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC845.

<400> 82
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccaacggta 60

tcgaatcctt ccacgcaaat gaatctatca ccagatgctc gtattgaaga tagcttgtgt 120

gtagccgagg tgaacaatat tgatccattt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccattttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

```
acttcaatta atcctgtaac attacagttt acgtctcgag acgtttatag aacagaatca    1200
aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt    1260
aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat    1320
cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa    1380
cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac    1440
actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaacataatt    1500
gcatcggata gtattactca aatccctgca gtgaagggaa actttctttt taatggttct    1560
gtaatttcag gaccaggatt tactggtggg gacttagtta gattaaatag tagtggaaat    1620
aacattcaga atagagggta tattgaagtt ccaattcact tcccatcgac atctaccaga    1680
tatcgagttc gtgtacggta tgcttctgta accccgattc acctcaacgt taattggggt    1740
aattcatcca ttttttccaa tacagtacca gctacagcta cgtcattaga taatctacaa    1800
tcaagtgatt ttggttattt tgaaagtgcc aatgctttta catcttcatt aggtaatata    1860
gtaggtgtta gaaattttag tgggactgca ggagtgataa tagacagatt tgaatttatt    1920
ccagttactg caacactcga ggctgaatat aatctggaaa gagcacaaaa ggcggtgaat    1980
gagctgttta cttcttccaa tcaaatcggg ttaaaaacag atgtgacgga ttatcatatt    2040
gatcaagtat ccaatttagt tgagtgttta tctgatgaat tttgtctgga tgaaaaaaaa    2100
gaattgtccg agaaagtcaa acatgcgaag cgacttagtg atgagcggaa tttacttcaa    2160
gatccaaact ttagagggat caatagacaa ctagaccgtg gctggagagg aagtacggat    2220
attaccatcc aaggaggcga tgacgtattc aaagagaatt acgttacgct attgggtacc    2280
tttgatgagt gctatccaac gtatttatat caaaaaatag atgagtcgaa attaaaagcc    2340
tatacccgtt accaattaag agggtatatc gaagatagtc aagacttaga aatctattta    2400
attcgctaca atgccaaaca cgaaacagta aatgtgccag gtacgggttc cttatggccg    2460
ctttcagccc caagtccaat cggaaaatgt gcccatcatt cccatcattt ctccttggac    2520
attgatgttg gatgtacaga cttaaatgag gacttaggtg tatgggtgat attcaagatt    2580
aagacgcaag atggccatgc aagactagga aatctagaat ttctcgaaga gaaaccatta    2640
gtaggagaag cactagctcg tgtgaaaaga gcggagaaaa aatggagaga caaacgtgaa    2700
aaattggaat gggaaacaaa tattgtttat aaagaggcaa aagaatctgt agatgcttta    2760
tttgtaaact ctcaatatga tagattacaa gcggatacca acatcgcgat gattcatgcg    2820
gcagataaac gcgttcatag cattcgagaa gcttatctgc ctgagctgtc tgtgattccg    2880
ggtgtcaatg cggctatttt tgaagaatta gaagggcgta ttttcactgc attctcccta    2940
tatgatgcga gaaatgtcat taaaaatggt gattttaata atggcttatc ctgctggaac    3000
```

gtgaaagggc atgtagatgt agaagaacaa aacaaccacc gttcggtcct tgttgttccg 3060

gaatgggaag cagaagtgtc acaagaagtt cgtgtctgtc cgggtcgtgg ctatatcctt 3120

cgtgtcacag cgtacaagga gggatatgga gaaggttgcg taaccattca tgagatcgag 3180

aacaatacag acgaactgaa gtttagcaac tgtgtagaag aggaagtata tccaaacaac 3240

acggtaacgt gtaatgatta tactgcgact caagaagaat atgagggtac gtacacttct 3300

cgtaatcgag gatatgacgg agcctatgaa agcaattctt ctgtaccagc tgattatgca 3360

tcagcctatg aagaaaaagc atatacagat ggacgaagag acaatccttg tgaatctaac 3420

agaggatatg gggattacac accactacca gctggctatg tgacaaaaga attagagtac 3480

ttcccagaaa ccgataaggt atggattgag atcggagaaa cggaaggaac attcatcgtg 3540

gacagcgtgg aattacttct tatggaggaa tag 3573

<210> 83
<211> 1190
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC845.

<400> 83

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Thr Val Ser Asn Pro Ser Thr Gln Met Asn Leu Ser Pro Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Val Ala Glu Val Asn Asn Ile Asp
35 40 45

Pro Phe Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Leu Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn
130 135 140

Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145 150 155 160

```
Leu Glu Leu Asp Ile Thr Thr Ala Ile Pro Leu Phe Arg Ile Arg Asn
                165                 170                 175

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
        195                 200                 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

Ala Thr Thr Tyr Ser Gln Pro Tyr Gln Gly Val Gly Ile Gln Leu Phe
        435                 440                 445

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
    450                 455                 460

Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
```

```
                485                 490                 495

Thr Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Ala Val Lys
            500                 505                 510

Gly Asn Phe Leu Phe Asn Gly Ser Val Ile Ser Gly Pro Gly Phe Thr
        515                 520                 525

Gly Gly Asp Leu Val Arg Leu Asn Ser Ser Gly Asn Asn Ile Gln Asn
    530                 535                 540

Arg Gly Tyr Ile Glu Val Pro Ile His Phe Pro Ser Thr Ser Thr Arg
545                 550                 555                 560

Tyr Arg Val Arg Val Arg Tyr Ala Ser Val Thr Pro Ile His Leu Asn
                565                 570                 575

Val Asn Trp Gly Asn Ser Ser Ile Phe Ser Asn Thr Val Pro Ala Thr
            580                 585                 590

Ala Thr Ser Leu Asp Asn Leu Gln Ser Ser Asp Phe Gly Tyr Phe Glu
        595                 600                 605

Ser Ala Asn Ala Phe Thr Ser Ser Leu Gly Asn Ile Val Gly Val Arg
    610                 615                 620

Asn Phe Ser Gly Thr Ala Gly Val Ile Ile Asp Arg Phe Glu Phe Ile
625                 630                 635                 640

Pro Val Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln
                645                 650                 655

Lys Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys
            660                 665                 670

Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu
        675                 680                 685

Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu
    690                 695                 700

Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
705                 710                 715                 720

Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg
                725                 730                 735

Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu
            740                 745                 750

Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr
        755                 760                 765

Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr
    770                 775                 780

Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu
785                 790                 795                 800

Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly
                805                 810                 815
```

```
Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His
            820                 825                 830

His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu
        835                 840                 845

Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp
    850                 855                 860

Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu
865                 870                 875                 880

Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg
                885                 890                 895

Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu
            900                 905                 910

Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg
        915                 920                 925

Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg
    930                 935                 940

Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro
945                 950                 955                 960

Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr
                965                 970                 975

Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe
            980                 985                 990

Asn Asn Gly Leu Ser Cys Trp Asn  Val Lys Gly His Val  Asp Val Glu
        995                 1000                1005

Glu Gln  Asn Asn His Arg Ser  Val Leu Val Val Pro  Glu Trp Glu
    1010                1015                1020

Ala Glu  Val Ser Gln Glu Val  Arg Val Cys Pro Gly  Arg Gly Tyr
    1025                1030                1035

Ile Leu  Arg Val Thr Ala Tyr  Lys Glu Gly Tyr Gly  Glu Gly Cys
    1040                1045                1050

Val Thr  Ile His Glu Ile Glu  Asn Asn Thr Asp Glu  Leu Lys Phe
    1055                1060                1065

Ser Asn  Cys Val Glu Glu Glu  Val Tyr Pro Asn Asn  Thr Val Thr
    1070                1075                1080

Cys Asn  Asp Tyr Thr Ala Thr  Gln Glu Glu Tyr Glu  Gly Thr Tyr
    1085                1090                1095

Thr Ser  Arg Asn Arg Gly Tyr  Asp Gly Ala Tyr Glu  Ser Asn Ser
    1100                1105                1110

Ser Val  Pro Ala Asp Tyr Ala  Ser Ala Tyr Glu Glu  Lys Ala Tyr
    1115                1120                1125
```

Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr
1130 1135 1140

Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu
1145 1150 1155

Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu
1160 1165 1170

Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met
1175 1180 1185

Glu Glu
1190

<210> 84
<211> 3567
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC846.

<400> 84
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccaacggta 60

tcgaatcctt ccacgcaaat gaatctatca ccagatgctc gtattgaaga tagcttgtgt 120

gtagccgagg tgaacaatat tgatccattt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccatttttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

```
acttcaatta atcctgtaac attacagttt acgtctcgag acgtttatag aacagaatca    1200
aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt    1260
aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat    1320
cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa    1380
cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac    1440
actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaatacgatt    1500
ggaccaaata gaattacaca aataccattg gtaaaagcac tgaatcttca ttcaggtgtt    1560
actgttgttg gagggccagg atttacaggt ggagatatcc ttcgaagaac gaatactggc    1620
acatttgcag atatgagagt aaatattact gggccattat cccaaagata tcgtgtaaga    1680
attcgctatg cttctacgac agatttacaa tttttcacga gaatcaatgg aacttctgta    1740
aatcaaggta atttccaaag aactatgaat agaggggata atttagaatc tggaaacttt    1800
aggactgcag gatttagtac gccttttagt ttttcaaatg cgcaaagtac attcacattg    1860
ggtactcagg ctttttcaaa tcaggaagtt tatatagatc gaattgaatt tgtcccggca    1920
gaagtaacat tcgaggcaga atctgattta gaaagagcac aaaaggcggt gaatgagctg    1980
tttacttctt ccaatcaaat cgggttaaaa acagatgtga cggattatca tattgatcaa    2040
gtatccaatt tagttgagtg tttatctgat gaattttgtc tggatgaaaa aaaagaattg    2100
tccgagaaag tcaaacatgc gaagcgactt agtgatgagc ggaatttact tcaagatcca    2160
aactttagag ggatcaatag acaactagac cgtggctgga gaggaagtac ggatattacc    2220
atccaaggag gcgatgacgt attcaaagag aattacgtta cgctattggg tacctttgat    2280
gagtgctatc caacgtattt atatcaaaaa atagatgagt cgaaattaaa agcctatacc    2340
cgttaccaat taagagggta tatcgaagat agtcaagact tagaaatcta tttaattcgc    2400
tacaatgcca aacacgaaac agtaaatgtg ccaggtacgg gttccttatg gccgctttca    2460
gccccaagtc caatcggaaa atgtgcccat cattcccatc atttctcctt ggacattgat    2520
gttggatgta cagacttaaa tgaggactta ggtgtatggg tgatattcaa gattaagacg    2580
caagatggcc atgcaagact aggaaatcta gaatttctcg aagagaaacc attagtagga    2640
gaagcactag ctcgtgtgaa aagagcggag aaaaaatgga gagacaaacg tgaaaaattg    2700
gaatgggaaa caaatattgt ttataaagag gcaaaagaat ctgtagatgc tttatttgta    2760
aactctcaat atgatagatt acaagcggat accaacatcg cgatgattca tgcggcagat    2820
aaacgcgttc atagcattcg agaagcttat ctgcctgagc tgtctgtgat tccgggtgtc    2880
aatgcggcta tttttgaaga attagaaggg cgtattttca ctgcattctc cctatatgat    2940
```

gcgagaaatg tcattaaaaa tggtgatttt aataatggct tatcctgctg gaacgtgaaa 3000

gggcatgtag atgtagaaga acaaaacaac caccgttcgg tccttgttgt tccggaatgg 3060

gaagcagaag tgtcacaaga agttcgtgtc tgtccgggtc gtggctatat ccttcgtgtc 3120

acagcgtaca aggagggata tggagaaggt tgcgtaacca ttcatgagat cgagaacaat 3180

acagacgaac tgaagtttag caactgtgta gaagaggaag tatatccaaa caacacggta 3240

acgtgtaatg attatactgc gactcaagaa gaatatgagg gtacgtacac ttctcgtaat 3300

cgaggatatg acggagccta tgaaagcaat tcttctgtac cagctgatta tgcatcagcc 3360

tatgaagaaa aagcatatac agatggacga agagacaatc cttgtgaatc taacagagga 3420

tatggggatt acacaccact accagctggc tatgtgacaa aagaattaga gtacttccca 3480

gaaaccgata aggtatggat tgagatcgga gaaacggaag gaacattcat cgtggacagc 3540

gtggaattac ttcttatgga ggaatag 3567

<210> 85
<211> 1188
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC846.

<220>
<221> ĐẶC TÍNH_KHÁC
<222> (1)..(1188)
<223> Trình tự axit amin của protein khảm TIC846.

<400> 85

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Thr Val Ser Asn Pro Ser Thr Gln Met Asn Leu Ser Pro Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Val Ala Glu Val Asn Asn Ile Asp
35 40 45

Pro Phe Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Leu Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

```
Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn
    130                 135                 140

Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145                 150                 155                 160

Leu Glu Leu Asp Ile Thr Thr Ala Ile Pro Leu Phe Arg Ile Arg Asn
                165                 170                 175

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
        195                 200                 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

Ala Thr Thr Tyr Ser Gln Pro Tyr Gln Gly Val Gly Ile Gln Leu Phe
        435                 440                 445

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
```

```
    450                 455                 460

Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
                485                 490                 495

Thr Asn Thr Ile Gly Pro Asn Arg Ile Thr Gln Ile Pro Leu Val Lys
            500                 505                 510

Ala Leu Asn Leu His Ser Gly Val Thr Val Val Gly Gly Pro Gly Phe
        515                 520                 525

Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Thr Phe Ala Asp
    530                 535                 540

Met Arg Val Asn Ile Thr Gly Pro Leu Ser Gln Arg Tyr Arg Val Arg
545                 550                 555                 560

Ile Arg Tyr Ala Ser Thr Thr Asp Leu Gln Phe Phe Thr Arg Ile Asn
                565                 570                 575

Gly Thr Ser Val Asn Gln Gly Asn Phe Gln Arg Thr Met Asn Arg Gly
            580                 585                 590

Asp Asn Leu Glu Ser Gly Asn Phe Arg Thr Ala Gly Phe Ser Thr Pro
        595                 600                 605

Phe Ser Phe Ser Asn Ala Gln Ser Thr Phe Thr Leu Gly Thr Gln Ala
    610                 615                 620

Phe Ser Asn Gln Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala
625                 630                 635                 640

Glu Val Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala
                645                 650                 655

Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp
            660                 665                 670

Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu
        675                 680                 685

Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val
    690                 695                 700

Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro
705                 710                 715                 720

Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser
                725                 730                 735

Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr
            740                 745                 750

Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr
        755                 760                 765

Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu
    770                 775                 780
```

```
Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg
785                 790                 795                 800

Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu
                805                 810                 815

Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser
            820                 825                 830

His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu
        835                 840                 845

Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His
    850                 855                 860

Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly
865                 870                 875                 880

Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys
                885                 890                 895

Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys
            900                 905                 910

Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln
        915                 920                 925

Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His
    930                 935                 940

Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val
945                 950                 955                 960

Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe
                965                 970                 975

Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn
            980                 985                 990

Gly Leu Ser Cys Trp Asn Val Lys  Gly His Val Asp Val  Glu Glu Gln
        995                 1000                1005

Asn Asn  His Arg Ser Val Leu  Val Val Pro Glu Trp  Glu Ala Glu
    1010                1015                1020

Val Ser  Gln Glu Val Arg Val  Cys Pro Gly Arg Gly  Tyr Ile Leu
    1025                1030                1035

Arg Val  Thr Ala Tyr Lys Glu  Gly Tyr Gly Glu Gly  Cys Val Thr
    1040                1045                1050

Ile His  Glu Ile Glu Asn Asn  Thr Asp Glu Leu Lys  Phe Ser Asn
    1055                1060                1065

Cys Val  Glu Glu Glu Val Tyr  Pro Asn Asn Thr Val  Thr Cys Asn
    1070                1075                1080

Asp Tyr  Thr Ala Thr Gln Glu  Glu Tyr Glu Gly Thr  Tyr Thr Ser
    1085                1090                1095
```

Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val
1100 1105 1110

Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
1115 1120 1125

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp
1130 1135 1140

Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr
1145 1150 1155

Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu
1160 1165 1170

Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180 1185

<210> 86
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC858.

<400> 86
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccaacggta 60

tcgaatcctt ccacgcaaat gaatctatca ccagatgctc gtattgaaga tagcttgtgt 120

gtagccgagg tgaacaatat tgatccattt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccatttttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

acttcaatta atcctgtaac attacagttt acgtctcgag acgtttatag aacagaatca 1200

aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt 1260

aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat 1320

cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa 1380

cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac 1440

actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaatacgatt 1500

ggaccaaata gaattaccca aatcccaatg gtaaaagcat ccgaacttcc tcaaggtacc 1560

actgttgtta gaggaccagg atttactggt ggggatattc ttcgaagaac gaatactggt 1620

ggatttggac cgataagagt aactgttaac ggaccattaa cacaaagata tcgtatagga 1680

ttccgctatg cttcaactgt agattttgat ttctttgtat cacgtggagg tactactgta 1740

aataatttta gattcctacg tacaatgaac agtggagacg aactaaaata cggaaatttt 1800

gtgagacgtg cttttactac acctttttact tttacacaaa ttcaagatat aattcgaacg 1860

tctattcaag gccttagtgg aaatggggaa gtgtatatag ataaaattga aattattcca 1920

gttactgcaa ccttcgaagc agaatatgat ttagaaagag cacaaaaggc ggtgaatgag 1980

ctgtttactt cttccaatca aatcgggtta aaaacagatg tgacggatta tcatattgat 2040

caagtatcca atttagttga gtgtttatct gatgaatttt gtctggatga aaaaaaagaa 2100

ttgtccgaga aagtcaaaca tgcgaagcga cttagtgatg agcggaattt acttcaagat 2160

ccaaacttta gagggatcaa tagacaacta gaccgtggct ggagaggaag tacggatatt 2220

accatccaag gaggcgatga cgtattcaaa gagaattacg ttacgctatt gggtaccttt 2280

gatgagtgct atccaacgta tttatatcaa aaaatagatg agtcgaaatt aaaagcctat 2340

acccgttacc aattaagagg gtatatcgaa gatagtcaag acttagaaat ctatttaatt 2400

cgctacaatg ccaaacacga aacagtaaat gtgccaggta cgggttcctt atggccgctt 2460

tcagccccaa gtccaatcgg aaaatgtgcc catcattccc atcatttctc cttggacatt 2520

gatgttggat gtacagactt aaatgaggac ttaggtgtat gggtgatatt caagattaag 2580

acgcaagatg gccatgcaag actaggaaat ctagaatttc tcgaagagaa accattagta 2640

ggagaagcac tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgtgaaaaa 2700

ttggaatggg aaacaaatat tgtttataaa gaggcaaaag aatctgtaga tgctttattt 2760

gtaaactctc aatatgatag attacaagcg gataccaaca tcgcgatgat tcatgcggca 2820

gataaacgcg ttcatagcat tcgagaagct tatctgcctg agctgtctgt gattccgggt 2880

gtcaatgcgg ctatttttga agaattagaa gggcgtattt tcactgcatt ctccctatat 2940

gatgcgagaa atgtcattaa aaatggtgat tttaataatg gcttatcctg ctggaacgtg 3000

aaagggcatg tagatgtaga agaacaaaac aaccaccgtt cggtccttgt tgttccggaa 3060

tgggaagcag aagtgtcaca agaagttcgt gtctgtccgg gtcgtggcta tatccttcgt 3120

gtcacagcgt acaaggaggg atatggagaa ggttgcgtaa ccattcatga gatcgagaac 3180

aatacagacg aactgaagtt tagcaactgt gtagaagagg aagtatatcc aaacaacacg 3240

gtaacgtgta atgattatac tgcgactcaa gaagaatatg agggtacgta cacttctcgt 3300

aatcgaggat atgacggagc ctatgaaagc aattcttctg taccagctga ttatgcatca 3360

gcctatgaag aaaaagcata tacagatgga cgaagagaca atccttgtga atctaacaga 3420

ggatatgggg attacacacc actaccagct ggctatgtga caaaagaatt agagtacttc 3480

ccagaaaccg ataaggtatg gattgagatc ggagaaacgg aaggaacatt catcgtggac 3540

agcgtggaat tacttcttat ggaggaatag 3570

<210> 87
<211> 1189
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC858.

<400> 87

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Thr Val Ser Asn Pro Ser Thr Gln Met Asn Leu Ser Pro Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Val Ala Glu Val Asn Asn Ile Asp
35 40 45

Pro Phe Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Leu Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn
130 135 140

```
Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145                 150                 155                 160

Leu Glu Leu Asp Ile Thr Thr Ala Ile Pro Leu Phe Arg Ile Arg Asn
                165                 170                 175

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
        195                 200                 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

Ala Thr Thr Tyr Ser Gln Pro Tyr Gln Gly Val Gly Ile Gln Leu Phe
        435                 440                 445

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
    450                 455                 460
```

```
Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
                485                 490                 495

Thr Asn Thr Ile Gly Pro Asn Arg Ile Thr Gln Ile Pro Met Val Lys
            500                 505                 510

Ala Ser Glu Leu Pro Gln Gly Thr Thr Val Val Arg Gly Pro Gly Phe
        515                 520                 525

Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Gly Phe Gly Pro
    530                 535                 540

Ile Arg Val Thr Val Asn Gly Pro Leu Thr Gln Arg Tyr Arg Ile Gly
545                 550                 555                 560

Phe Arg Tyr Ala Ser Thr Val Asp Phe Asp Phe Phe Val Ser Arg Gly
                565                 570                 575

Gly Thr Thr Val Asn Asn Phe Arg Phe Leu Arg Thr Met Asn Ser Gly
            580                 585                 590

Asp Glu Leu Lys Tyr Gly Asn Phe Val Arg Arg Ala Phe Thr Thr Pro
        595                 600                 605

Phe Thr Phe Thr Gln Ile Gln Asp Ile Ile Arg Thr Ser Ile Gln Gly
    610                 615                 620

Leu Ser Gly Asn Gly Glu Val Tyr Ile Asp Lys Ile Glu Ile Ile Pro
625                 630                 635                 640

Val Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys
                645                 650                 655

Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr
            660                 665                 670

Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys
        675                 680                 685

Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys
    690                 695                 700

Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp
705                 710                 715                 720

Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly
                725                 730                 735

Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn
            740                 745                 750

Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu
        755                 760                 765

Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln
    770                 775                 780

Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile
```

```
785                 790                 795                 800

Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser
                805                 810                 815

Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His
            820                 825                 830

Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn
        835                 840                 845

Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly
    850                 855                 860

His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val
865                 870                 875                 880

Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp
                885                 890                 895

Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala
            900                 905                 910

Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu
        915                 920                 925

Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val
    930                 935                 940

His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly
945                 950                 955                 960

Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala
                965                 970                 975

Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn
            980                 985                 990

Asn Gly Leu Ser Cys Trp Asn Val  Lys Gly His Val Asp  Val Glu Glu
        995                 1000                1005

Gln Asn  Asn His Arg Ser Val  Leu Val Val Pro Glu  Trp Glu Ala
    1010                1015                1020

Glu Val  Ser Gln Glu Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile
    1025                1030                1035

Leu Arg  Val Thr Ala Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val
    1040                1045                1050

Thr Ile  His Glu Ile Glu Asn  Asn Thr Asp Glu Leu  Lys Phe Ser
    1055                1060                1065

Asn Cys  Val Glu Glu Glu Val  Tyr Pro Asn Asn Thr  Val Thr Cys
    1070                1075                1080

Asn Asp  Tyr Thr Ala Thr Gln  Glu Glu Tyr Glu Gly  Thr Tyr Thr
    1085                1090                1095

Ser Arg  Asn Arg Gly Tyr Asp  Gly Ala Tyr Glu Ser  Asn Ser Ser
    1100                1105                1110
```

Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr
1115 1120 1125

Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly
1130 1135 1140

Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu
1145 1150 1155

Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr
1160 1165 1170

Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu
1175 1180 1185

Glu

<210> 88
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC865.

<400> 88
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ccttatcgat tccagctgta 60

tcgaatcatt ccgcacaaat ggatctatcg ctagatgctc gtattgagga ttctttgtgt 120

atagccgagg ggaataatat caatccactt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccattttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

acttcaatta atcctgtaac attacagttt acgtctcgtg acgtttatag aacagaatca 1200

aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt 1260

aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat 1320

cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa 1380

cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac 1440

actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaacacaatt 1500

gatccggaga ggattactca aataccattg gtaaaagcac atacacttca gtcaggtact 1560

actgttgtaa gagggcccgg gtttacggga ggagatattc ttcgacgaac aagtggagga 1620

ccatttgctt atactattgt taatataaat gggcaattac cccaaaggta tcgtgcaaga 1680

atacgctatg cctctactac aaatctaaga atttacgtaa cggttgcagg tgaacggatt 1740

tttgctggtc aatttaacaa aacaatggat accggtgacc cattaacatt ccaatctttt 1800

agttacgcaa ctattaatac agcttttaca ttcccaatga gccagagtag tttcacagta 1860

ggtgctgata cttttagttc agggaatgaa gtttatatag acagatttga attgattcca 1920

gttactgcaa catttgaagc agaatatgat ttagaaagag cacaaaaggc ggtgaatgag 1980

ctgtttactt cttccaatca aatcgggtta aaaacagatg tgacggatta tcatattgat 2040

caagtatcca atttagttga gtgtttatct gatgaatttt gtctggatga aaaaaaagaa 2100

ttgtccgaga aagtcaaaca tgcgaagcga cttagtgatg agcggaattt acttcaagat 2160

ccaaacttta gagggatcaa tagacaacta gaccgtggct ggagaggaag tacggatatt 2220

accatccaag gaggcgatga cgtattcaaa gagaattacg ttacgctatt gggtaccttt 2280

gatgagtgct atccaacgta tttatatcaa aaaatagatg agtcgaaatt aaaagcctat 2340

acccgttacc aattaagagg gtatatcgaa gatagtcaag acttagaaat ctatttaatt 2400

cgctacaatg ccaaacacga aacagtaaat gtgccaggta cgggttcctt atggccgctt 2460

tcagccccaa gtccaatcgg aaaatgtgcc catcattccc atcatttctc cttggacatt 2520

gatgttggat gtacagactt aaatgaggac ttaggtgtat gggtgatatt caagattaag 2580

acgcaagatg gccatgcaag actaggaaat ctagaatttc tcgaagagaa accattagta 2640

ggagaagcac tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgtgaaaaa 2700

ttggaatggg aaacaaatat tgtttataaa gaggcaaaag aatctgtaga tgctttattt 2760

gtaaactctc aatatgatag attacaagcg gataccaaca tcgcgatgat tcatgcggca 2820

gataaacgcg ttcatagcat tcgagaagct tatctgcctg agctgtctgt gattccgggt 2880

gtcaatgcgg ctatttttga agaattagaa gggcgtattt tcactgcatt ctccctatat 2940

gatgcgagaa atgtcattaa aaatggtgat tttaataatg gcttatcctg ctggaacgtg 3000

aaagggcatg tagatgtaga agaacaaaac aaccaccgtt cggtccttgt tgttccggaa 3060

tgggaagcag aagtgtcaca agaagttcgt gtctgtccgg gtcgtggcta tatccttcgt 3120

gtcacagcgt acaaggaggg atatggagaa ggttgcgtaa ccattcatga gatcgagaac 3180

aatacagacg aactgaagtt tagcaactgt gtagaagagg aagtatatcc aaacaacacg 3240

gtaacgtgta atgattatac tgcgactcaa gaagaatatg agggtacgta cacttctcgt 3300

aatcgaggat atgacggagc ctatgaaagc aattcttctg taccagctga ttatgcatca 3360

gcctatgaag aaaaagcata tacagatgga cgaagagaca atccttgtga atctaacaga 3420

ggatatgggg attacacacc actaccagct ggctatgtga caaaagaatt agagtacttc 3480

ccagaaaccg ataaggtatg gattgagatc ggagaaacgg aaggaacatt catcgtggac 3540

agcgtggaat tacttcttat ggaggaatag 3570

<210> 89
<211> 1189
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC865.

<400> 89

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asp Leu Ser Leu Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asn
35 40 45

Pro Leu Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Leu Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn

```
    130                 135                 140

Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145                 150                 155                 160

Leu Glu Leu Asp Ile Thr Thr Ala Ile Pro Leu Phe Arg Ile Arg Asn
                165                 170                 175

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
        195                 200                 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

Ala Thr Thr Tyr Ser Gln Pro Tyr Gln Gly Val Gly Ile Gln Leu Phe
        435                 440                 445

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
    450                 455                 460
```

```
Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
                485                 490                 495

Thr Asn Thr Ile Asp Pro Glu Arg Ile Thr Gln Ile Pro Leu Val Lys
            500                 505                 510

Ala His Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe
        515                 520                 525

Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr
    530                 535                 540

Thr Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg
545                 550                 555                 560

Ile Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala
                565                 570                 575

Gly Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly
            580                 585                 590

Asp Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala
        595                 600                 605

Phe Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr
    610                 615                 620

Phe Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro
625                 630                 635                 640

Val Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys
                645                 650                 655

Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr
            660                 665                 670

Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys
        675                 680                 685

Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys
    690                 695                 700

Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp
705                 710                 715                 720

Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly
                725                 730                 735

Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn
            740                 745                 750

Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu
        755                 760                 765

Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln
    770                 775                 780
```

```
Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile
785                 790                 795                 800

Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser
                805                 810                 815

Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His
            820                 825                 830

Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn
        835                 840                 845

Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly
    850                 855                 860

His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val
865                 870                 875                 880

Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp
                885                 890                 895

Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala
            900                 905                 910

Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu
        915                 920                 925

Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val
    930                 935                 940

His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly
945                 950                 955                 960

Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala
                965                 970                 975

Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn
            980                 985                 990

Asn Gly Leu Ser Cys Trp Asn Val  Lys Gly His Val Asp  Val Glu Glu
        995                 1000                1005

Gln Asn  Asn His Arg Ser Val  Leu Val Val Pro Glu  Trp Glu Ala
    1010                1015                1020

Glu Val  Ser Gln Glu Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile
    1025                1030                1035

Leu Arg  Val Thr Ala Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val
    1040                1045                1050

Thr Ile  His Glu Ile Glu Asn  Asn Thr Asp Glu Leu  Lys Phe Ser
    1055                1060                1065

Asn Cys  Val Glu Glu Glu Val  Tyr Pro Asn Asn Thr  Val Thr Cys
    1070                1075                1080

Asn Asp  Tyr Thr Ala Thr Gln  Glu Glu Tyr Glu Gly  Thr Tyr Thr
    1085                1090                1095

Ser Arg  Asn Arg Gly Tyr Asp  Gly Ala Tyr Glu Ser  Asn Ser Ser
```

```
    1100                1105                1110

Val Pro  Ala Asp Tyr Ala Ser  Ala Tyr Glu Glu Lys  Ala Tyr Thr
    1115                1120                1125

Asp Gly  Arg Arg Asp Asn Pro  Cys Glu Ser Asn Arg  Gly Tyr Gly
    1130                1135                1140

Asp Tyr  Thr Pro Leu Pro Ala  Gly Tyr Val Thr Lys  Glu Leu Glu
    1145                1150                1155

Tyr Phe  Pro Glu Thr Asp Lys  Val Trp Ile Glu Ile  Gly Glu Thr
    1160                1165                1170

Glu Gly  Thr Phe Ile Val Asp  Ser Val Glu Leu Leu  Leu Met Glu
    1175                1180                1185

Glu
```

<210> 90
<211> 3564
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC866.

<400> 90

```
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccagctgta       60

tcgaatcatt ccgcacaaat gaatctatca accgatgctc gtattgagga tagcttgtgt      120

atagccgagg ggaacaatat cgatccattt gttagcgcat caacagtcca aacgggtatt      180

aacatagctg gtagaatact aggtgtatta ggcgtaccgt ttgctggaca aatagctagt      240

ttttatagtt ttcttgttgg tgaattatgg ccccgcggca gagatccttg ggaaattttc      300

ctagaacatg tcgaacaact tataagacaa caagtaacag aaaatactag ggatacggct      360

cttgctcgat tacaaggttt aggaaattcc tttagagcct atcaacagtc acttgaagat      420

tggctagaaa accgtgatga tgcaagaacg agaagtgttc tttataccca atatatagcc      480

ttagaacttg attttcttaa tgcgatgccg cttttcgcaa ttagaaacca agaagttcca      540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agatgcctct      600

ctttttggta gtgaatttgg gcttacatcc caagaaattc aacgttatta tgagcgccaa      660

gtggaaaaaa cgagagaata ttctgattat tgcgcaagat ggtataatac gggtttaaat      720

aatttgagag ggacaaatgc tgaaagttgg ttgcgatata atcaattccg tagagactta      780

acgctaggag tattagatct agtggcacta ttcccaagct atgacacgcg tgtttatcca      840

atgaatacca gtgctcaatt aacaagagaa atttatacag atccaattgg gagaacaaat      900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc      960
```

atagaggctg ccgttattag gcctccgcat ctacttgatt ttccagaaca gcttacaatt 1020

ttcagcgtat taagtcgatg gagtaatact caatatatga attactgggt gggacataga 1080

cttgaatcgc gaacaataag ggggtcatta agtacctcga cacacggaaa taccaatact 1140

tctattaatc ctgtaacatt acagttcaca tctcgagacg tttatagaac agaatcattt 1200

gcagggataa atatacttct aactactcct gtgaatggag taccttgggc tagatttaat 1260

tggagaaatc ccctgaattc tcttagaggt agccttctct atactatagg gtatactgga 1320

gtggggacac aactatttga ttcagaaact gaattaccac cagaaacaac agaacgacca 1380

aattatgaat cttacagtca tagattatct aatataagac taatatcagg aaacactttg 1440

agagcaccag tatattcttg gacgcaccgt agtgcagatc gtacaaatac cattagttca 1500

gatagcatta ctcaaatacc attggtaaaa gcacatacac ttcagtcagg tactactgtt 1560

gtaagagggc ccgggtttac gggaggagat attcttcgac gaacaagtgg aggaccattt 1620

gcttatacta ttgttaatat aaatgggcaa ttaccccaaa ggtatcgtgc aagaatacgc 1680

tatgcctcta ctacaaatct aagaatttac gtaacggttg caggtgaacg gatttttgct 1740

ggtcaattta acaaaacaat ggataccggt gacccattaa cattccaatc ttttagttac 1800

gcaactatta atacagcttt tacattccca atgagccaga gtagtttcac agtaggtgct 1860

gatactttta gttcagggaa tgaagtttat atagacagat ttgaattgat tccagttact 1920

gcaacatttg aagcagaata tgatttagaa agagcacaaa aggcggtgaa tgagctgttt 1980

acttcttcca atcaaatcgg gttaaaaaca gatgtgacgg attatcatat tgatcaagta 2040

tccaatttag ttgagtgttt atctgatgaa ttttgtctgg atgaaaaaaa agaattgtcc 2100

gagaaagtca aacatgcgaa gcgacttagt gatgagcgga atttacttca agatccaaac 2160

tttagaggga tcaatagaca actagaccgt ggctggagag gaagtacgga tattaccatc 2220

caaggaggcg atgacgtatt caaagagaat tacgttacgc tattgggtac ctttgatgag 2280

tgctatccaa cgtatttata tcaaaaaata gatgagtcga aattaaaagc ctatacccgt 2340

taccaattaa gagggtatat cgaagatagt caagacttag aaatctattt aattcgctac 2400

aatgccaaac acgaaacagt aaatgtgcca ggtacgggtt ccttatggcc gctttcagcc 2460

ccaagtccaa tcggaaaatg tgcccatcat tcccatcatt tctccttgga cattgatgtt 2520

ggatgtacag acttaaatga ggacttaggt gtatgggtga tattcaagat taagacgcaa 2580

gatggccatg caagactagg aaatctagaa tttctcgaag agaaaccatt agtaggagaa 2640

gcactagctc gtgtgaaaag agcggagaaa aaatggagag acaaacgtga aaaattggaa 2700

tgggaaacaa atattgttta taaagaggca aaagaatctg tagatgcttt atttgtaaac 2760

tctcaatatg atagattaca agcggatacc aacatcgcga tgattcatgc ggcagataaa 2820

cgcgttcata gcattcgaga agcttatctg cctgagctgt ctgtgattcc gggtgtcaat 2880

gcggctattt ttgaagaatt agaagggcgt attttcactg cattctccct atatgatgcg 2940

agaaatgtca ttaaaaatgg tgattttaat aatggcttat cctgctggaa cgtgaaaggg 3000

catgtagatg tagaagaaca aaacaaccac cgttcggtcc ttgttgttcc ggaatgggaa 3060

gcagaagtgt cacaagaagt tcgtgtctgt ccgggtcgtg gctatatcct tcgtgtcaca 3120

gcgtacaagg agggatatgg agaaggttgc gtaaccattc atgagatcga gaacaataca 3180

gacgaactga agtttagcaa ctgtgtagaa gaggaagtat atccaaacaa cacggtaacg 3240

tgtaatgatt atactgcgac tcaagaagaa tatgagggta cgtacacttc tcgtaatcga 3300

ggatatgacg gagcctatga aagcaattct tctgtaccag ctgattatgc atcagcctat 3360

gaagaaaaag catatacaga tggacgaaga gacaatcctt gtgaatctaa cagaggatat 3420

ggggattaca caccactacc agctggctat gtgacaaaag aattagagta cttcccagaa 3480

accgataagg tatggattga gatcggagaa acggaaggaa cattcatcgt ggacagcgtg 3540

gaattacttc ttatggagga atag 3564

<210> 91
<211> 1187
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC866.

<400> 91

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Ala Val Ser Asn His Ser Ala Gln Met Asn Leu Ser Thr Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Ile Ala Glu Gly Asn Asn Ile Asp
35 40 45

Pro Phe Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Ile Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Arg Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asp Thr Ala Leu Ala Arg Leu Gln Gly Leu Gly
115 120 125

```
Asn Ser Phe Arg Ala Tyr Gln Gln Ser Leu Glu Asp Trp Leu Glu Asn
    130                 135                 140

Arg Asp Asp Ala Arg Thr Arg Ser Val Leu Tyr Thr Gln Tyr Ile Ala
145                 150                 155                 160

Leu Glu Leu Asp Phe Leu Asn Ala Met Pro Leu Phe Ala Ile Arg Asn
                165                 170                 175

Gln Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Phe Gly Leu
        195                 200                 205

Thr Ser Gln Glu Ile Gln Arg Tyr Tyr Glu Arg Gln Val Glu Lys Thr
    210                 215                 220

Arg Glu Tyr Ser Asp Tyr Cys Ala Arg Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Val Tyr Pro Met Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Val Ile Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Phe Ser Val Leu Ser Arg Trp Ser Asn Thr Gln Tyr
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Glu Ser Arg Thr Ile Arg Gly
        355                 360                 365

Ser Leu Ser Thr Ser Thr His Gly Asn Thr Asn Thr Ser Ile Asn Pro
    370                 375                 380

Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser Phe
385                 390                 395                 400

Ala Gly Ile Asn Ile Leu Leu Thr Thr Pro Val Asn Gly Val Pro Trp
                405                 410                 415

Ala Arg Phe Asn Trp Arg Asn Pro Leu Asn Ser Leu Arg Gly Ser Leu
            420                 425                 430

Leu Tyr Thr Ile Gly Tyr Thr Gly Val Gly Thr Gln Leu Phe Asp Ser
        435                 440                 445

Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr Glu Ser
```

```
    450                 455                 460

Tyr Ser His Arg Leu Ser Asn Ile Arg Leu Ile Ser Gly Asn Thr Leu
465                 470                 475                 480

Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg Thr Asn
                485                 490                 495

Thr Ile Ser Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala His
            500                 505                 510

Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
        515                 520                 525

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
    530                 535                 540

Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg
545                 550                 555                 560

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
                565                 570                 575

Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
            580                 585                 590

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
        595                 600                 605

Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
    610                 615                 620

Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr
625                 630                 635                 640

Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val
                645                 650                 655

Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
            660                 665                 670

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
        675                 680                 685

Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys
    690                 695                 700

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
705                 710                 715                 720

Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr
                725                 730                 735

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
            740                 745                 750

Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
        755                 760                 765

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
    770                 775                 780
```

```
Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
785                 790                 795                 800

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
                805                 810                 815

Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His
            820                 825                 830

His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp
        835                 840                 845

Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala
    850                 855                 860

Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu
865                 870                 875                 880

Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg
                885                 890                 895

Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu
            900                 905                 910

Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala
        915                 920                 925

Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser
    930                 935                 940

Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn
945                 950                 955                 960

Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser
                965                 970                 975

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly
            980                 985                 990

Leu Ser Cys Trp Asn Val Lys Gly  His Val Asp Val Glu  Glu Gln Asn
        995                 1000                1005

Asn His  Arg Ser Val Leu Val  Val Pro Glu Trp Glu  Ala Glu Val
    1010                1015                1020

Ser Gln  Glu Val Arg Val Cys  Pro Gly Arg Gly Tyr  Ile Leu Arg
    1025                1030                1035

Val Thr  Ala Tyr Lys Glu Gly  Tyr Gly Glu Gly Cys  Val Thr Ile
    1040                1045                1050

His Glu  Ile Glu Asn Asn Thr  Asp Glu Leu Lys Phe  Ser Asn Cys
    1055                1060                1065

Val Glu  Glu Glu Val Tyr Pro  Asn Asn Thr Val Thr  Cys Asn Asp
    1070                1075                1080

Tyr Thr  Ala Thr Gln Glu Glu  Tyr Glu Gly Thr Tyr  Thr Ser Arg
    1085                1090                1095
```

Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro
1100 1105 1110

Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly
1115 1120 1125

Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr
1130 1135 1140

Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe
1145 1150 1155

Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly
1160 1165 1170

Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180 1185

<210> 92
<211> 3525
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC838.

<400> 92
atggaggaaa ataatcaaaa tcaatgcata ccttacaatt gtttaagtaa tcctgaagaa 60

gtacttttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca 120

cttgttcagt ttctggtatc taactttgta ccaggggggag gatttttagt tggattaata 180

gattttgtat ggggaatagt tggcccttct caatgggatg catttctagt acaaattgaa 240

caattaatta atgaaagaat agctgaattt gctaggaatg ctgctattgc taatttagaa 300

ggattaggaa acaatttcaa tatatatgtg gaagcattta aagaatggga agaagatcct 360

aataatccag caaccaggac cagagtaatt gatcgctttc gtatacttga tgggctactt 420

gaaagggaca ttccttcgtt tgcaatttct ggatttgaag taccccttt atccgtttat 480

gctcaagcgg ccaatctgca tctagctata ttaagagatt ctgtaatttt tggagaaaga 540

tggggattga caacgataaa tgtcaatgaa aactataata gactaattag gcatattgat 600

gaatatgctg atcactgtgc aaatacgtat aatcggggat taaataattt accgaaatct 660

acgtatcaag attggataac atataatcga ttacggagag acttaacatt gactgtatta 720

gatatcgccg ctttctttcc aaactatgac aataggagat atccaattca gtcagttggt 780

caactaacaa gggaaattta tacggaccca ttaattactt ttaatccaca gttacagtct 840

gtagctcaat tacctacttt taacgttatg gaaagcaacg caattagaac tcctcattta 900

tttgatgtat tgaataatct tacaattttt acagattggt ttagtgttgg acgcaacttt 960

tattggggag gacatcgagt aatatctaac cgtataggag gaggtaacat aacatctcct 1020

atatatggaa gggaggcgaa tcaggagcct ccaagatctt ttacttttaa tgggcctgtt 1080

tttaggactt tatcaaatcc tacttttaga cctttacagc aaccttggcc agcgccacca 1140

tttaatttac gtggtgttga aggagtagaa ttttctacac ctttaaatag ctttacgtat 1200

cgaggaaggg gtacggttga ttctttaact gagttaccgc ctgaggataa tagtgtgcct 1260

cctcgcgaag gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaacc 1320

ccatttttaa caactggtcc agtattttct tggacgcatc gtagtgctac tgatcgaaat 1380

ataattgcat cggatagtat tactcaaatc cctgcagtga agggaaactt tctttttaat 1440

ggttctgtaa tttcaggacc aggatttact ggtggggact tagttagatt aaatagtagt 1500

ggaaataaca ttcagaatag agggtatatt gaagttccaa ttcacttccc atcgacatct 1560

accagatatc gagttcgtgt acggtatgct tctgtaaccc cgattcacct caacgttaat 1620

tggggtaatt catccatttt ttccaataca gtaccagcta cagctacgtc attagataat 1680

ctacaatcaa gtgattttgg ttattttgaa agtgccaatg cttttacatc ttcattaggt 1740

aatatagtag gtgttagaaa ttttagtggg actgcaggag tgataataga cagatttgaa 1800

tttattcctg ttactgcaac actcgaggct gaatataatc tagaaagagc gcagaaggcg 1860

gtgaatgcgc tgtttacgtc tacaaaccaa ctagggctaa aaacaaatgt aacggattat 1920

catattgatc aagtgtccaa tttagttacg tatttatcgg atgaattttg tctggatgaa 1980

aagcgagaat tgtccgagaa agtcaaacat gcgaagcgac tcagtgatga acgcaattta 2040

ctccaagatt caaatttcaa agacattaat aggcaaccag aacgtgggtg gggcggaagt 2100

acagggatta ccatccaagg aggggatgac gtatttaaag aaaattacgt cacactatca 2160

ggtacctttg atgagtgcta tccaacatat ttgtatcaaa aaatcgatga atcaaaatta 2220

aaagccttta cccgttatca attaagaggg tatatcgaag atagtcaaga cttagaaatc 2280

tatttaattc gctacaatgc aaaacatgaa acagtaaatg tgccaggtac gggttcctta 2340

tggccgcttt cagcccaaag tccaatcgga aagtgtggag agccgaatcg atgcgcgcca 2400

caccttgaat ggaatcctga cttagattgt tcgtgtaggg atggagaaaa gtgtgcccat 2460

cattcgcatc atttctcctt agacattgat gtaggatgta cagacttaaa tgaggaccta 2520

ggtgtatggg tgatctttaa gattaagacg caagatgggc acgcaagact agggaatcta 2580

gagtttctcg aagaaaaacc attagtagga gaagcgctag ctcgtgtgaa aagagcggag 2640

aaaaaatgga gagacaaacg tgaaaaattg gaatgggaaa caaatatcgt ttataaagag 2700

gcaaaagaat ctgtagatgc tttatttgta aactctcaat atgatcaatt acaagcggat 2760

acgaatattg ccatgattca tgcggcagat aaacgtgttc atagcattcg agaagcttat 2820

ctgcctgagc tgtctgtgat tccgggtgtc aatgcggcta tttttgaaga attagaaggg 2880

cgtattttca ctgcattctc cctatatgat gcgagaaatg tcattaaaaa tggtgatttt 2940

aataatggct tatcctgctg gaacgtgaaa gggcatgtag atgtagaaga acaaaacaac 3000

caacgttcgg tccttgttgt tccggaatgg gaagcagaag tgtcacaaga agttcgtgtc 3060

tgtccgggtc gtggctatat ccttcgtgtc acagcgtaca aggagggata tggagaaggt 3120

tgcgtaacca ttcatgagat cgagaacaat acagacgaac tgaagtttag caactgcgta 3180

gaggaggaaa tctatccaaa taacacggta acgtgtaatg attatactgt aaatcaagaa 3240

gaatacggag gtgcgtacac ttctcgtaat cgaggatata acgaagctcc ttccgtacca 3300

gctgattatg cgtcagtcta tgaagaaaaa tcgtatacag atggacgtag agagaatcct 3360

tgtgaattta acagagggta tagggattac acgccactac cagttggtta tgtgacaaaa 3420

gaattagaat acttcccaga aaccgataag gtatggattg agattggaga aacggaagga 3480

acatttatcg tggacagcgt ggaattactc cttatggagg aatag 3525

<210> 93
<211> 1174
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC838.

<400> 93

Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
20 25 30

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
50 55 60

Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65 70 75 80

Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
85 90 95

Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
100 105 110

Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Thr Arg
115 120 125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
130 135 140

Pro Ser Phe Ala Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr

```
145                 150                 155                 160

Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
                165                 170                 175

Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr
            180                 185                 190

Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
        195                 200                 205

Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
    210                 215                 220

Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
                245                 250                 255

Gln Ser Val Gly Gln Leu Thr Arg Glu Ile Tyr Thr Asp Pro Leu Ile
            260                 265                 270

Thr Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
        275                 280                 285

Val Met Glu Ser Asn Ala Ile Arg Thr Pro His Leu Phe Asp Val Leu
    290                 295                 300

Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
305                 310                 315                 320

Tyr Trp Gly Gly His Arg Val Ile Ser Asn Arg Ile Gly Gly Gly Asn
                325                 330                 335

Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
            340                 345                 350

Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
        355                 360                 365

Phe Arg Pro Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
    370                 375                 380

Gly Val Glu Gly Val Glu Phe Ser Thr Pro Leu Asn Ser Phe Thr Tyr
385                 390                 395                 400

Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu Asp
                405                 410                 415

Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
            420                 425                 430

Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Pro Val
        435                 440                 445

Phe Ser Trp Thr His Arg Ser Ala Thr Asp Arg Asn Ile Ile Ala Ser
    450                 455                 460

Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Asn Phe Leu Phe Asn
465                 470                 475                 480
```

```
Gly Ser Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp Leu Val Arg
                485                 490                 495

Leu Asn Ser Ser Gly Asn Asn Ile Gln Asn Arg Gly Tyr Ile Glu Val
            500                 505                 510

Pro Ile His Phe Pro Ser Thr Ser Thr Arg Tyr Arg Val Arg Val Arg
        515                 520                 525

Tyr Ala Ser Val Thr Pro Ile His Leu Asn Val Asn Trp Gly Asn Ser
    530                 535                 540

Ser Ile Phe Ser Asn Thr Val Pro Ala Thr Ala Thr Ser Leu Asp Asn
545                 550                 555                 560

Leu Gln Ser Ser Asp Phe Gly Tyr Phe Glu Ser Ala Asn Ala Phe Thr
                565                 570                 575

Ser Ser Leu Gly Asn Ile Val Gly Val Arg Asn Phe Ser Gly Thr Ala
            580                 585                 590

Gly Val Ile Ile Asp Arg Phe Glu Phe Ile Pro Val Thr Ala Thr Leu
        595                 600                 605

Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu
    610                 615                 620

Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val Thr Asp Tyr
625                 630                 635                 640

His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe
                645                 650                 655

Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys
            660                 665                 670

Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn Phe Lys Asp
        675                 680                 685

Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr
    690                 695                 700

Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser
705                 710                 715                 720

Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp
                725                 730                 735

Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile
            740                 745                 750

Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys
        755                 760                 765

His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser
    770                 775                 780

Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro
785                 790                 795                 800
```

```
His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu
                805                 810                 815

Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly
            820                 825                 830

Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile
        835                 840                 845

Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu
    850                 855                 860

Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu
865                 870                 875                 880

Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile
                885                 890                 895

Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser
            900                 905                 910

Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala
        915                 920                 925

Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu
    930                 935                 940

Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly
945                 950                 955                 960

Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys
                965                 970                 975

Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His
            980                 985                 990

Val Asp Val Glu Glu Gln Asn Asn  Gln Arg Ser Val Leu  Val Val Pro
        995                 1000                1005

Glu Trp  Glu Ala Glu Val Ser  Gln Glu Val Arg Val  Cys Pro Gly
    1010                1015                1020

Arg Gly  Tyr Ile Leu Arg Val  Thr Ala Tyr Lys Glu  Gly Tyr Gly
    1025                1030                1035

Glu Gly  Cys Val Thr Ile His  Glu Ile Glu Asn Asn  Thr Asp Glu
    1040                1045                1050

Leu Lys  Phe Ser Asn Cys Val  Glu Glu Glu Ile Tyr  Pro Asn Asn
    1055                1060                1065

Thr Val  Thr Cys Asn Asp Tyr  Thr Val Asn Gln Glu  Glu Tyr Gly
    1070                1075                1080

Gly Ala  Tyr Thr Ser Arg Asn  Arg Gly Tyr Asn Glu  Ala Pro Ser
    1085                1090                1095

Val Pro  Ala Asp Tyr Ala Ser  Val Tyr Glu Glu Lys  Ser Tyr Thr
    1100                1105                1110

Asp Gly  Arg Arg Glu Asn Pro  Cys Glu Phe Asn Arg  Gly Tyr Arg
```

1115 1120 1125

Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys Glu Leu Glu
1130 1135 1140

Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr
1145 1150 1155

Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu
1160 1165 1170

Glu

<210> 94
<211> 3558
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC839.

<400> 94
atggaggaaa ataatcaaaa tcaatgcata ccttacaatt gtttaagtaa tcctgaagaa 60

gtacttttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca 120

cttgttcagt ttctggtatc taactttgta ccaggggggag gatttttagt tggattaata 180

gattttgtat ggggaatagt tggcccttct caatgggatg catttctagt acaaattgaa 240

caattaatta atgaaagaat agctgaattt gctaggaatg ctgctattgc taatttagaa 300

ggattaggaa acaatttcaa tatatatgtg gaagcattta aagaatggga agaagatcct 360

aataatccag caaccaggac cagagtaatt gatcgctttc gtatacttga tgggctactt 420

gaaagggaca ttccttcgtt tgcaatttct ggatttgaag taccccttttt atccgtttat 480

gctcaagcgg ccaatctgca tctagctata ttaagagatt ctgtaatttt tggagaaaga 540

tggggattga caacgataaa tgtcaatgaa aactataata gactaattag gcatattgat 600

gaatatgctg atcactgtgc aaatacgtat aatcggggat taaataattt accgaaatct 660

acgtatcaag attggataac atataatcga ttacggagag acttaacatt gactgtatta 720

gatatcgccg ctttctttcc aaactatgac aataggagat atccaattca gtcagttggt 780

caactaacaa gggaaattta tacggaccca ttaattactt ttaatccaca gttacagtct 840

gtagctcaat tacctacttt taacgttatg gaaagcaacg caattagaac tcctcattta 900

tttgatgtat tgaataatct tacaattttt acagattggt ttagtgttgg acgcaacttt 960

tattggggag gacatcgagt aatatctaac cgtataggag gaggtaacat aacatctcct 1020

atatatggaa gggaggcgaa tcaggagcct ccaagatctt ttacttttaa tgggcctgtt 1080

tttaggactt tatcaaatcc tacttttaga cctttacagc aaccttggcc agcgccacca 1140

tttaatttac gtggtgttga aggagtagaa ttttctacac ctttaaatag ctttacgtat 1200

cgaggaaggg gtacggttga ttctttaact gagttaccgc ctgaggataa tagtgtgcct 1260

cctcgcgaag gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaacc 1320

ccatttttaa caactggtcc agtattttct tggacgcatc gtagtgctac tgatcgaaat 1380

acaattgatc cagagagaat taatcaaata cctttagtga aaggatttag agtttggggg 1440

ggcacctctg tcattacagg accaggattt acaggagggg atatccttcg aagaaatacc 1500

tttggtgatt ttgtatctct acaagtcaat attaattcac caattaccca aagataccgt 1560

ttaagatttc gttacgcttc cagtagggat gcacgagtta tagtattaac aggagcggca 1620

tccacaggag tgggaggcca agttagtgta aatatgcctc tccagaaaac tatggaaata 1680

ggggagaact taacatctag aacatttaga tataccgatt ttagtaatcc tttttcattt 1740

agagctaatc cagatataat tgggataagt gaacaacctc tatttggtgc aggttctatt 1800

agtagcggtg aactttatat agataaaatt gaaattattc ctgttactgc aacactcgag 1860

gctgaatata atctagaaag agcgcagaag gcggtgaatg cgctgtttac gtctacaaac 1920

caactagggc taaaaacaaa tgtaacggat tatcatattg atcaagtgtc caatttagtt 1980

acgtatttat cggatgaatt ttgtctggat gaaaagcgag aattgtccga gaaagtcaaa 2040

catgcgaagc gactcagtga tgaacgcaat ttactccaag attcaaattt caaagacatt 2100

aataggcaac cagaacgtgg gtggggcgga agtacaggga ttaccatcca aggaggggat 2160

gacgtattta aagaaaatta cgtcacacta tcaggtacct ttgatgagtg ctatccaaca 2220

tatttgtatc aaaaaatcga tgaatcaaaa ttaaaagcct ttacccgtta tcaattaaga 2280

gggtatatcg aagatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacat 2340

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagccca aagtccaatc 2400

ggaaagtgtg gagagccgaa tcgatgcgcg ccacaccttg aatggaatcc tgacttagat 2460

tgttcgtgta gggatggaga aaagtgtgcc catcattcgc atcatttctc cttagacatt 2520

gatgtaggat gtacagactt aaatgaggac ctaggtgtat gggtgatctt taagattaag 2580

acgcaagatg ggcacgcaag actagggaat ctagagtttc tcgaagaaaa accattagta 2640

ggagaagcgc tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgtgaaaaa 2700

ttggaatggg aaacaaatat cgtttataaa gaggcaaaag aatctgtaga tgctttattt 2760

gtaaactctc aatatgatca attacaagcg gatacgaata ttgccatgat tcatgcggca 2820

gataaacgtg ttcatagcat tcgagaagct tatctgcctg agctgtctgt gattccgggt 2880

gtcaatgcgg ctatttttga agaattagaa gggcgtattt tcactgcatt ctccctatat 2940

gatgcgagaa atgtcattaa aaatggtgat tttaataatg gcttatcctg ctggaacgtg 3000

aaagggcatg tagatgtaga agaacaaaac aaccaacgtt cggtccttgt tgttccggaa 3060

tgggaagcag aagtgtcaca agaagttcgt gtctgtccgg gtcgtggcta tatccttcgt 3120

gtcacagcgt acaaggaggg atatggagaa ggttgcgtaa ccattcatga gatcgagaac 3180

aatacagacg aactgaagtt tagcaactgc gtagaggagg aaatctatcc aaataacacg 3240

gtaacgtgta atgattatac tgtaaatcaa gaagaatacg gaggtgcgta cacttctcgt 3300

aatcgaggat ataacgaagc tccttccgta ccagctgatt atgcgtcagt ctatgaagaa 3360

aaatcgtata cagatggacg tagagagaat ccttgtgaat ttaacagagg gtatagggat 3420

tacacgccac taccagttgg ttatgtgaca aaagaattag aatacttccc agaaaccgat 3480

aaggtatgga ttgagattgg agaaacggaa ggaacattta tcgtggacag cgtggaatta 3540

ctccttatgg aggaatag 3558

<210> 95
<211> 1185
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC839.

<400> 95

Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
20 25 30

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
50 55 60

Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65 70 75 80

Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
85 90 95

Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
100 105 110

Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Thr Arg
115 120 125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
130 135 140

Pro Ser Phe Ala Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145 150 155 160

```
Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
                165                 170                 175

Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr
            180                 185                 190

Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
        195                 200                 205

Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
    210                 215                 220

Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
                245                 250                 255

Gln Ser Val Gly Gln Leu Thr Arg Glu Ile Tyr Thr Asp Pro Leu Ile
            260                 265                 270

Thr Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
        275                 280                 285

Val Met Glu Ser Asn Ala Ile Arg Thr Pro His Leu Phe Asp Val Leu
    290                 295                 300

Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
305                 310                 315                 320

Tyr Trp Gly Gly His Arg Val Ile Ser Asn Arg Ile Gly Gly Gly Asn
                325                 330                 335

Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
            340                 345                 350

Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
        355                 360                 365

Phe Arg Pro Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
    370                 375                 380

Gly Val Glu Gly Val Glu Phe Ser Thr Pro Leu Asn Ser Phe Thr Tyr
385                 390                 395                 400

Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu Asp
                405                 410                 415

Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
            420                 425                 430

Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Pro Val
        435                 440                 445

Phe Ser Trp Thr His Arg Ser Ala Thr Asp Arg Asn Thr Ile Asp Pro
    450                 455                 460

Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp Gly
465                 470                 475                 480

Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu
```

```
                485                 490                 495

Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile Asn
            500                 505                 510

Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser Ser
        515                 520                 525

Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly Val
    530                 535                 540

Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu Ile
545                 550                 555                 560

Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser Asn
                565                 570                 575

Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu Gln
            580                 585                 590

Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp
        595                 600                 605

Lys Ile Glu Ile Ile Pro Val Thr Ala Thr Leu Glu Ala Glu Tyr Asn
    610                 615                 620

Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn
625                 630                 635                 640

Gln Leu Gly Leu Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val
                645                 650                 655

Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys
            660                 665                 670

Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu
        675                 680                 685

Arg Asn Leu Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro
    690                 695                 700

Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp
705                 710                 715                 720

Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu
                725                 730                 735

Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys
            740                 745                 750

Ala Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp
        755                 760                 765

Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn
    770                 775                 780

Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile
785                 790                 795                 800

Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn
                805                 810                 815
```

```
Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala His His
            820                 825                 830

Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn
        835                 840                 845

Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly
    850                 855                 860

His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val
865                 870                 875                 880

Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp
                885                 890                 895

Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala
            900                 905                 910

Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Gln Leu
        915                 920                 925

Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val
    930                 935                 940

His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly
945                 950                 955                 960

Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala
                965                 970                 975

Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn
            980                 985                 990

Asn Gly Leu Ser Cys Trp Asn Val  Lys Gly His Val Asp  Val Glu Glu
        995                 1000                1005

Gln Asn  Asn Gln Arg Ser Val  Leu Val Val Pro Glu  Trp Glu Ala
    1010                1015                1020

Glu Val  Ser Gln Glu Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile
    1025                1030                1035

Leu Arg  Val Thr Ala Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val
    1040                1045                1050

Thr Ile  His Glu Ile Glu Asn  Asn Thr Asp Glu Leu  Lys Phe Ser
    1055                1060                1065

Asn Cys  Val Glu Glu Glu Ile  Tyr Pro Asn Asn Thr  Val Thr Cys
    1070                1075                1080

Asn Asp  Tyr Thr Val Asn Gln  Glu Glu Tyr Gly Gly  Ala Tyr Thr
    1085                1090                1095

Ser Arg  Asn Arg Gly Tyr Asn  Glu Ala Pro Ser Val  Pro Ala Asp
    1100                1105                1110

Tyr Ala  Ser Val Tyr Glu Glu  Lys Ser Tyr Thr Asp  Gly Arg Arg
    1115                1120                1125
```

Glu Asn Pro Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr Pro
1130 1135 1140

Leu Pro Val Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu
1145 1150 1155

Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe
1160 1165 1170

Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1175 1180 1185

<210> 96
<211> 3519
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC841.

<400> 96
atggaggaaa ataatcaaaa tcaatgcata ccttacaatt gtttaagtaa tcctgaagaa 60

gtacttttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca 120

cttgttcagt ttctggtatc taactttgta ccaggggggag gatttttagt tggattaata 180

gattttgtat ggggaatagt tggcccttct caatgggatg catttctagt acaaattgaa 240

caattaatta atgaaagaat agctgaattt gctaggaatg ctgctattgc taatttagaa 300

ggattaggaa acaatttcaa tatatatgtg gaagcattta aagaatggga agaagatcct 360

aataatccag caaccaggac cagagtaatt gatcgctttc gtatacttga tgggctactt 420

gaaagggaca ttccttcgtt tgcaatttct ggatttgaag taccccttttt atccgtttat 480

gctcaagcgg ccaatctgca tctagctata ttaagagatt ctgtaatttt tggagaaaga 540

tggggattga caacgataaa tgtcaatgaa aactataata gactaattag gcatattgat 600

gaatatgctg atcactgtgc aaatacgtat aatcggggat taaataattt accgaaatct 660

acgtatcaag attggataac atataatcga ttacggagag acttaacatt gactgtatta 720

gatatcgccg ctttctttcc aaactatgac aataggagat atccaattca gtcagttggt 780

caactaacaa gggaaattta tacggaccca ttaattactt ttaatccaca gttacagtct 840

gtagctcaat tacctacttt taacgttatg gaaagcaacg caattagaac tcctcattta 900

tttgatgtat taaataatct tacaattttt acagattggt ttagtgttgg acgcaacttt 960

tattggggag gacatcgagt aatatctaac cgtataggag gaggtaacat aacatctcct 1020

atatatggaa gggaggcgaa tcaggagcct ccaagatctt ttacttttaa tgggcctgtt 1080

tttaggactt tatcaaatcc tacttttaga cctttacagc aaccttggcc agcgccacca 1140

tttaatttac gtggtgttga aggagtagaa ttttctacac ctttaaatag ctttacgtat 1200

cgaggaaggg gtacggttga ttctttaact gagttaccgc ctgaggataa tagtgtgcct 1260

cctcgcgaag gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaacc 1320

ccatttttaa caactggtcc agtattttct tggacgcatc gtagtgctac tgatcgaaat 1380

acaattgatc cagagagaat taatcaaata cctttagtga aaggatttag agtttggggg 1440

ggcacctctg tcattacagg accaggattt acaggagggg atatccttag aagaacgaat 1500

actggcacat ttgcagatat gagagtaaat attactgggc cattatccca aagatatcgt 1560

gtaagaattc gctatgcttc tacgacagat ttacaatttt tcacgagaat caatggaact 1620

tctgtaaatc aaggtaattt ccaaagaact atgaatagag gggataattt agaatctgga 1680

aactttagga ctgcaggatt tagtacgcct tttagttttt caaatgcgca aagtacattc 1740

acattgggta ctcaggcttt ttcaaatcag gaagtttata tagatcgaat tgaatttgtc 1800

ccggcagaag taacattcga ggcagaatct gatttagaaa gagcgcagaa ggcggtgaat 1860

gcgctgttta cgtctacaaa ccaactaggg ctaaaaacaa atgtaacgga ttatcatatt 1920

gatcaagtgt ccaatttagt tacgtattta tcggatgaat tttgtctgga tgaaaagcga 1980

gaattgtccg agaaagtcaa acatgcgaag cgactcagtg atgaacgcaa tttactccaa 2040

gattcaaatt tcaaagacat taataggcaa ccagaacgtg ggtggggcgg aagtacaggg 2100

attaccatcc aaggagggga tgacgtattt aaagaaaatt acgtcacact atcaggtacc 2160

tttgatgagt gctatccaac atatttgtat caaaaaatcg atgaatcaaa attaaaagcc 2220

tttacccgtt atcaattaag agggtatatc gaagatagtc aagacttaga aatctattta 2280

attcgctaca atgcaaaaca tgaaacagta aatgtgccag gtacgggttc cttatggccg 2340

ctttcagccc aaagtccaat cggaaagtgt ggagagccga atcgatgcgc gccacacctt 2400

gaatggaatc ctgacttaga ttgttcgtgt agggatggag aaaagtgtgc ccatcattcg 2460

catcatttct ccttagacat tgatgtagga tgtacagact taaatgagga cctaggtgta 2520

tgggtgatct ttaagattaa gacgcaagat gggcacgcaa gactagggaa tctagagttt 2580

ctcgaagaaa aaccattagt aggagaagcg ctagctcgtg tgaaaagagc ggagaaaaaa 2640

tggagagaca aacgtgaaaa attggaatgg gaaacaaata tcgtttataa agaggcaaaa 2700

gaatctgtag atgctttatt tgtaaactct caatatgatc aattacaagc ggatacgaat 2760

attgccatga ttcatgcggc agataaacgt gttcatagca ttcgagaagc ttatctgcct 2820

gagctgtctg tgattccggg tgtcaatgcg gctatttttg aagaattaga agggcgtatt 2880

ttcactgcat tctccctata tgatgcgaga aatgtcatta aaaatggtga ttttaataat 2940

ggcttatcct gctggaacgt gaaagggcat gtagatgtag aagaacaaaa caaccaacgt 3000

tcggtccttg ttgttccgga atgggaagca gaagtgtcac aagaagttcg tgtctgtccg 3060

ggtcgtggct atatccttcg tgtcacagcg tacaaggagg gatatggaga aggttgcgta 3120

accattcatg agatcgagaa caatacagac gaactgaagt ttagcaactg cgtagaggag 3180

gaaatctatc caaataacac ggtaacgtgt aatgattata ctgtaaatca agaagaatac 3240

ggaggtgcgt acacttctcg taatcgagga tataacgaag ctccttccgt accagctgat 3300

tatgcgtcag tctatgaaga aaaatcgtat acagatggac gtagagagaa tccttgtgaa 3360

tttaacagag ggtataggga ttacacgcca ctaccagttg gttatgtgac aaaagaatta 3420

gaatacttcc cagaaaccga taaggtatgg attgagattg gagaaacgga aggaacattt 3480

atcgtggaca gcgtggaatt actccttatg gaggaataa 3519

<210> 97
<211> 1172
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC841.

<400> 97

Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
20 25 30

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
50 55 60

Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65 70 75 80

Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
85 90 95

Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
100 105 110

Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Thr Arg
115 120 125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
130 135 140

Pro Ser Phe Ala Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145 150 155 160

Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
165 170 175

Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr

```
            180                 185                 190

Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
        195                 200                 205

Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
    210                 215                 220

Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
                245                 250                 255

Gln Ser Val Gly Gln Leu Thr Arg Glu Ile Tyr Thr Asp Pro Leu Ile
            260                 265                 270

Thr Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
        275                 280                 285

Val Met Glu Ser Asn Ala Ile Arg Thr Pro His Leu Phe Asp Val Leu
    290                 295                 300

Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
305                 310                 315                 320

Tyr Trp Gly Gly His Arg Val Ile Ser Asn Arg Ile Gly Gly Gly Asn
                325                 330                 335

Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
            340                 345                 350

Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
        355                 360                 365

Phe Arg Pro Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
    370                 375                 380

Gly Val Glu Gly Val Glu Phe Ser Thr Pro Leu Asn Ser Phe Thr Tyr
385                 390                 395                 400

Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu Asp
                405                 410                 415

Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
            420                 425                 430

Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Pro Val
        435                 440                 445

Phe Ser Trp Thr His Arg Ser Ala Thr Asp Arg Asn Thr Ile Asp Pro
    450                 455                 460

Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp Gly
465                 470                 475                 480

Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu
                485                 490                 495

Arg Arg Thr Asn Thr Gly Thr Phe Ala Asp Met Arg Val Asn Ile Thr
            500                 505                 510
```

```
Gly Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr Ala Ser Thr
        515                 520                 525

Thr Asp Leu Gln Phe Phe Thr Arg Ile Asn Gly Thr Ser Val Asn Gln
    530                 535                 540

Gly Asn Phe Gln Arg Thr Met Asn Arg Gly Asp Asn Leu Glu Ser Gly
545                 550                 555                 560

Asn Phe Arg Thr Ala Gly Phe Ser Thr Pro Phe Ser Phe Ser Asn Ala
                565                 570                 575

Gln Ser Thr Phe Thr Leu Gly Thr Gln Ala Phe Ser Asn Gln Glu Val
            580                 585                 590

Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu Val Thr Phe Glu Ala
        595                 600                 605

Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr
    610                 615                 620

Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val Thr Asp Tyr His Ile
625                 630                 635                 640

Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser Asp Glu Phe Cys Leu
                645                 650                 655

Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu
            660                 665                 670

Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn Phe Lys Asp Ile Asn
        675                 680                 685

Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr Gly Ile Thr Ile Gln
    690                 695                 700

Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Ser Gly Thr
705                 710                 715                 720

Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser
                725                 730                 735

Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp
            740                 745                 750

Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu
        755                 760                 765

Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln
    770                 775                 780

Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu
785                 790                 795                 800

Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys
                805                 810                 815

Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr
            820                 825                 830
```

```
Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr
        835                 840                 845

Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys
    850                 855                 860

Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys
865                 870                 875                 880

Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr
                885                 890                 895

Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr
            900                 905                 910

Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp
        915                 920                 925

Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val
    930                 935                 940

Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile
945                 950                 955                 960

Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly
                965                 970                 975

Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp
            980                 985                 990

Val Glu Glu Gln Asn Asn Gln Arg  Ser Val Leu Val Val  Pro Glu Trp
        995                 1000                1005

Glu Ala  Glu Val Ser Gln Glu  Val Arg Val Cys Pro  Gly Arg Gly
    1010                1015                1020

Tyr Ile  Leu Arg Val Thr Ala  Tyr Lys Glu Gly Tyr  Gly Glu Gly
    1025                1030                1035

Cys Val  Thr Ile His Glu Ile  Glu Asn Asn Thr Asp  Glu Leu Lys
    1040                1045                1050

Phe Ser  Asn Cys Val Glu Glu  Glu Ile Tyr Pro Asn  Asn Thr Val
    1055                1060                1065

Thr Cys  Asn Asp Tyr Thr Val  Asn Gln Glu Glu Tyr  Gly Gly Ala
    1070                1075                1080

Tyr Thr  Ser Arg Asn Arg Gly  Tyr Asn Glu Ala Pro  Ser Val Pro
    1085                1090                1095

Ala Asp  Tyr Ala Ser Val Tyr  Glu Glu Lys Ser Tyr  Thr Asp Gly
    1100                1105                1110

Arg Arg  Glu Asn Pro Cys Glu  Phe Asn Arg Gly Tyr  Arg Asp Tyr
    1115                1120                1125

Thr Pro  Leu Pro Val Gly Tyr  Val Thr Lys Glu Leu  Glu Tyr Phe
    1130                1135                1140

Pro Glu  Thr Asp Lys Val Trp  Ile Glu Ile Gly Glu  Thr Glu Gly
```

1145 1150 1155

Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1160 1165 1170

<210> 98
<211> 3516
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC842.

<400> 98
atggcagaga ataatattca aaatcaatgc gtaccttaca attgtttaaa taatcctgaa 60

gtagaaatat taaatgaaga aagaagtact ggcagattac cgttagatat atccttatcg 120

cttacacgtt tccttttgag tgaatttgtt ccaggtgtgg gagttgcgtt tggattattt 180

gatttaatat ggggttttat aactccttct gattggagct tatttctttt acagattgaa 240

caattgattg agcaaagaat agaaacattg gaaaggaacc gggcaattac tacattacga 300

gggttagcag atagctatga aatttatatt gaagcactaa gagagtggga agcaaatcct 360

aataatgcac aattaaggga agatgtgcgt attcgatttg ctaatacaga cgacgcttta 420

ataacagcaa taaataattt tacacttaca agttttgaaa tccctctttt atcggtctat 480

gttcaagcgg cgaatttaca tttatcacta ttaagagacg ctgtatcgtt tgggcagggt 540

tggggactgg atatagctac tgttaataat cattataata gattaataaa tcttattcat 600

agatatacga aacattgttt ggacacatac aatcaaggat tagaaaactt aagaggtact 660

aatactcgac aatgggcaag attcaatcag tttaggagag atttaacgct aactgtatta 720

gatatcgttg ctctttttcc gaactacgat gttagaacat atccaattca aacgtcatcc 780

caattaacaa gggaaattta tacaagttca gtaattgagg attctccagt ttctgctaat 840

atacctaatg gttttaatag ggcggaattt ggagttagac cgccccatct tatggacttt 900

atgaattctt tgtttgtaac tgcagagact gttagaagtc aaactgtgtg gggaggacac 960

ttagttagtt cacgaaatac ggctggtaac cgtataaatt tccctagtta cggggtcttc 1020

aatcctggtg gcgccatttg gattgcagat gaggatccac gtccttttta tcggacatta 1080

tcagatcctg tttttgtccg aggaggattt gggaatcctc attatgtact ggggcttagg 1140

ggagtagcat ttcaacaaac tggtacgaac cacacccgaa catttagaaa tagtgggacc 1200

atagattctc tagatgaaat cccacctcag gataatagtg gggcaccttg gaatgattat 1260

agtcatgtat taaatcatgt tacatttgta cgatggccag gtgagatttc aggaagtgat 1320

tcatggagag ctccaatgtt ttcttggacg caccgtagtg caacccctac aaatacaatt 1380

gatccggaga ggattactca aataccattg gtaaaagcac atacacttca gtcaggtact 1440

actgttgtaa gagggcccgg gtttacggga ggagatattc ttcgacgaac aagtggagga 1500

ccatttgctt atactattgt taatataaat gggcaattac cccaaaggta tcgtgcaaga 1560

atacgctatg cctctactac aaatctaaga atttacgtaa cggttgcagg tgaacggatt 1620

tttgctggtc aatttaacaa aacaatggat accggtgacc cattaacatt ccaatctttt 1680

agttacgcaa ctattaatac agcttttaca ttcccaatga gccagagtag tttcacagta 1740

ggtgctgata cttttagttc agggaatgaa gtttatatag acagatttga attgattcca 1800

gttactgcaa catttgaagc agaatatgat ttagaaagag cacaaaaggc ggtgaatgcg 1860

ctgtttactt ctataaacca aatagggata aaaacagatg tgacggatta tcatattgat 1920

caagtatcca atttagtgga ttgtttatca gatgaatttt gtctggatga aaagcgagaa 1980

ttgtccgaga aagtcaaaca tgcgaagcga ctcagtgatg agcggaattt acttcaagat 2040

ccaaacttca aaggcatcaa taggcaacta gaccgtggtt ggagaggaag tacggatatt 2100

accatccaaa gaggagatga cgtattcaaa gaaaattatg tcacactacc aggtaccttt 2160

gatgagtgct atccaacata tttgtatcaa aaaatcgatg aatcaaaatt aaaagccttt 2220

acccgttatc aattaagagg gtatatcgaa gatagtcaag acttagaaat ctatttaatt 2280

cgctacaatg caaaacatga aacagtaaat gtgccaggta cgggttcctt atggccgctt 2340

tcagcccaaa gtccaatcgg aaagtgtgga gagccgaatc gatgcgcgcc acaccttgaa 2400

tggaatcctg acttagattg ttcgtgtagg gatggagaaa agtgtgccca tcattcgcat 2460

catttctcct tagacattga tgtaggatgt acagacttaa atgaggacct aggtgtatgg 2520

gtgatcttta agattaagac gcaagatggg cacgcaagac tagggaatct agagtttctc 2580

gaagagaaac cattagtagg agaagcgcta gctcgtgtga aaagagcgga gaaaaaatgg 2640

agagacaaac gtgaaaaatt ggaatgggaa acaaatatcg tttataaaga ggcaaaagaa 2700

tctgtagatg ctttatttgt aaactctcaa tatgatcaat tacaagcgga tacgaatatt 2760

gccatgattc atgcggcaga taaacgtgtt catagcattc gagaagctta tctgcctgag 2820

ctgtctgtga ttccgggtgt caatgcggct atttttgaag aattagaagg gcgtattttc 2880

actgcattct ccctatatga tgcgagaaat gtcattaaaa atggtgattt taataatggc 2940

ttatcctgct ggaacgtgaa agggcatgta gatgtagaag aacaaaacaa ccaacgttcg 3000

gtccttgttg ttccggaatg ggaagcagaa gtgtcacaag aagttcgtgt ctgtccgggt 3060

cgtggctata tccttcgtgt cacagcgtac aaggagggat atggagaagg ttgcgtaacc 3120

attcatgaga tcgagaacaa tacagacgaa ctgaagttta gcaactgcgt agaagaggaa 3180

atctatccaa ataacacggt aacgtgtaat gattatactg taaatcaaga agaatacgga 3240

ggtgcgtaca cttctcgtaa tcgaggatat aacgaagctc cttccgtacc agctgattat 3300

gcgtcagtct atgaagaaaa atcgtataca gatggacgaa gagagaatcc ttgtgaattt 3360

aacagagggt atagggatta cacgccacta ccagttggtt atgtgacaaa agaattagaa 3420

tacttcccag aaaccgataa ggtatggatt gagattggag aaacggaagg aacatttatc 3480

gtggacagcg tggaattact ccttatggag gaataa 3516

<210> 99
<211> 1171
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC842.

<400> 99

Met Ala Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1 5 10 15

Asn Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg
20 25 30

Leu Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu
35 40 45

Phe Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp
50 55 60

Gly Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu
65 70 75 80

Gln Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile
85 90 95

Thr Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala
100 105 110

Leu Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp
115 120 125

Val Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile
130 135 140

Asn Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr
145 150 155 160

Val Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser
165 170 175

Phe Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr
180 185 190

Asn Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp
195 200 205

Thr Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln
210 215 220

```
Trp Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile
                245                 250                 255

Gln Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile
            260                 265                 270

Glu Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala
        275                 280                 285

Glu Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu
    290                 295                 300

Phe Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His
305                 310                 315                 320

Leu Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser
                325                 330                 335

Tyr Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp
            340                 345                 350

Pro Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly
        355                 360                 365

Gly Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe
    370                 375                 380

Gln Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr
385                 390                 395                 400

Ile Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro
                405                 410                 415

Trp Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp
            420                 425                 430

Pro Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser
        435                 440                 445

Trp Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg
    450                 455                 460

Ile Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr
465                 470                 475                 480

Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg
                485                 490                 495

Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln
            500                 505                 510

Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn
        515                 520                 525

Leu Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln
    530                 535                 540
```

```
Phe Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe
545                 550                 555                 560

Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser
                565                 570                 575

Ser Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr
            580                 585                 590

Ile Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu Ala Glu
        595                 600                 605

Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser
    610                 615                 620

Ile Asn Gln Ile Gly Ile Lys Thr Asp Val Thr Asp Tyr His Ile Asp
625                 630                 635                 640

Gln Val Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp
                645                 650                 655

Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser
            660                 665                 670

Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Lys Gly Ile Asn Arg
        675                 680                 685

Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Arg
    690                 695                 700

Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr Phe
705                 710                 715                 720

Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys
                725                 730                 735

Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser
            740                 745                 750

Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr
        755                 760                 765

Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser
    770                 775                 780

Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu
785                 790                 795                 800

Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg Asp Gly Glu Lys Cys Ala
                805                 810                 815

His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp
            820                 825                 830

Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln
        835                 840                 845

Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro
    850                 855                 860

Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp
```

865 870 875 880

Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys
885 890 895

Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp
900 905 910

Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys
915 920 925

Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile
930 935 940

Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe
945 950 955 960

Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp
965 970 975

Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val
980 985 990

Glu Glu Gln Asn Asn Gln Arg Ser Val Leu Val Val Pro Glu Trp Glu
995 1000 1005

Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr
1010 1015 1020

Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys
1025 1030 1035

Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe
1040 1045 1050

Ser Asn Cys Val Glu Glu Glu Ile Tyr Pro Asn Asn Thr Val Thr
1055 1060 1065

Cys Asn Asp Tyr Thr Val Asn Gln Glu Glu Tyr Gly Gly Ala Tyr
1070 1075 1080

Thr Ser Arg Asn Arg Gly Tyr Asn Glu Ala Pro Ser Val Pro Ala
1085 1090 1095

Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ser Tyr Thr Asp Gly Arg
1100 1105 1110

Arg Glu Asn Pro Cys Glu Phe Asn Arg Gly Tyr Arg Asp Tyr Thr
1115 1120 1125

Pro Leu Pro Val Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro
1130 1135 1140

Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr
1145 1150 1155

Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1160 1165 1170

<210> 100
<211> 3444

<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC850.

<400> 100
atggagaata atattcaaaa tcaatgcgta ccttacaatt gtttaaataa tcctgaagta 60

gaaatattaa atgaagaaag aagtactggc agattaccgt tagatatatc cttatcgctt 120

acacgtttcc ttttgagtga atttgttcca ggtgtgggag ttgcgtttgg attatttgat 180

ttaatatggg gttttataac tccttctgat tggagcttat ttcttttaca gattgaacaa 240

ttgattgagc aaagaataga aacattggaa aggaaccggg caattactac attacgaggg 300

ttagcagata gctatgaaat ttatattgaa gcactaagag agtgggaagc aaatcctaat 360

aatgcacaat taagggaaga tgtgcgtatt cgatttgcta atacagacga cgctttaata 420

acagcaataa ataattttac acttacaagt tttgaaatcc ctcttttatc ggtctatgtt 480

caagcggcga atttacattt atcactatta agagacgctg tatcgtttgg gcagggttgg 540

ggactggata tagctactgt taataatcat tataatagat taataaatct tattcataga 600

tatacgaaac attgtttgga cacatacaat caaggattag aaaacttaag aggtactaat 660

actcgacaat gggcaagatt caatcagttt aggagagatt taacgctaac tgtattagat 720

atcgttgctc tttttccgaa ctacgatgtt agaacatatc caattcaaac gtcatcccaa 780

ttaacaaggg aaatttatac aagttcagta attgaggatt ctccagtttc tgctaatata 840

cctaatggtt ttaatagggc ggaatttgga gttagaccgc cccatcttat ggactttatg 900

aattctttgt ttgtaactgc agagactgtt agaagtcaaa ctgtgtgggg aggacactta 960

gttagttcac gaaatacggc tggtaaccgt ataaatttcc ctagttacgg ggtcttcaat 1020

cctggtggcg ccatttggat tgcagatgag gatccacgtc ctttttatcg gacattatca 1080

gatcctgttt ttgtccgagg aggatttggg aatcctcatt atgtactggg gcttagggga 1140

gtagcatttc aacaaactgg tacgaaccac acccgaacat ttagaaatag tgggaccata 1200

gattctctag atgaaatccc acctcaggat aatagtgggg caccttggaa tgattatagt 1260

catgtattaa atcatgttac atttgtacga tggccaggtg agatttcagg aagtgattca 1320

tggagagctc caatgttttc ttggacgcac cgtagtgcaa cccctacaaa tacaattgat 1380

ccggagagga ttactcaaat accattggta aaagcacata cacttcagtc aggtactact 1440

gttgtaagag ggcccgggtt tacgggagga gatatccttc gaagaacgaa tactggcaca 1500

tttgcagata tgagagtaaa tattactggg ccattatccc aaagatatcg tgtaagaatt 1560

cgctatgctt ctacgacaga tttacaattt ttcacgagaa tcaatggaac ttctgtaaat 1620

caaggtaatt tccaaagaac tatgaataga ggggataatt tagaatctgg aaactttagg 1680

actgcaggat ttagtacgcc ttttagtttt tcaaatgcgc aaagtacatt cacattgggt 1740

actcaggctt tttcaaatca ggaagtttat atagatcgaa ttgaatttgt cccggcagaa 1800

gtaacattcg aggcagaatc tgatttagaa agagcacaaa aggcggtgaa tgagctgttt 1860

acttcttcca atcaaatcgg gttaaaaaca gatgtgacgg attatcatat tgatcaagta 1920

tccaatttag ttgagtgttt atctgatgaa ttttgtctgg atgaaaaaaa agaattgtcc 1980

gagaaagtca aacatgcgaa gcgacttagt gatgagcgga atttacttca agatccaaac 2040

tttagaggga tcaatagaca actagaccgt ggctggagag gaagtacgga tattaccatc 2100

caaggaggcg atgacgtatt caaagagaat tacgttacgc tattgggtac ctttgatgag 2160

tgctatccaa cgtatttata tcaaaaaata gatgagtcga aattaaaagc ctatacccgt 2220

taccaattaa gagggtatat cgaagatagt caagacttag aaatctattt aattcgctac 2280

aatgccaaac acgaaacagt aaatgtgcca ggtacgggtt ccttatggcc gctttcagcc 2340

ccaagtccaa tcggaaaatg tgcccatcat tcccatcatt tctccttgga cattgatgtt 2400

ggatgtacag acttaaatga ggacttaggt gtatgggtga tattcaagat taagacgcaa 2460

gatggccatg caagactagg aaatctagaa tttctcgaag agaaaccatt agtaggagaa 2520

gcactagctc gtgtgaaaag agcggagaaa aaatggagag acaaacgtga aaaattggaa 2580

tgggaaacaa atattgttta taaagaggca aaagaatctg tagatgcttt atttgtaaac 2640

tctcaatatg atagattaca agcggatacc aacatcgcga tgattcatgc ggcagataaa 2700

cgcgttcata gcattcgaga agcttatctg cctgagctgt ctgtgattcc gggtgtcaat 2760

gcggctattt ttgaagaatt agaagggcgt attttcactg cattctccct atatgatgcg 2820

agaaatgtca ttaaaaatgg tgattttaat aatggcttat cctgctggaa cgtgaaaggg 2880

catgtagatg tagaagaaca aaacaaccac cgttcggtcc ttgttgttcc ggaatgggaa 2940

gcagaagtgt cacaagaagt tcgtgtctgt ccgggtcgtg gctatatcct tcgtgtcaca 3000

gcgtacaagg agggatatgg agaaggttgc gtaaccattc atgagatcga gaacaataca 3060

gacgaactga agtttagcaa ctgtgtagaa gaggaagtat atccaaacaa cacggtaacg 3120

tgtaatgatt atactgcgac tcaagaagaa tatgagggta cgtacacttc tcgtaatcga 3180

ggatatgacg gagcctatga aagcaattct tctgtaccag ctgattatgc atcagcctat 3240

gaagaaaaag catatacaga tggacgaaga gacaatcctt gtgaatctaa cagaggatat 3300

ggggattaca caccactacc agctggctat gtgacaaaag aattagagta cttcccagaa 3360

accgataagg tatggattga gatcggagaa acggaaggaa cattcatcgt ggacagcgtg 3420

gaattacttc ttatggagga atag 3444

<210> 101
<211> 1147
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC850.

<220>
<221> ĐẶC TÍNH_KHÁC
<222> (1)..(1147)
<223> Trình tự axit amin của protein khảm TIC850.

<400> 101

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1 5 10 15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
20 25 30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
35 40 45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
50 55 60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65 70 75 80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
85 90 95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
100 105 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
115 120 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
130 135 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145 150 155 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
165 170 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
180 185 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
195 200 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
210 215 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225 230 235 240

```
Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
                245                 250                 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
            260                 265                 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
    450                 455                 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Asn Thr Gly Thr Phe Ala Asp Met Arg Val Asn Ile Thr Gly Pro Leu
            500                 505                 510

Ser Gln Arg Tyr Arg Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu
        515                 520                 525

Gln Phe Phe Thr Arg Ile Asn Gly Thr Ser Val Asn Gln Gly Asn Phe
    530                 535                 540

Gln Arg Thr Met Asn Arg Gly Asp Asn Leu Glu Ser Gly Asn Phe Arg
545                 550                 555                 560

Thr Ala Gly Phe Ser Thr Pro Phe Ser Phe Ser Asn Ala Gln Ser Thr
```

565 570 575

Phe Thr Leu Gly Thr Gln Ala Phe Ser Asn Gln Glu Val Tyr Ile Asp
580 585 590

Arg Ile Glu Phe Val Pro Ala Glu Val Thr Phe Glu Ala Glu Ser Asp
595 600 605

Leu Glu Arg Ala Gln Lys Ala Val Asn Glu Leu Phe Thr Ser Ser Asn
610 615 620

Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val
625 630 635 640

Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys
645 650 655

Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu
660 665 670

Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu
675 680 685

Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp
690 695 700

Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu
705 710 715 720

Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys
725 730 735

Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp
740 745 750

Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn
755 760 765

Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile
770 775 780

Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val
785 790 795 800

Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys
805 810 815

Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu
820 825 830

Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala
835 840 845

Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn
850 855 860

Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn
865 870 875 880

Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His
885 890 895

```
Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu
            900                 905                 910

Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu
        915                 920                 925

Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile
    930                 935                 940

Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly
945                 950                 955                 960

His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Val
                965                 970                 975

Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly
            980                 985                 990

Arg Gly Tyr Ile Leu Arg Val Thr  Ala Tyr Lys Glu Gly  Tyr Gly Glu
        995                 1000                1005

Gly Cys  Val Thr Ile His Glu  Ile Glu Asn Asn Thr  Asp Glu Leu
    1010                1015                1020

Lys Phe  Ser Asn Cys Val Glu  Glu Glu Val Tyr Pro  Asn Asn Thr
    1025                1030                1035

Val Thr  Cys Asn Asp Tyr Thr  Ala Thr Gln Glu Glu  Tyr Glu Gly
    1040                1045                1050

Thr Tyr  Thr Ser Arg Asn Arg  Gly Tyr Asp Gly Ala  Tyr Glu Ser
    1055                1060                1065

Asn Ser  Ser Val Pro Ala Asp  Tyr Ala Ser Ala Tyr  Glu Glu Lys
    1070                1075                1080

Ala Tyr  Thr Asp Gly Arg Arg  Asp Asn Pro Cys Glu  Ser Asn Arg
    1085                1090                1095

Gly Tyr  Gly Asp Tyr Thr Pro  Leu Pro Ala Gly Tyr  Val Thr Lys
    1100                1105                1110

Glu Leu  Glu Tyr Phe Pro Glu  Thr Asp Lys Val Trp  Ile Glu Ile
    1115                1120                1125

Gly Glu  Thr Glu Gly Thr Phe  Ile Val Asp Ser Val  Glu Leu Leu
    1130                1135                1140

Leu Met  Glu Glu
    1145
```

<210> 102
<211> 3447
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC859.

<400> 102

atggagaata atattcaaaa tcaatgcgta ccttacaatt gtttaaataa tcctgaagta 60

gaaatattaa atgaagaaag aagtactggc agattaccgt tagatatatc cttatcgctt 120

acacgtttcc ttttgagtga atttgttcca ggtgtgggag ttgcgtttgg attatttgat 180

ttaatatggg gttttataac tccttctgat tggagcttat ttcttttaca gattgaacaa 240

ttgattgagc aaagaataga aacattggaa aggaaccggg caattactac attacgaggg 300

ttagcagata gctatgaaat ttatattgaa gcactaagag agtgggaagc aaatcctaat 360

aatgcacaat taagggaaga tgtgcgtatt cgatttgcta atacagacga cgctttaata 420

acagcaataa ataattttac acttacaagt tttgaaatcc ctcttttatc ggtctatgtt 480

caagcggcga atttacattt atcactatta agagacgctg tatcgtttgg gcagggttgg 540

ggactggata tagctactgt taataatcat tataatagat taataaatct tattcataga 600

tatacgaaac attgtttgga cacatacaat caaggattag aaaacttaag aggtactaat 660

actcgacaat gggcaagatt caatcagttt aggagagatt taacgctaac tgtattagat 720

atcgttgctc tttttccgaa ctacgatgtt agaacatatc caattcaaac gtcatcccaa 780

ttaacaaggg aaatttatac aagttcagta attgaggatt ctccagtttc tgctaatata 840

cctaatggtt ttaatagggc ggaatttgga gttagaccgc cccatcttat ggactttatg 900

aattctttgt ttgtaactgc agagactgtt agaagtcaaa ctgtgtgggg aggacactta 960

gttagttcac gaaatacggc tggtaaccgt ataaatttcc ctagttacgg ggtcttcaat 1020

cctggtggcg ccatttggat tgcagatgag gatccacgtc ctttttatcg gacattatca 1080

gatcctgttt ttgtccgagg aggatttggg aatcctcatt atgtactggg gcttagggga 1140

gtagcatttc aacaaactgg tacgaaccac acccgaacat ttagaaatag tgggaccata 1200

gattctctag atgaaatccc acctcaggat aatagtgggg caccttggaa tgattatagt 1260

catgtattaa atcatgttac atttgtacga tggccaggtg agatttcagg aagtgattca 1320

tggagagctc caatgttttc ttggacgcac cgtagtgcaa cccctacaaa cgaagttcgt 1380

gtatctagaa ttacacaatt accgatggtg aaggcgcata cgcttcatgc gggagctact 1440

gttgttagag gaccaggatt tacaggagga gatatactcc gaagaactac ttcaggctca 1500

tttggggata tgagaataac aaatttttca agttcatcat cgaggtatcg tgtaagaata 1560

cgttatgctt ctactacaga tttacaattt ttcttgagtg ttggaggaac ccctgtcaat 1620

gtagccgatt tcccgaaaac catagataga ggggaaaact tagaatatgg aagctttaga 1680

acggcaggtt ttactacccc ttttagtttt gtaagttcaa caaataattt cacattaggt 1740

gttcagagtg tttcttcagg taacgagatt tttgtagatc gaattgaatt tgttccggca 1800

gatgcaacct ttgaggcaga atatgattta gaaagagcac aaaaggcggt gaatgagctg 1860

tttacttctt ccaatcaaat cgggttaaaa acagatgtga cggattatca tattgatcaa 1920

gtatccaatt tagttgagtg tttatctgat gaattttgtc tggatgaaaa aaaagaattg 1980

tccgagaaag tcaaacatgc gaagcgactt agtgatgagc ggaatttact tcaagatcca 2040

aactttagag ggatcaatag acaactagac cgtggctgga gaggaagtac ggatattacc 2100

atccaaggag gcgatgacgt attcaaagag aattacgtta cgctattggg tacctttgat 2160

gagtgctatc caacgtattt atatcaaaaa atagatgagt cgaaattaaa agcctatacc 2220

cgttaccaat taagagggta tatcgaagat agtcaagact tagaaatcta tttaattcgc 2280

tacaatgcca aacacgaaac agtaaatgtg ccaggtacgg gttccttatg gccgctttca 2340

gccccaagtc caatcggaaa atgtgcccat cattcccatc atttctcctt ggacattgat 2400

gttggatgta cagacttaaa tgaggactta ggtgtatggg tgatattcaa gattaagacg 2460

caagatggcc atgcaagact aggaaatcta gaatttctcg aagagaaacc attagtagga 2520

gaagcactag ctcgtgtgaa aagagcggag aaaaaatgga gagacaaacg tgaaaaattg 2580

gaatgggaaa caaatattgt ttataaagag gcaaaagaat ctgtagatgc tttatttgta 2640

aactctcaat atgatagatt acaagcggat accaacatcg cgatgattca tgcggcagat 2700

aaacgcgttc atagcattcg agaagcttat ctgcctgagc tgtctgtgat tccgggtgtc 2760

aatgcggcta tttttgaaga attagaaggg cgtattttca ctgcattctc cctatatgat 2820

gcgagaaatg tcattaaaaa tggtgatttt aataatggct tatcctgctg gaacgtgaaa 2880

gggcatgtag atgtagaaga acaaaacaac caccgttcgg tccttgttgt tccggaatgg 2940

gaagcagaag tgtcacaaga agttcgtgtc tgtccgggtc gtggctatat ccttcgtgtc 3000

acagcgtaca aggagggata tggagaaggt tgcgtaacca ttcatgagat cgagaacaat 3060

acagacgaac tgaagtttag caactgtgta gaagaggaag tatatccaaa caacacggta 3120

acgtgtaatg attatactgc gactcaagaa gaatatgagg gtacgtacac ttctcgtaat 3180

cgaggatatg acggagccta tgaaagcaat tcttctgtac cagctgatta tgcatcagcc 3240

tatgaagaaa aagcatatac agatggacga agagacaatc cttgtgaatc taacagagga 3300

tatggggatt acacaccact accagctggc tatgtgacaa aagaattaga gtacttccca 3360

gaaaccgata aggtatggat tgagatcgga gaaacggaag gaacattcat cgtggacagc 3420

gtggaattac ttcttatgga ggaatag 3447

<210> 103
<211> 1148
<212> PRT
<213> Nhân tạo

<220>

<223> Trình tự axit amin của protein khảm TIC859.

<400> 103

```
Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1               5                  10                  15

Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
            20                  25                  30

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
        35                  40                  45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
    50                  55                  60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65                  70                  75                  80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
                85                  90                  95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
            100                 105                 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
        115                 120                 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
    130                 135                 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
145                 150                 155                 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
                165                 170                 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
            180                 185                 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
        195                 200                 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
    210                 215                 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
225                 230                 235                 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
                245                 250                 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
            260                 265                 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
        275                 280                 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
    290                 295                 300
```

```
Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
305                 310                 315                 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
                325                 330                 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
            340                 345                 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
        355                 360                 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
    370                 375                 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
385                 390                 395                 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
                405                 410                 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
            420                 425                 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
        435                 440                 445

Thr His Arg Ser Ala Thr Pro Thr Asn Glu Val Arg Val Ser Arg Ile
    450                 455                 460

Thr Gln Leu Pro Met Val Lys Ala His Thr Leu His Ala Gly Ala Thr
465                 470                 475                 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
                485                 490                 495

Thr Ser Gly Ser Phe Gly Asp Met Arg Ile Thr Asn Phe Ser Ser Ser
            500                 505                 510

Ser Ser Arg Tyr Arg Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu
        515                 520                 525

Gln Phe Phe Leu Ser Val Gly Gly Thr Pro Val Asn Val Ala Asp Phe
    530                 535                 540

Pro Lys Thr Ile Asp Arg Gly Glu Asn Leu Glu Tyr Gly Ser Phe Arg
545                 550                 555                 560

Thr Ala Gly Phe Thr Thr Pro Phe Ser Phe Val Ser Ser Thr Asn Asn
                565                 570                 575

Phe Thr Leu Gly Val Gln Ser Val Ser Ser Gly Asn Glu Ile Phe Val
            580                 585                 590

Asp Arg Ile Glu Phe Val Pro Ala Asp Ala Thr Phe Glu Ala Glu Tyr
        595                 600                 605

Asp Leu Glu Arg Ala Gln Lys Ala Val Asn Glu Leu Phe Thr Ser Ser
    610                 615                 620

Asn Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln
```

625 630 635 640

Val Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu
645 650 655

Lys Lys Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp
660 665 670

Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln
675 680 685

Leu Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly
690 695 700

Asp Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp
705 710 715 720

Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu
725 730 735

Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln
740 745 750

Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val
755 760 765

Asn Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro
770 775 780

Ile Gly Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
785 790 795 800

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
805 810 815

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
820 825 830

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
835 840 845

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
850 855 860

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
865 870 875 880

Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
885 890 895

His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
900 905 910

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
915 920 925

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
930 935 940

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys
945 950 955 960

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val
965 970 975

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro
980 985 990

Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly
995 1000 1005

Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr Asp Glu
1010 1015 1020

Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro Asn Asn
1025 1030 1035

Thr Val Thr Cys Asn Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu
1040 1045 1050

Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu
1055 1060 1065

Ser Asn Ser Ser Val Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu
1070 1075 1080

Lys Ala Tyr Thr Asp Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn
1085 1090 1095

Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr
1100 1105 1110

Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu
1115 1120 1125

Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu
1130 1135 1140

Leu Leu Met Glu Glu
1145

<210> 104
<211> 3612
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC861.

<400> 104
atgaaactaa agaatccaga taagcatcaa agtttttcta gcaatgcgaa agtagataaa 60

atctctacgg attcactaaa aaatgaaaca gatatagaat tacaaaacat taatcatgaa 120

gattgtttga aaatatctga gtatgaaaat gtagagccgt ttgttagtgc atcaacaatt 180

caaacaggta ttagtattgc gggtaaaata cttggcaccc taggcgttcc ttttgcagga 240

caagtagcta gtctttatag ttttatctta ggtgagctat ggcctaaggg gaaaaatcaa 300

tgggaaatct ttatggaaca tgtagaagag attattaatc aaaaaatatc aacttatgca 360

agaaataaag cacttacaga cttgaaagga ttaggagatg ccttagctgt ctaccatgaa 420

tcgcttgaaa gttgggttgg aaatcgtaag aacacaaggg ctaggagtgt tgtcaagagc 480

caatatatcg cattagaatt gatgttcgtt cagaaactac cttcttttgc agtgtctgga 540

gaggaggtac cattattacc gatatatgcc caagctgcaa atttacattt gttgctatta 600

agagatgcat ctatttttgg aaaagagtgg ggattatcat cttcagaaat ttcaacattt 660

tataaccgtc aagtcgaacg agcaggagat tattccgacc attgtgtgaa atggtatagt 720

acaggtctaa ataacttgag gggtacaaat gccgaaagct gggttcgtta taatcaattt 780

cgtaaagata tgacattaat ggtacttgat ttagtcgcac tattcccaag ctatgataca 840

cttgtatatc caattaaaac tacttctcaa cttacaagag aagtatatac agacgcaatt 900

gggacagtac atccgaatgc aagttttgca agtacgactt ggtataataa taatgcccct 960

tcgttctcta ccatagagtc tgctgttgtt cgaaacccgc atctactcga ttttctagaa 1020

caagttacaa tttacagctt attaagtagg tggagtaaca ctcagtatat gaatatgtgg 1080

ggaggacata gacttgaatt ccgaacaatc ggaggaatgt taaatacctc aacacaagga 1140

tctactaata cttctattaa tcctgtaaca ttaccgttca cgtctcgaga cgtctatagg 1200

actgaatcat tggcagggct gaatctattt ttaactcaac ctgttaatgg agtacctagg 1260

gttgattttc attggaaatt cgtcacacat ccgatcgcat ctgataattt ctattatcca 1320

gggtatgctg gaattgggac gcaattacaa gattcagaaa atgaattacc acctgaaaca 1380

acaggacagc caaattatga atcatatagt catagattat ctcatatagg actcatttca 1440

gcatcccatg tgaaagcatt ggtatattct tggacgcatc gtagtgcaga tcgtacaaac 1500

acaattgatc cagagagaat taatcaaata cctttagtga aaggatttag agtttggggg 1560

ggcacctctg tcattacagg accaggattt acaggagggg atatccttcg aagaaatacc 1620

tttggtgatt ttgtatctct acaagtcaat attaattcac caattaccca aagataccgt 1680

ttaagatttc gttacgcttc cagtagggat gcacgagtta tagtattaac aggagcggca 1740

tccacaggag tgggaggcca agttagtgta aatatgcctc ttcagaaaac tatggaaata 1800

ggggagaact taacatctag aacatttaga tataccgatt ttagtaatcc tttttcattt 1860

agagctaatc cagatataat tgggataagt gaacaacctc tatttggtgc aggttctatt 1920

agtagcggtg aactttatat agataaaatt gaaattattc tagcagatgc aacatttgaa 1980

gcagaatctg atttagaaag agcacaaaag gcggtgaatg agctgtttac ttcttccaat 2040

caaatcgggt taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 2100

gagtgtttat ctgatgaatt ttgtctggat gaaaaaaaag aattgtccga gaaagtcaaa 2160

catgcgaagc gacttagtga tgagcggaat ttacttcaag atccaaactt tagagggatc 2220

aatagacaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggcgat 2280

gacgtattca aagagaatta cgttacgcta ttgggtacct ttgatgagtg ctatccaacg 2340

tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgtta ccaattaaga 2400

gggtatatcg aagatagtca agacttagaa atctatttaa ttcgctacaa tgccaaacac 2460

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagcccc aagtccaatc 2520

ggaaaatgtg cccatcattc ccatcatttc tccttggaca ttgatgttgg atgtacagac 2580

ttaaatgagg acttaggtgt atgggtgata ttcaagatta agacgcaaga tggccatgca 2640

agactaggaa atctagaatt tctcgaagag aaaccattag taggagaagc actagctcgt 2700

gtgaaaagag cggagaaaaa atggagagac aaacgtgaaa aattggaatg ggaaacaaat 2760

attgtttata aagaggcaaa agaatctgta gatgctttat ttgtaaactc tcaatatgat 2820

agattacaag cggataccaa catcgcgatg attcatgcgg cagataaacg cgttcatagc 2880

attcgagaag cttatctgcc tgagctgtct gtgattccgg gtgtcaatgc ggctattttt 2940

gaagaattag aagggcgtat tttcactgca ttctccctat atgatgcgag aaatgtcatt 3000

aaaaatggtg attttaataa tggcttatcc tgctggaacg tgaaagggca tgtagatgta 3060

gaagaacaaa acaaccaccg ttcggtcctt gttgttccgg aatgggaagc agaagtgtca 3120

caagaagttc gtgtctgtcc gggtcgtggc tatatccttc gtgtcacagc gtacaaggag 3180

ggatatggag aaggttgcgt aaccattcat gagatcgaga acaatacaga cgaactgaag 3240

tttagcaact gtgtagaaga ggaagtatat ccaaacaaca cggtaacgtg taatgattat 3300

actgcgactc aagaagaata tgagggtacg tacacttctc gtaatcgagg atatgacgga 3360

gcctatgaaa gcaattcttc tgtaccagct gattatgcat cagcctatga agaaaaagca 3420

tatacagatg gacgaagaga caatccttgt gaatctaaca gaggatatgg ggattacaca 3480

ccactaccag ctggctatgt gacaaaagaa ttagagtact tcccagaaac cgataaggta 3540

tggattgaga tcggagaaac ggaaggaaca ttcatcgtgg acagcgtgga attacttctt 3600

atggaggaat ag 3612

<210> 105
<211> 1203
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC861.

<400> 105

Met Lys Leu Lys Asn Pro Asp Lys His Gln Ser Phe Ser Ser Asn Ala
1 5 10 15

```
Lys Val Asp Lys Ile Ser Thr Asp Ser Leu Lys Asn Glu Thr Asp Ile
            20                  25                  30

Glu Leu Gln Asn Ile Asn His Glu Asp Cys Leu Lys Ile Ser Glu Tyr
        35                  40                  45

Glu Asn Val Glu Pro Phe Val Ser Ala Ser Thr Ile Gln Thr Gly Ile
    50                  55                  60

Ser Ile Ala Gly Lys Ile Leu Gly Thr Leu Gly Val Pro Phe Ala Gly
65                  70                  75                  80

Gln Val Ala Ser Leu Tyr Ser Phe Ile Leu Gly Glu Leu Trp Pro Lys
                85                  90                  95

Gly Lys Asn Gln Trp Glu Ile Phe Met Glu His Val Glu Glu Ile Ile
            100                 105                 110

Asn Gln Lys Ile Ser Thr Tyr Ala Arg Asn Lys Ala Leu Thr Asp Leu
        115                 120                 125

Lys Gly Leu Gly Asp Ala Leu Ala Val Tyr His Glu Ser Leu Glu Ser
    130                 135                 140

Trp Val Gly Asn Arg Lys Asn Thr Arg Ala Arg Ser Val Val Lys Ser
145                 150                 155                 160

Gln Tyr Ile Ala Leu Glu Leu Met Phe Val Gln Lys Leu Pro Ser Phe
                165                 170                 175

Ala Val Ser Gly Glu Glu Val Pro Leu Leu Pro Ile Tyr Ala Gln Ala
            180                 185                 190

Ala Asn Leu His Leu Leu Leu Leu Arg Asp Ala Ser Ile Phe Gly Lys
        195                 200                 205

Glu Trp Gly Leu Ser Ser Ser Glu Ile Ser Thr Phe Tyr Asn Arg Gln
    210                 215                 220

Val Glu Arg Ala Gly Asp Tyr Ser Asp His Cys Val Lys Trp Tyr Ser
225                 230                 235                 240

Thr Gly Leu Asn Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Val Arg
                245                 250                 255

Tyr Asn Gln Phe Arg Lys Asp Met Thr Leu Met Val Leu Asp Leu Val
            260                 265                 270

Ala Leu Phe Pro Ser Tyr Asp Thr Leu Val Tyr Pro Ile Lys Thr Thr
        275                 280                 285

Ser Gln Leu Thr Arg Glu Val Tyr Thr Asp Ala Ile Gly Thr Val His
    290                 295                 300

Pro Asn Ala Ser Phe Ala Ser Thr Thr Trp Tyr Asn Asn Asn Ala Pro
305                 310                 315                 320

Ser Phe Ser Thr Ile Glu Ser Ala Val Val Arg Asn Pro His Leu Leu
                325                 330                 335
```

```
Asp Phe Leu Glu Gln Val Thr Ile Tyr Ser Leu Leu Ser Arg Trp Ser
            340                 345                 350

Asn Thr Gln Tyr Met Asn Met Trp Gly Gly His Arg Leu Glu Phe Arg
        355                 360                 365

Thr Ile Gly Gly Met Leu Asn Thr Ser Thr Gln Gly Ser Thr Asn Thr
    370                 375                 380

Ser Ile Asn Pro Val Thr Leu Pro Phe Thr Ser Arg Asp Val Tyr Arg
385                 390                 395                 400

Thr Glu Ser Leu Ala Gly Leu Asn Leu Phe Leu Thr Gln Pro Val Asn
                405                 410                 415

Gly Val Pro Arg Val Asp Phe His Trp Lys Phe Val Thr His Pro Ile
            420                 425                 430

Ala Ser Asp Asn Phe Tyr Tyr Pro Gly Tyr Ala Gly Ile Gly Thr Gln
        435                 440                 445

Leu Gln Asp Ser Glu Asn Glu Leu Pro Pro Glu Thr Thr Gly Gln Pro
    450                 455                 460

Asn Tyr Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ser
465                 470                 475                 480

Ala Ser His Val Lys Ala Leu Val Tyr Ser Trp Thr His Arg Ser Ala
                485                 490                 495

Asp Arg Thr Asn Thr Ile Asp Pro Glu Arg Ile Asn Gln Ile Pro Leu
            500                 505                 510

Val Lys Gly Phe Arg Val Trp Gly Gly Thr Ser Val Ile Thr Gly Pro
        515                 520                 525

Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Asn Thr Phe Gly Asp Phe
    530                 535                 540

Val Ser Leu Gln Val Asn Ile Asn Ser Pro Ile Thr Gln Arg Tyr Arg
545                 550                 555                 560

Leu Arg Phe Arg Tyr Ala Ser Ser Arg Asp Ala Arg Val Ile Val Leu
                565                 570                 575

Thr Gly Ala Ala Ser Thr Gly Val Gly Gly Gln Val Ser Val Asn Met
            580                 585                 590

Pro Leu Gln Lys Thr Met Glu Ile Gly Glu Asn Leu Thr Ser Arg Thr
        595                 600                 605

Phe Arg Tyr Thr Asp Phe Ser Asn Pro Phe Ser Phe Arg Ala Asn Pro
    610                 615                 620

Asp Ile Ile Gly Ile Ser Glu Gln Pro Leu Phe Gly Ala Gly Ser Ile
625                 630                 635                 640

Ser Ser Gly Glu Leu Tyr Ile Asp Lys Ile Glu Ile Ile Leu Ala Asp
                645                 650                 655

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
```

660 665 670

Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
675 680 685

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
690 695 700

Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys
705 710 715 720

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
725 730 735

Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr
740 745 750

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
755 760 765

Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
770 775 780

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
785 790 795 800

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
805 810 815

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
820 825 830

Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His
835 840 845

His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp
850 855 860

Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala
865 870 875 880

Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu
885 890 895

Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg
900 905 910

Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu
915 920 925

Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala
930 935 940

Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser
945 950 955 960

Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn
965 970 975

Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser
980 985 990

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly
995 1000 1005

Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln
1010 1015 1020

Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu
1025 1030 1035

Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu
1040 1045 1050

Arg Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr
1055 1060 1065

Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn
1070 1075 1080

Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn
1085 1090 1095

Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser
1100 1105 1110

Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val
1115 1120 1125

Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
1130 1135 1140

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp
1145 1150 1155

Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr
1160 1165 1170

Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu
1175 1180 1185

Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1190 1195 1200

<210> 106
<211> 3432
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC848.

<400> 106
atggagataa ataatcagaa gcaatgcata ccatataatt gcttaagtaa tcctgaggaa 60

gtacttttgg atggggagag gatattacct gatatcgatc cactcgaagt ttctttgtcg 120

cttttgcaat ttcttttgaa taactttgtt ccaggggggag gctttatttc aggattagtt 180

gataaaatat gggggggcttt gagaccatct gaatgggact tatttcttgc acagattgaa 240

cggttgattg atcaaagaat agaagcaaca gtaagagcaa aagcaatcac tgaattagaa 300

ggattaggga gaaattatca aatatacgct gaagcattta aagaatggga atcagatcct 360

gataacgaag cggctaaaag tagagtaatt gatcgctttc gtatacttga tggtctaatt 420

gaagcaaata tcccttcatt tcggataatt ggatttgaag tgccactttt atcggtttat 480

gttcaagcag ctaatctaca tctcgctcta ttgagagatt ctgttatttt tggagagaga 540

tggggattga cgacaaaaaa tgtcaatgat atctataata gacaaattag agaaattcat 600

gaatatagca atcattgcgt agatacgtat aacacagaac tagaacgtct agggtttaga 660

tctatagcgc agtggagaat atataatcag tttagaagag aactaacact aactgtatta 720

gatattgtcg ctcttttccc gaactatgac agtagactgt atccgatcca aacttttct 780

caattgacaa gagaaattgt tacatcccca gtaagcgaat tttattatgg tgttattaat 840

agtggtaata taattggtac tcttactgaa cagcagataa ggcgaccaca tcttatggac 900

ttctttaact ccatgatcat gtatacatca gataatagac gggaacatta ttggtcagga 960

cttgaaatga cggcttattt tacaggattt gcaggagctc aagtgtcatt ccctttagtc 1020

gggactagag gggagtcagc tccaccatta actgttagaa gtgttaatga tggaatttat 1080

agaatattat cggcaccgtt ttattcagcg ccttttctag gcaccattgt attgggaagt 1140

cgtggagaaa aatttgattt tgcgcttaat aatatttcac ctccgccatc tacaatatac 1200

agacatcctg gaacagtaga ttcactagtc agtataccgc cacaggataa tagcgtacca 1260

ccgcacaggg gatctagtca tcgattaagt catgttacaa tgcgcgcaag ttcccctata 1320

ttccattgga cgcatcgcag cgcaaccact acaaatacaa ttaatccaaa tgctattatc 1380

caaataccac tagtaaaatc tactaatctt ggctctggaa cttctgtcgt taaaggacca 1440

ggatttacag gaggagatat tcttcgaaga acttcacctg gccagatttc aaccttaaga 1500

gtaaatatta ctgcaccatt atcacaaaga tatcgggtaa gaattcgcta cgcttctacc 1560

acaaatttac aattccatac atcaattgac ggaagaccta ttaatcaggg gaatttttca 1620

gcaactatga gtagtgggag taatttacag tccggaagct ttaggactgt aggttttact 1680

actccgttta acttttcaaa tggatcaagt gtatttacgt taagtgctca tgtcttcaat 1740

tcaggcaatg aagtttatat agatcgaatt gaatttgttc cggcagaagt aacctttgag 1800

gcagaatatg atttagaaag agcacaaaag gcggtgaatg agctgtttac ttcttccaat 1860

caaatcgggt taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1920

gagtgtttat ctgatgaatt ttgtctggat gaaaaaaaag aattgtccga gaaagtcaaa 1980

catgcgaagc gacttagtga tgagcggaat ttacttcaag atccaaactt tagagggatc 2040

aatagacaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggcgat 2100

gacgtattca aagagaatta cgttacgcta ttgggtacct ttgatgagtg ctatccaacg 2160

tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgtta ccaattaaga 2220

gggtatatcg aagatagtca agacttagaa atctatttaa ttcgctacaa tgccaaacac 2280

gaaacagtaa atgtgccagg tacgggttcc ttatggccgc tttcagcccc aagtccaatc 2340

ggaaaatgtg cccatcattc ccatcatttc tccttggaca ttgatgttgg atgtacagac 2400

ttaaatgagg acttaggtgt atgggtgata ttcaagatta agacgcaaga tggccatgca 2460

agactaggaa atctagaatt tctcgaagag aaaccattag taggagaagc actagctcgt 2520

gtgaaaagag cggagaaaaa atggagagac aaacgtgaaa aattggaatg ggaaacaaat 2580

attgtttata aagaggcaaa agaatctgta gatgctttat ttgtaaactc tcaatatgat 2640

agattacaag cggataccaa catcgcgatg attcatgcgg cagataaacg cgttcatagc 2700

attcgagaag cttatctgcc tgagctgtct gtgattccgg gtgtcaatgc ggctattttt 2760

gaagaattag aagggcgtat tttcactgca ttctccctat atgatgcgag aaatgtcatt 2820

aaaaatggtg attttaataa tggcttatcc tgctggaacg tgaaagggca tgtagatgta 2880

gaagaacaaa acaaccaccg ttcggtcctt gttgttccgg aatgggaagc agaagtgtca 2940

caagaagttc gtgtctgtcc gggtcgtggc tatatccttc gtgtcacagc gtacaaggag 3000

ggatatggag aaggttgcgt aaccattcat gagatcgaga acaatacaga cgaactgaag 3060

tttagcaact gtgtagaaga ggaagtatat ccaaacaaca cggtaacgtg taatgattat 3120

actgcgactc aagaagaata tgagggtacg tacacttctc gtaatcgagg atatgacgga 3180

gcctatgaaa gcaattcttc tgtaccagct gattatgcat cagcctatga agaaaaagca 3240

tatacagatg gacgaagaga caatccttgt gaatctaaca gaggatatgg ggattacaca 3300

ccactaccag ctggctatgt gacaaaagaa ttagagtact tcccagaaac cgataaggta 3360

tggattgaga tcggagaaac ggaaggaaca ttcatcgtgg acagcgtgga attacttctt 3420

atggaggaat ag 3432

<210> 107
<211> 1143
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC848.

<400> 107

Met Glu Ile Asn Asn Gln Lys Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Leu Pro Asp Ile
20 25 30

```
Asp Pro Leu Glu Val Ser Leu Ser Leu Leu Gln Phe Leu Leu Asn Asn
        35                  40                  45

Phe Val Pro Gly Gly Gly Phe Ile Ser Gly Leu Val Asp Lys Ile Trp
    50                  55                  60

Gly Ala Leu Arg Pro Ser Glu Trp Asp Leu Phe Leu Ala Gln Ile Glu
65                  70                  75                  80

Arg Leu Ile Asp Gln Arg Ile Glu Ala Thr Val Arg Ala Lys Ala Ile
                85                  90                  95

Thr Glu Leu Glu Gly Leu Gly Arg Asn Tyr Gln Ile Tyr Ala Glu Ala
            100                 105                 110

Phe Lys Glu Trp Glu Ser Asp Pro Asp Asn Glu Ala Ala Lys Ser Arg
        115                 120                 125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Ile Glu Ala Asn Ile
    130                 135                 140

Pro Ser Phe Arg Ile Ile Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145                 150                 155                 160

Val Gln Ala Ala Asn Leu His Leu Ala Leu Leu Arg Asp Ser Val Ile
                165                 170                 175

Phe Gly Glu Arg Trp Gly Leu Thr Thr Lys Asn Val Asn Asp Ile Tyr
            180                 185                 190

Asn Arg Gln Ile Arg Glu Ile His Glu Tyr Ser Asn His Cys Val Asp
        195                 200                 205

Thr Tyr Asn Thr Glu Leu Glu Arg Leu Gly Phe Arg Ser Ile Ala Gln
    210                 215                 220

Trp Arg Ile Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Leu Tyr Pro Ile
                245                 250                 255

Gln Thr Phe Ser Gln Leu Thr Arg Glu Ile Val Thr Ser Pro Val Ser
            260                 265                 270

Glu Phe Tyr Tyr Gly Val Ile Asn Ser Gly Asn Ile Ile Gly Thr Leu
        275                 280                 285

Thr Glu Gln Gln Ile Arg Arg Pro His Leu Met Asp Phe Phe Asn Ser
    290                 295                 300

Met Ile Met Tyr Thr Ser Asp Asn Arg Arg Glu His Tyr Trp Ser Gly
305                 310                 315                 320

Leu Glu Met Thr Ala Tyr Phe Thr Gly Phe Ala Gly Ala Gln Val Ser
                325                 330                 335

Phe Pro Leu Val Gly Thr Arg Gly Glu Ser Ala Pro Pro Leu Thr Val
            340                 345                 350
```

```
Arg Ser Val Asn Asp Gly Ile Tyr Arg Ile Leu Ser Ala Pro Phe Tyr
        355                 360                 365

Ser Ala Pro Phe Leu Gly Thr Ile Val Leu Gly Ser Arg Gly Glu Lys
    370                 375                 380

Phe Asp Phe Ala Leu Asn Asn Ile Ser Pro Pro Pro Ser Thr Ile Tyr
385                 390                 395                 400

Arg His Pro Gly Thr Val Asp Ser Leu Val Ser Ile Pro Pro Gln Asp
                405                 410                 415

Asn Ser Val Pro Pro His Arg Gly Ser Ser His Arg Leu Ser His Val
            420                 425                 430

Thr Met Arg Ala Ser Ser Pro Ile Phe His Trp Thr His Arg Ser Ala
        435                 440                 445

Thr Thr Thr Asn Thr Ile Asn Pro Asn Ala Ile Ile Gln Ile Pro Leu
    450                 455                 460

Val Lys Ser Thr Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro
465                 470                 475                 480

Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile
                485                 490                 495

Ser Thr Leu Arg Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg
            500                 505                 510

Val Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser
        515                 520                 525

Ile Asp Gly Arg Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser
    530                 535                 540

Ser Gly Ser Asn Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr
545                 550                 555                 560

Thr Pro Phe Asn Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala
                565                 570                 575

His Val Phe Asn Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe
            580                 585                 590

Val Pro Ala Glu Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala
        595                 600                 605

Gln Lys Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu
    610                 615                 620

Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val
625                 630                 635                 640

Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser
                645                 650                 655

Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu
            660                 665                 670

Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp
```

```
        675                 680                 685

Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys
    690                 695                 700

Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr
705                 710                 715                 720

Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg
                725                 730                 735

Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr
            740                 745                 750

Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr
        755                 760                 765

Gly Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala
    770                 775                 780

His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp
785                 790                 795                 800

Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln
                805                 810                 815

Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro
            820                 825                 830

Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp
        835                 840                 845

Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys
    850                 855                 860

Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp
865                 870                 875                 880

Arg Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys
                885                 890                 895

Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile
            900                 905                 910

Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe
        915                 920                 925

Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp
    930                 935                 940

Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val
945                 950                 955                 960

Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu
                965                 970                 975

Ala Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile
            980                 985                 990

Leu Arg Val Thr Ala Tyr Lys Glu  Gly Tyr Gly Glu Gly  Cys Val Thr
        995                 1000                1005
```

Ile His Glu Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn
1010 1015 1020

Cys Val Glu Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn
1025 1030 1035

Asp Tyr Thr Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser
1040 1045 1050

Arg Asn Arg Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val
1055 1060 1065

Pro Ala Asp Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp
1070 1075 1080

Gly Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp
1085 1090 1095

Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr
1100 1105 1110

Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu
1115 1120 1125

Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1130 1135 1140

<210> 108
<211> 3429
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC849.

<400> 108
atggagataa ataatcagaa gcaatgcata ccatataatt gcttaagtaa tcctgaggaa 60

gtacttttgg atggggagag gatattacct gatatcgatc cactcgaagt ttctttgtcg 120

cttttgcaat ttcttttgaa taactttgtt ccaggggggag gctttatttc aggattagtt 180

gataaaatat ggggggcttt gagaccatct gaatgggact tatttcttgc acagattgaa 240

cggttgattg atcaaagaat agaagcaaca gtaagagcaa aagcaatcac tgaattagaa 300

ggattaggga gaaattatca aatatacgct gaagcattta aagaatggga atcagatcct 360

gataacgaag cggctaaaag tagagtaatt gatcgctttc gtatacttga tggtctaatt 420

gaagcaaata tcccttcatt tcggataatt ggatttgaag tgccactttt atcggtttat 480

gttcaagcag ctaatctaca tctcgctcta ttgagagatt ctgttatttt tggagagaga 540

tggggattga cgacaaaaaa tgtcaatgat atctataata gacaaattag agaaattcat 600

gaatatagca atcattgcgt agatacgtat aacacagaac tagaacgtct agggtttaga 660

tctatagcgc agtggagaat atataatcag tttagaagag aactaacact aactgtatta 720

gatattgtcg ctcttttccc gaactatgac agtagactgt atccgatcca aacttttttct 780

caattgacaa gagaaattgt tacatcccca gtaagcgaat tttattatgg tgttattaat 840

agtggtaata taattggtac tcttactgaa cagcagataa ggcgaccaca tcttatggac 900

ttctttaact ccatgatcat gtatacatca gataatagac gggaacatta ttggtcagga 960

cttgaaatga cggcttattt tacaggattt gcaggagctc aagtgtcatt ccctttagtc 1020

gggactagag gggagtcagc tccaccatta actgttagaa gtgttaatga tggaatttat 1080

agaatattat cggcaccgtt ttattcagcg ccttttctag gcaccattgt attgggaagt 1140

cgtggagaaa aatttgattt tgcgcttaat aatatttcac ctccgccatc tacaatatac 1200

agacatcctg gaacagtaga ttcactagtc agtataccgc cacaggataa tagcgtacca 1260

ccgcacaggg gatctagtca tcgattaagt catgttacaa tgcgcgcaag ttcccctata 1320

ttccattgga cgcatcgcag cgcaaccact acaaatacaa ttaatccaaa tgctattatc 1380

caaataccac tagtaaaagc atttaacctt cattcaggtg ccactgttgt tagaggacca 1440

gggtttacag gtggagatat ccttcgtaga acaaatacgg gtacatttgg agatatacga 1500

ttaaatatta atgtgccatt atcccaaaga tatcgcgtaa ggattcgtta tgcttctact 1560

acagatttac aatttttcac gagaattaat ggaaccactg ttaatattgg taatttctca 1620

agaactatga atagggggga taatttagaa tatagaagtt ttagaactgc aggatttagt 1680

actcctttta atttttaaa tgcccaaagc acattcacat tgggtgctca gagtttttca 1740

aatcaggaag tttatataga tagagtcgaa tttgttccag cagaggtaac atttgaggca 1800

gaatatgatt tagaaagagc acaaaaggcg gtgaatgagc tgtttacttc ttccaatcaa 1860

atcgggttaa aaacagatgt gacggattat catattgatc aagtatccaa tttagttgag 1920

tgtttatctg atgaattttg tctggatgaa aaaaaagaat tgtccgagaa agtcaaacat 1980

gcgaagcgac ttagtgatga gcggaattta cttcaagatc caaactttag agggatcaat 2040

agacaactag accgtggctg gagaggaagt acggatatta ccatccaagg aggcgatgac 2100

gtattcaaag agaattacgt tacgctattg ggtacctttg atgagtgcta tccaacgtat 2160

ttatatcaaa aaatagatga gtcgaaatta aaagcctata cccgttacca attaagaggg 2220

tatatcgaag atagtcaaga cttagaaatc tatttaattc gctacaatgc caaacacgaa 2280

acagtaaatg tgccaggtac gggttcctta tggccgcttt cagccccaag tccaatcgga 2340

aaatgtgccc atcattccca tcatttctcc ttggacattg atgttggatg tacagactta 2400

aatgaggact taggtgtatg ggtgatattc aagattaaga cgcaagatgg ccatgcaaga 2460

ctaggaaatc tagaatttct cgaagagaaa ccattagtag gagaagcact agctcgtgtg 2520

aaaagagcgg agaaaaaatg gagagacaaa cgtgaaaaat tggaatggga aacaaatatt 2580

gtttataaag aggcaaaaga atctgtagat gctttatttg taaactctca atatgataga 2640

ttacaagcgg ataccaacat cgcgatgatt catgcggcag ataaacgcgt tcatagcatt 2700

cgagaagctt atctgcctga gctgtctgtg attccgggtg tcaatgcggc tatttttgaa 2760

gaattagaag ggcgtatttt cactgcattc tccctatatg atgcgagaaa tgtcattaaa 2820

aatggtgatt ttaataatgg cttatcctgc tggaacgtga aagggcatgt agatgtagaa 2880

gaacaaaaca accaccgttc ggtccttgtt gttccggaat gggaagcaga agtgtcacaa 2940

gaagttcgtg tctgtccggg tcgtggctat atccttcgtg tcacagcgta caaggaggga 3000

tatggagaag gttgcgtaac cattcatgag atcgagaaca atacagacga actgaagttt 3060

agcaactgtg tagaagagga agtatatcca aacaacacgg taacgtgtaa tgattatact 3120

gcgactcaag aagaatatga gggtacgtac acttctcgta atcgaggata tgacggagcc 3180

tatgaaagca attcttctgt accagctgat tatgcatcag cctatgaaga aaaagcatat 3240

acagatggac gaagagacaa tccttgtgaa tctaacagag gatatgggga ttacacacca 3300

ctaccagctg gctatgtgac aaaagaatta gagtacttcc cagaaaccga taaggtatgg 3360

attgagatcg gagaaacgga aggaacattc atcgtggaca gcgtggaatt acttcttatg 3420

gaggaatag 3429

<210> 109
<211> 1142
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC849.

<400> 109

Met Glu Ile Asn Asn Gln Lys Gln Cys Ile Pro Tyr Asn Cys Leu Ser
1 5 10 15

Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Leu Pro Asp Ile
20 25 30

Asp Pro Leu Glu Val Ser Leu Ser Leu Leu Gln Phe Leu Leu Asn Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Ile Ser Gly Leu Val Asp Lys Ile Trp
50 55 60

Gly Ala Leu Arg Pro Ser Glu Trp Asp Leu Phe Leu Ala Gln Ile Glu
65 70 75 80

Arg Leu Ile Asp Gln Arg Ile Glu Ala Thr Val Arg Ala Lys Ala Ile
85 90 95

Thr Glu Leu Glu Gly Leu Gly Arg Asn Tyr Gln Ile Tyr Ala Glu Ala
100 105 110

```
Phe Lys Glu Trp Glu Ser Asp Pro Asp Asn Glu Ala Ala Lys Ser Arg
        115                 120                 125

Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Ile Glu Ala Asn Ile
    130                 135                 140

Pro Ser Phe Arg Ile Ile Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145                 150                 155                 160

Val Gln Ala Ala Asn Leu His Leu Ala Leu Leu Arg Asp Ser Val Ile
                165                 170                 175

Phe Gly Glu Arg Trp Gly Leu Thr Thr Lys Asn Val Asn Asp Ile Tyr
            180                 185                 190

Asn Arg Gln Ile Arg Glu Ile His Glu Tyr Ser Asn His Cys Val Asp
        195                 200                 205

Thr Tyr Asn Thr Glu Leu Glu Arg Leu Gly Phe Arg Ser Ile Ala Gln
    210                 215                 220

Trp Arg Ile Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu
225                 230                 235                 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Leu Tyr Pro Ile
                245                 250                 255

Gln Thr Phe Ser Gln Leu Thr Arg Glu Ile Val Thr Ser Pro Val Ser
            260                 265                 270

Glu Phe Tyr Tyr Gly Val Ile Asn Ser Gly Asn Ile Ile Gly Thr Leu
        275                 280                 285

Thr Glu Gln Gln Ile Arg Arg Pro His Leu Met Asp Phe Phe Asn Ser
    290                 295                 300

Met Ile Met Tyr Thr Ser Asp Asn Arg Arg Glu His Tyr Trp Ser Gly
305                 310                 315                 320

Leu Glu Met Thr Ala Tyr Phe Thr Gly Phe Ala Gly Ala Gln Val Ser
                325                 330                 335

Phe Pro Leu Val Gly Thr Arg Gly Glu Ser Ala Pro Pro Leu Thr Val
            340                 345                 350

Arg Ser Val Asn Asp Gly Ile Tyr Arg Ile Leu Ser Ala Pro Phe Tyr
        355                 360                 365

Ser Ala Pro Phe Leu Gly Thr Ile Val Leu Gly Ser Arg Gly Glu Lys
    370                 375                 380

Phe Asp Phe Ala Leu Asn Asn Ile Ser Pro Pro Pro Ser Thr Ile Tyr
385                 390                 395                 400

Arg His Pro Gly Thr Val Asp Ser Leu Val Ser Ile Pro Pro Gln Asp
                405                 410                 415

Asn Ser Val Pro Pro His Arg Gly Ser Ser His Arg Leu Ser His Val
            420                 425                 430
```

```
Thr Met Arg Ala Ser Ser Pro Ile Phe His Trp Thr His Arg Ser Ala
        435                 440                 445

Thr Thr Thr Asn Thr Ile Asn Pro Asn Ala Ile Ile Gln Ile Pro Leu
    450                 455                 460

Val Lys Ala Phe Asn Leu His Ser Gly Ala Thr Val Val Arg Gly Pro
465                 470                 475                 480

Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Asn Thr Gly Thr Phe
                485                 490                 495

Gly Asp Ile Arg Leu Asn Ile Asn Val Pro Leu Ser Gln Arg Tyr Arg
            500                 505                 510

Val Arg Ile Arg Tyr Ala Ser Thr Thr Asp Leu Gln Phe Phe Thr Arg
        515                 520                 525

Ile Asn Gly Thr Thr Val Asn Ile Gly Asn Phe Ser Arg Thr Met Asn
    530                 535                 540

Arg Gly Asp Asn Leu Glu Tyr Arg Ser Phe Arg Thr Ala Gly Phe Ser
545                 550                 555                 560

Thr Pro Phe Asn Phe Leu Asn Ala Gln Ser Thr Phe Thr Leu Gly Ala
                565                 570                 575

Gln Ser Phe Ser Asn Gln Glu Val Tyr Ile Asp Arg Val Glu Phe Val
            580                 585                 590

Pro Ala Glu Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln
        595                 600                 605

Lys Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys
    610                 615                 620

Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu
625                 630                 635                 640

Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu
                645                 650                 655

Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln
            660                 665                 670

Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg
        675                 680                 685

Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu
    690                 695                 700

Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr
705                 710                 715                 720

Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr
                725                 730                 735

Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu
            740                 745                 750

Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly
```

```
        755                 760                 765

Ser Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His
    770                 775                 780

His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu
785                 790                 795                 800

Asn Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp
                805                 810                 815

Gly His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu
            820                 825                 830

Val Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg
        835                 840                 845

Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu
    850                 855                 860

Ala Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg
865                 870                 875                 880

Leu Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg
                885                 890                 895

Val His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro
            900                 905                 910

Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr
        915                 920                 925

Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe
    930                 935                 940

Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu
945                 950                 955                 960

Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala
                965                 970                 975

Glu Val Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu
            980                 985                 990

Arg Val Thr Ala Tyr Lys Glu Gly  Tyr Gly Glu Gly Cys  Val Thr Ile
        995                 1000                1005

His Glu  Ile Glu Asn Asn Thr  Asp Glu Leu Lys Phe  Ser Asn Cys
    1010                1015                1020

Val Glu  Glu Glu Val Tyr Pro  Asn Asn Thr Val Thr  Cys Asn Asp
    1025                1030                1035

Tyr Thr  Ala Thr Gln Glu Glu  Tyr Glu Gly Thr Tyr  Thr Ser Arg
    1040                1045                1050

Asn Arg  Gly Tyr Asp Gly Ala  Tyr Glu Ser Asn Ser  Ser Val Pro
    1055                1060                1065

Ala Asp  Tyr Ala Ser Ala Tyr  Glu Glu Lys Ala Tyr  Thr Asp Gly
    1070                1075                1080
```

Arg Arg Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr
1085 1090 1095

Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe
1100 1105 1110

Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly
1115 1120 1125

Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu
1130 1135 1140

<210> 110
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự mã hóa khảm ghi mã TIC847.

<400> 110
atgacttcaa ataggaaaaa tgagaatgaa attataaatg ctttatcgat tccaacggta 60

tcgaatcctt ccacgcaaat gaatctatca ccagatgctc gtattgaaga tagcttgtgt 120

gtagccgagg tgaacaatat tgatccattt gttagcgcat caacagtcca aacgggtata 180

aacatagctg gtagaatatt gggcgtatta ggtgtgccgt ttgctggaca actagctagt 240

ttttatagtt ttcttgttgg ggaattatgg cctagtggca gagatccatg ggaaattttc 300

ctggaacatg tagaacaact tataagacaa caagtaacag aaaatactag gaatacggct 360

attgctcgat tagaaggtct aggaagaggc tatagatctt accagcaggc tcttgaaact 420

tggttagata accgaaatga tgcaagatca agaagcatta ttcttgagcg ctatgttgct 480

ttagaacttg acattactac tgctataccg cttttcagaa tacgaaatga agaagttcca 540

ttattaatgg tatatgctca agctgcaaat ttacacctat tattattgag agacgcatcc 600

ctttttggta gtgaatgggg gatggcatct tccgatgtta accaatatta ccaagaacaa 660

atcagatata cagaggaata ttctaaccat tgcgtacaat ggtataatac agggctaaat 720

aacttaagag ggacaaatgc tgaaagttgg ttgcggtata atcaattccg tagagaccta 780

acgttagggg tattagattt agtagcccta ttcccaagct atgatactcg cacttatcca 840

atcaatacga gtgctcagtt aacaagagaa atttatacag atccaattgg gagaacaaat 900

gcaccttcag gatttgcaag tacgaattgg tttaataata atgcaccatc gttttctgcc 960

atagaggctg ccattttcag gcctccgcat ctacttgatt ttccagaaca acttacaatt 1020

tacagtgcat caagccgttg gagtagcact caacatatga attattgggt gggacatagg 1080

cttaacttcc gcccaatagg agggacatta aatacctcaa cacaaggact tactaataat 1140

acttcaatta atcctgtaac attacagttt acgtctcgag acgtttatag aacagaatca 1200

aatgcaggga caaatatact atttactact cctgtgaatg gagtaccttg ggctagattt 1260

aattttataa accctcagaa tatttatgaa agaggcgcca ctacctacag tcaaccgtat 1320

cagggagttg ggattcaatt atttgattca gaaactgaat taccaccaga aacaacagaa 1380

cgaccaaatt atgaatcata tagtcataga ttatctcata taggactaat cataggaaac 1440

actttgagag caccagtcta ttcttggacg catcgtagtg cagatcgtac gaatacgatt 1500

ggaccaaata gaattacaca aataccattg gtaaaggcgc ataccctcca atcgggtacc 1560

actgtagtaa aagggccagg gtttacagga ggggatatcc tccgtcgaac aagtggagga 1620

ccatttgctt ttagtaatgt taatctagat tttaacttgt cacaaaggta tcgtgctaga 1680

attcgttatg cctctactac taacctaaga atttacgtaa cggttgcagg tgaacgaatt 1740

tttgctggtc aatttgacaa aactatggat gctggtgccc cattaacatt ccaatctttt 1800

agttacgcaa ctattaatac agcttttaca ttcccagaaa gatcgagcag cttgactgta 1860

ggtgccgata cgtttagttc aggtaatgaa gtttatgtag atagatttga attaatccca 1920

gttactgcaa ccttcgaggc agaatctgat ttagaaagag cacaaaaggc ggtgaatgag 1980

ctgtttactt cttccaatca aatcgggtta aaaacagatg tgacggatta tcatattgat 2040

caagtatcca atttagttga gtgtttatct gatgaatttt gtctggatga aaaaaaagaa 2100

ttgtccgaga aagtcaaaca tgcgaagcga cttagtgatg agcggaattt acttcaagat 2160

ccaaacttta gagggatcaa tagacaacta gaccgtggct ggagaggaag tacggatatt 2220

accatccaag gaggcgatga cgtattcaaa gagaattacg ttacgctatt gggtaccttt 2280

gatgagtgct atccaacgta tttatatcaa aaaatagatg agtcgaaatt aaaagcctat 2340

acccgttacc aattaagagg gtatatcgaa gatagtcaag acttagaaat ctatttaatt 2400

cgctacaatg ccaaacacga aacagtaaat gtgccaggta cgggttcctt atggccgctt 2460

tcagccccaa gtccaatcgg aaaatgtgcc catcattccc atcatttctc cttggacatt 2520

gatgttggat gtacagactt aaatgaggac ttaggtgtat gggtgatatt caagattaag 2580

acgcaagatg gccatgcaag actaggaaat ctagaatttc tcgaagagaa accattagta 2640

ggagaagcac tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgtgaaaaa 2700

ttggaatggg aaacaaatat tgtttataaa gaggcaaaag aatctgtaga tgctttattt 2760

gtaaactctc aatatgatag attacaagcg gataccaaca tcgcgatgat tcatgcggca 2820

gataaacgcg ttcatagcat tcgagaagct tatctgcctg agctgtctgt gattccgggt 2880

gtcaatgcgg ctatttttga agaattagaa gggcgtattt tcactgcatt ctccctatat 2940

gatgcgagaa atgtcattaa aaatggtgat tttaataatg gcttatcctg ctggaacgtg 3000

aaagggcatg tagatgtaga agaacaaaac aaccaccgtt cggtccttgt tgttccggaa 3060

tgggaagcag aagtgtcaca agaagttcgt gtctgtccgg gtcgtggcta tatccttcgt 3120

gtcacagcgt acaaggaggg atatggagaa ggttgcgtaa ccattcatga gatcgagaac 3180

aatacagacg aactgaagtt tagcaactgt gtagaagagg aagtatatcc aaacaacacg 3240

gtaacgtgta atgattatac tgcgactcaa gaagaatatg agggtacgta cacttctcgt 3300

aatcgaggat atgacggagc ctatgaaagc aattcttctg taccagctga ttatgcatca 3360

gcctatgaag aaaaagcata tacagatgga cgaagagaca atccttgtga atctaacaga 3420

ggatatgggg attacacacc actaccagct ggctatgtga caaaagaatt agagtacttc 3480

ccagaaaccg ataaggtatg gattgagatc ggagaaacgg aaggaacatt catcgtggac 3540

agcgtggaat tacttcttat ggaggaatag 3570

<210> 111
<211> 1189
<212> PRT
<213> Nhân tạo

<220>
<223> Trình tự axit amin của protein khảm TIC847.

<400> 111

Met Thr Ser Asn Arg Lys Asn Glu Asn Glu Ile Ile Asn Ala Leu Ser
1 5 10 15

Ile Pro Thr Val Ser Asn Pro Ser Thr Gln Met Asn Leu Ser Pro Asp
20 25 30

Ala Arg Ile Glu Asp Ser Leu Cys Val Ala Glu Val Asn Asn Ile Asp
35 40 45

Pro Phe Val Ser Ala Ser Thr Val Gln Thr Gly Ile Asn Ile Ala Gly
50 55 60

Arg Ile Leu Gly Val Leu Gly Val Pro Phe Ala Gly Gln Leu Ala Ser
65 70 75 80

Phe Tyr Ser Phe Leu Val Gly Glu Leu Trp Pro Ser Gly Arg Asp Pro
85 90 95

Trp Glu Ile Phe Leu Glu His Val Glu Gln Leu Ile Arg Gln Gln Val
100 105 110

Thr Glu Asn Thr Arg Asn Thr Ala Ile Ala Arg Leu Glu Gly Leu Gly
115 120 125

Arg Gly Tyr Arg Ser Tyr Gln Gln Ala Leu Glu Thr Trp Leu Asp Asn
130 135 140

Arg Asn Asp Ala Arg Ser Arg Ser Ile Ile Leu Glu Arg Tyr Val Ala
145 150 155 160

Leu Glu Leu Asp Ile Thr Thr Ala Ile Pro Leu Phe Arg Ile Arg Asn

```
                 165                 170                 175

Glu Glu Val Pro Leu Leu Met Val Tyr Ala Gln Ala Ala Asn Leu His
            180                 185                 190

Leu Leu Leu Leu Arg Asp Ala Ser Leu Phe Gly Ser Glu Trp Gly Met
        195                 200                 205

Ala Ser Ser Asp Val Asn Gln Tyr Tyr Gln Glu Gln Ile Arg Tyr Thr
    210                 215                 220

Glu Glu Tyr Ser Asn His Cys Val Gln Trp Tyr Asn Thr Gly Leu Asn
225                 230                 235                 240

Asn Leu Arg Gly Thr Asn Ala Glu Ser Trp Leu Arg Tyr Asn Gln Phe
                245                 250                 255

Arg Arg Asp Leu Thr Leu Gly Val Leu Asp Leu Val Ala Leu Phe Pro
            260                 265                 270

Ser Tyr Asp Thr Arg Thr Tyr Pro Ile Asn Thr Ser Ala Gln Leu Thr
        275                 280                 285

Arg Glu Ile Tyr Thr Asp Pro Ile Gly Arg Thr Asn Ala Pro Ser Gly
    290                 295                 300

Phe Ala Ser Thr Asn Trp Phe Asn Asn Asn Ala Pro Ser Phe Ser Ala
305                 310                 315                 320

Ile Glu Ala Ala Ile Phe Arg Pro Pro His Leu Leu Asp Phe Pro Glu
                325                 330                 335

Gln Leu Thr Ile Tyr Ser Ala Ser Ser Arg Trp Ser Ser Thr Gln His
            340                 345                 350

Met Asn Tyr Trp Val Gly His Arg Leu Asn Phe Arg Pro Ile Gly Gly
        355                 360                 365

Thr Leu Asn Thr Ser Thr Gln Gly Leu Thr Asn Asn Thr Ser Ile Asn
    370                 375                 380

Pro Val Thr Leu Gln Phe Thr Ser Arg Asp Val Tyr Arg Thr Glu Ser
385                 390                 395                 400

Asn Ala Gly Thr Asn Ile Leu Phe Thr Thr Pro Val Asn Gly Val Pro
                405                 410                 415

Trp Ala Arg Phe Asn Phe Ile Asn Pro Gln Asn Ile Tyr Glu Arg Gly
            420                 425                 430

Ala Thr Thr Tyr Ser Gln Pro Tyr Gln Gly Val Gly Ile Gln Leu Phe
        435                 440                 445

Asp Ser Glu Thr Glu Leu Pro Pro Glu Thr Thr Glu Arg Pro Asn Tyr
    450                 455                 460

Glu Ser Tyr Ser His Arg Leu Ser His Ile Gly Leu Ile Ile Gly Asn
465                 470                 475                 480

Thr Leu Arg Ala Pro Val Tyr Ser Trp Thr His Arg Ser Ala Asp Arg
                485                 490                 495
```

```
Thr Asn Thr Ile Gly Pro Asn Arg Ile Thr Gln Ile Pro Leu Val Lys
            500                 505                 510

Ala His Thr Leu Gln Ser Gly Thr Thr Val Val Lys Gly Pro Gly Phe
        515                 520                 525

Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Phe
    530                 535                 540

Ser Asn Val Asn Leu Asp Phe Asn Leu Ser Gln Arg Tyr Arg Ala Arg
545                 550                 555                 560

Ile Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala
                565                 570                 575

Gly Glu Arg Ile Phe Ala Gly Gln Phe Asp Lys Thr Met Asp Ala Gly
            580                 585                 590

Ala Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala
        595                 600                 605

Phe Thr Phe Pro Glu Arg Ser Ser Ser Leu Thr Val Gly Ala Asp Thr
    610                 615                 620

Phe Ser Ser Gly Asn Glu Val Tyr Val Asp Arg Phe Glu Leu Ile Pro
625                 630                 635                 640

Val Thr Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys
                645                 650                 655

Ala Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr
            660                 665                 670

Asp Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys
        675                 680                 685

Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys
    690                 695                 700

Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp
705                 710                 715                 720

Pro Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly
                725                 730                 735

Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn
            740                 745                 750

Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu
        755                 760                 765

Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln
    770                 775                 780

Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile
785                 790                 795                 800

Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser
                805                 810                 815
```

```
Leu Trp Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His
            820                 825                 830

Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn
        835                 840                 845

Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly
    850                 855                 860

His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val
865                 870                 875                 880

Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp
                885                 890                 895

Lys Arg Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala
            900                 905                 910

Lys Glu Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu
        915                 920                 925

Gln Ala Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val
    930                 935                 940

His Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly
945                 950                 955                 960

Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala
                965                 970                 975

Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn
            980                 985                 990

Asn Gly Leu Ser Cys Trp Asn Val  Lys Gly His Val Asp  Val Glu Glu
        995                 1000                1005

Gln Asn  Asn His Arg Ser Val  Leu Val Val Pro Glu  Trp Glu Ala
    1010                1015                1020

Glu Val  Ser Gln Glu Val Arg  Val Cys Pro Gly Arg  Gly Tyr Ile
    1025                1030                1035

Leu Arg  Val Thr Ala Tyr Lys  Glu Gly Tyr Gly Glu  Gly Cys Val
    1040                1045                1050

Thr Ile  His Glu Ile Glu Asn  Asn Thr Asp Glu Leu  Lys Phe Ser
    1055                1060                1065

Asn Cys  Val Glu Glu Glu Val  Tyr Pro Asn Asn Thr  Val Thr Cys
    1070                1075                1080

Asn Asp  Tyr Thr Ala Thr Gln  Glu Glu Tyr Glu Gly  Thr Tyr Thr
    1085                1090                1095

Ser Arg  Asn Arg Gly Tyr Asp  Gly Ala Tyr Glu Ser  Asn Ser Ser
    1100                1105                1110

Val Pro  Ala Asp Tyr Ala Ser  Ala Tyr Glu Glu Lys  Ala Tyr Thr
    1115                1120                1125

Asp Gly  Arg Arg Asp Asn Pro  Cys Glu Ser Asn Arg  Gly Tyr Gly
```

1130 1135 1140

Asp Tyr Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu
1145 1150 1155

Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr
1160 1165 1170

Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu
1175 1180 1185

Glu

<210> 112
<211> 3588
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC713.

<400> 112
atggataaca acccaaacat caacgaatgc attccgtaca actgcctgag caaccctgag 60

gtggaggtgc taggtggcga gcgcatcgaa acgggttaca cgccgattga catctccctg 120

tccctcaccc agttcctcct tagcgagttc gtgccgggcg cgggtttcgt gctgggcctc 180

gttgacatca tctgggggat cttcggtccc tcccaatggg atgcgttcct ggtgcagatc 240

gagcagctta tcaaccagcg gatcgaagag ttcgctagga accaggccat cagccgcctc 300

gaaggtctca gcaacctcta tcaaatctac gctgagtcct tccgcgagtg ggaggcggac 360

cccaccaacc ctgccctccg cgaggaaatg cgtatccaat tcaatgacat gaacagcgcc 420

ctgaccacgg cgatcccgct gctggccgtc cagaactacc aggtcccccc gctcagcgtc 480

tacgtccagg ccgctaacct ccacctcagc gtgctgcgcg acgtctccgt cttcggccag 540

aggtgggggt tcgacgcggc gacgatcaac agccgctaca acgacctgac ccggctgatc 600

gggaactaca cggactacgc tgtccggtgg tacaacaccg gcctggagcg cgtgtgggga 660

ccagactccc gtgactgggt ccggtacaac cagttccgta gggaactcac tctgacggtc 720

ctggacatcg tggctctgtt cccgaactac gacagccgcc gctaccccat ccggactgtg 780

tcccaactga cccgcgaaat ctacacgaac cctgtccttg agaacttcga tgggagcttc 840

cgtggctccg cccagggcat cgagcgctcc atccgctccc cgcacctcat ggacatcctc 900

aactccatca cgatctacac cgacgcccat cgcggctact attactggtc cggccaccag 960

atcatggcca gccccgtggg cttctccggc cccgagttca ccttcccgct ctacggcacg 1020

atgggcaacg ccgctcctca gcaacgcatc gtggcccagc tcggccaggg cgtgtacagg 1080

accctctcca gcaccctcta caggcgtcct ttcaacatcg gcatcaacaa tcagcaactc 1140

tccgtgctcg acgggactga gttcgcctac ggcaccagca gcaacctgcc tagcgccgtg 1200

taccgcaaga gcggcaccgt ggacagcctg gacgagatcc ctccgcagaa caacaacgtg 1260

cctccgaggc aaggcttcag ccacaggctg agccacgtga gcatgttccg tagcggcttc 1320

agcaacagct ctgtgagcat catcagggcc ccgatgttct cctggattca caggagcgcc 1380

gagttcaaca ataccatcgg cccgaacagg atcacccaga tccctctggt gaaggccctg 1440

aacctgcact ctggcgtcac cgtcgttggc ggtcctggct tcaccggtgg ggacatcctg 1500

aggcgtacca acaccggaac cttcggtgac atcaggctga acatcaacgt ccctctgtct 1560

cagaggtaca gggtcaggat caggtacgcc tctaccactg accttcagtt ctttaccagg 1620

atcaacggta ctactgtcaa catcggtaac ttcagccgca ctatgaacag gggtgacaac 1680

ctggagtacc gtagtttccg tactgctgga ttctctactc ctttcaactt cctgaacgct 1740

cagtctactt tcactctggg tgctcagtct ttctcgaacc aggaggtcta catcgaccgt 1800

gtcgagttcg tcccggctga ggtcactttc gaggctgagt acgacctgga gcgcgctcaa 1860

aaggctgtca acgctctgtt cacttctact aacccgcgca gactgaagac tgacgtcact 1920

gactaccaca tcgaccaagt ctcgaacatg gtcgcttgcc tgtctgacga gttctgcctg 1980

gacgagaagc gtgagctgtt cgagaaggtc aagtacgcta agcgtctgtc tgacgagcgt 2040

aacctgctcc aggacccgaa cttcactttc atctcgggtc aactgtcgtt tgctagtatt 2100

gacggtcagt cgaactttcc gtcgattaac gagctgtcgg agcacggttg gtggggtagt 2160

gctaacgtca ctatccagga gggtaacgac gtctttaagg agaactacgt cactctgccg 2220

ggtactttca acgagtgcta cccgaactac ctttaccaga agattggtga gtcggagctt 2280

aaggcttaca ctcgttacca gcttcgtggt tacattgagg actcgcagga ccttgaaatc 2340

taccttattc ggtacaatgc caaacacgag actcttgacg tccctgggac ggatagtctt 2400

tggccgcttt cggtcgagtc gccgattggg cggtgcgggg agccgaatcg ttgcgctccg 2460

cactttgagt ggaatccaga ccttgattgc tcatgccggg atggggagcg gtgcgctcac 2520

catagtcacc actttacgct tgatattgat gtcgggtgca cggatcttca cgagaatctt 2580

ggggtctggg tcgtgtttaa gattaagacg caagaggggt acgcgcggct tgggaatctt 2640

gagtttattg aggaaaagcc acttattggt gaggcgcttt cacgggtcaa gcgggcggaa 2700

aagaaatgga gggataagcg ggagaagctc caattggaga cgaagcgggt ctacacggag 2760

gcgaaggagg cggtcgatgc gctgtttgtc gatagtcaat acgatcaatt gcaagcggat 2820

acgaacattg ggatgattca cgcggcagat aagttggtcc accggattag agaggcgtac 2880

ttgtcagagt tgccagtcat tccaggggtt aatgcggaga ttttcgagga attggagggg 2940

cacatcataa cggcgatgag tctgtacgat gcgagaaatg ttgtaaagaa tggtgatttc 3000

aacaatgggt tgacgtgctg gaatgttaag ggccacgttg atgttcaaca atcacatcat 3060

cggtcggatc tcgttattcc agagtgggag gcggaggttt cacaagcagt tagagtttgc 3120

ccagggagag gttacatcct aagagttacg gcatacaagg agggctatgg agaaggatgc 3180

gttacgatcc atgagattga gaacaatacg gatgagctaa agtttaagaa ttgtgaggaa 3240

gaagaagttt acccaacgga tacgggaaca tgcaacgatt acacagcaca tcaaggaaca 3300

gcagcatgta attcacgaaa tgccgggtat gaagatgctt acgaagttga tacaacagca 3360

tcagttaatt acaagccaac atacgaagaa gaaacataca cagatgttcg acgagataat 3420

cattgtgagt atgatcgagg gtatgtaaat tatccaccag ttccagccgg gtatgtcaca 3480

aaggaactag aatactttcc agaaacagat acagtatgga tagagatcgg agaaacagaa 3540

ggaaagttca tagtggactc agtagaacta ttactgatgg aagaatga 3588

<210> 113
<211> 3588
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC713.

<400> 113
atggataaca acccaaacat caacgaatgc attccgtaca actgcctgag caaccctgag 60

gtggaggtgc taggtggcga gcgcatcgaa acgggttaca cgccgattga catctccctg 120

tccctcaccc agttcctcct tagcgagttc gtgccgggcg cgggtttcgt gctgggcctc 180

gttgacatca tctgggggat cttcggtccc tcccaatggg atgcgttcct ggtgcagatc 240

gagcagctta tcaaccagcg gatcgaagag ttcgctagga accaggccat cagccgcctc 300

gaaggtctca gcaacctcta tcaaatctac gctgagtcct tccgcgagtg ggaggcggac 360

cccaccaacc ctgccctccg cgaggaaatg cgtatccaat tcaatgacat gaacagcgcc 420

ctgaccacgg cgatcccgct gctggccgtc cagaactacc aggtcccccc gctcagcgtc 480

tacgtccagg ccgctaacct ccacctcagc gtgctgcgcg acgtctccgt cttcggccag 540

aggtgggggt tcgacgcggc gacgatcaac agccgctaca acgacctgac ccggctgatc 600

gggaactaca cggactacgc tgtccggtgg tacaacaccg gcctggagcg cgtgtgggga 660

ccagactccc gtgactgggt ccggtacaac cagttccgta gggaactcac tctgacggtc 720

ctggacatcg tggctctgtt cccgaactac gacagccgcc gctaccccat ccggactgtg 780

tcccaactga cccgcgaaat ctacacgaac cctgtccttg agaacttcga tgggagcttc 840

cgtggctccg cccagggcat cgagcgctcc atccgctccc cgcacctcat ggacatcctc 900

aactccatca cgatctacac cgacgcccat cgcggctact attactggtc cggccaccag 960

atcatggcca gccccgtggg cttctccggc cccgagttca ccttcccgct ctacggcacg 1020

atgggcaacg ccgctcctca gcaacgcatc gtggcccagc tcggccaggg cgtgtacagg 1080

accctctcca gcaccctcta caggcgtcct ttcaacatcg gcatcaacaa tcagcaactc 1140

tccgtgctcg acgggactga gttcgcctac ggcaccagca gcaacctgcc tagcgccgtg 1200

taccgcaaga gcggcaccgt ggacagcctg gacgagatcc ctccgcagaa caacaacgtg 1260

cctccgaggc aaggcttcag ccacaggctg agccacgtga gcatgttccg tagcggcttc 1320

agcaacagct ctgtgagcat catcagggcc ccgatgttct cctggattca caggagcgcc 1380

gagttcaaca ataccatcgg cccgaacagg atcacccaga tccctctggt gaaggccctg 1440

aacctgcact ctggcgtcac cgtcgttggc ggtcctggct tcaccggtgg ggacatcctg 1500

aggcgtacca acaccggaac cttcggtgac atcaggctga acatcaacgt ccctctgtct 1560

cagaggtaca gggtcaggat caggtacgcc tctaccactg accttcagtt ctttaccagg 1620

atcaacggta ctactgtcaa catcggtaac ttcagccgca ctatgaacag gggtgacaac 1680

ctggagtacc gtagtttccg tactgctgga ttctctactc ctttcaactt cctgaacgct 1740

cagtctactt tcactctggg tgctcagtct ttctcgaacc aggaggtcta catcgaccgt 1800

gtcgagttcg tcccggctga ggtcactttc gaggctgagt acgacctgga gcgcgctcaa 1860

aaggctgtca acgctctgtt cacttctact aacccgcgca gactgaagac tgacgtcact 1920

gactaccaca tcgaccaagt ctcgaacatg gtcgcttgcc tgtctgacga gttctgcctg 1980

gacgagaagc gtgagctgtt cgagaaggtc aagtacgcta agcgtctgtc tgacgagcgt 2040

aacctgctcc aggacccgaa cttcactttc atctcgggtc aactgtcgtt tgctagtatt 2100

gacggtcagt cgaactttcc gtcgattaac gagctgtcgg agcacggttg gtggggtagt 2160

gctaacgtca ctatccagga gggtaacgac gtctttaagg agaactacgt cactctgccg 2220

ggtactttca acgagtgcta cccgaactac ctttaccaga agattggtga gtcggagctt 2280

aaggcttaca ctcgttacca gcttcgtggt tacattgagg actcgcagga ccttgaaatc 2340

taccttattc ggtacaatgc caaacacgag actcttgacg tccctgggac ggatagtctt 2400

tggccgcttt cggtcgagtc gccgattggg cggtgcgggg agccgaatcg ttgcgctccg 2460

cactttgagt ggaatccaga ccttgattgc tcatgccggg atggggagcg gtgcgctcac 2520

catagtcacc actttacgct tgatattgat gtcgggtgca cggatcttca cgagaatctt 2580

ggggtctggg tcgtgtttaa gattaagacg caagaggggt acgcgcggct tgggaatctt 2640

gagtttattg aggaaaagcc acttattggt gaggcgcttt cacgggtcaa gcgggcggaa 2700

aagaaatgga gggataagcg ggagaagctc caattggaga cgaagcgggt ctacacggag 2760

gcgaaggagg cggtcgatgc gctgtttgtc gatagtcaat acgatcaatt gcaagcggat 2820

acgaacattg ggatgattca cgcggcagat aagttggtcc accggattag agaggcgtac 2880

ttgtcagagt tgccagtcat tccaggggtt aatgcggaga ttttcgagga attggagggg 2940

cacatcataa cggcgatgag tctgtacgat gcgagaaatg ttgtaaagaa tggtgatttc 3000

aacaatgggt tgacgtgctg gaatgttaag ggccacgttg atgttcaaca atcacatcat 3060

cggtcggatc tcgttattcc agagtgggag gcggaggttt cacaagcagt tagagtttgc 3120

ccagggagag gttacatcct aagagttacg gcatacaagg agggctatgg agaaggatgc 3180

gttacgatcc atgagattga gaacaatacg gatgagctaa agtttaagaa ttgtgaggaa 3240

gaagaagttt acccaacgga tacgggaaca tgcaacgatt acacagcaca tcaaggaaca 3300

gcagcatgta attcacgaaa tgccgggtat gaagatgctt acgaagttga tacaacagca 3360

tcagttaatt acaagccaac atacgaagaa gaaacataca cagatgttcg acgagataat 3420

cattgtgagt atgatcgagg gtatgtaaat tatccaccag ttccagccgg gtatgtcaca 3480

aaggaactag aatactttcc agaaacagat acagtatgga tagagatcgg agaaacagaa 3540

ggaaagttca tagtggactc agtagaacta ttactgatgg aagaatag 3588

<210> 114
<211> 3531
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC843.

<400> 114
atggataaca acccaaacat caacgaatgc attccgtaca actgcctgag caaccctgag 60

gtggaggtgc taggtggcga gcgcatcgaa acgggttaca cgccgattga catctccctg 120

tccctcaccc agttcctcct tagcgagttc gtgccgggcg cgggtttcgt gctgggcctc 180

gttgacatca tctgggggat cttcggtccc tcccaatggg atgcgttcct ggtgcagatc 240

gagcagctta tcaaccagcg gatcgaagag ttcgctagga accaggccat cagccgcctc 300

gaaggtctca gcaacctcta tcaaatctac gctgagtcct tccgcgagtg ggaggcggac 360

cccaccaacc ctgccctccg cgaggaaatg cgtatccaat tcaatgacat gaacagcgcc 420

ctgaccacgg cgatcccgct gctggccgtc cagaactacc aggtcccccc gctcagcgtc 480

tacgtccagg ccgctaacct ccacctcagc gtgctgcgcg acgtctccgt cttcggccag 540

aggtgggggt tcgacgcggc gacgatcaac agccgctaca acgacctgac ccggctgatc 600

gggaactaca cggactacgc tgtccggtgg tacaacaccg gcctggagcg cgtgtgggga 660

ccagactccc gtgactgggt ccggtacaac cagttccgta gggaactcac tctgacggtc 720

ctggacatcg tggctctgtt cccgaactac gacagccgcc gctaccccat ccggactgtg 780

tcccaactga cccgcgaaat ctacacgaac cctgtccttg agaacttcga tgggagcttc 840

cgtggctccg cccagggcat cgagcgctcc atccgctccc cgcacctcat ggacatcctc 900

aactccatca cgatctacac cgacgcccat cgcggctact attactggtc cggccaccag 960

atcatggcca gccccgtggg cttctccggc cccgagttca ccttcccgct ctacggcacg 1020

atgggcaacg ccgctcctca gcaacgcatc gtggcccagc tcggccaggg cgtgtacagg 1080

accctctcca gcaccctcta caggcgtcct ttcaacatcg gcatcaacaa tcagcaactc 1140

tccgtgctcg acgggactga gttcgcctac ggcaccagca gcaacctgcc tagcgccgtg 1200

taccgcaaga gcggcaccgt ggacagcctg gacgagatcc ctccgcagaa caacaacgtg 1260

cctccgaggc aaggcttcag ccacaggctg agccacgtga gcatgttccg tagcggcttc 1320

agcaacagct ctgtgagcat catcagggcc ccgatgttct cctggattca caggagcgcc 1380

gagttcaaca ataccatcgg cccgaacagg atcacccaga tccctctggt gaaggccctg 1440

aacctgcact ctggcgtcac cgtcgttggc ggtcctggct tcaccggtgg ggacatcctg 1500

aggcgtacca acaccggaac cttcggtgac atcaggctga acatcaacgt ccctctgtct 1560

cagaggtaca gggtcaggat caggtacgcc tctaccactg accttcagtt ctttaccagg 1620

atcaacggta ctactgtcaa catcggtaac ttcagccgca ctatgaacag gggtgacaac 1680

ctggagtacc gtagtttccg tactgctgga ttctctactc ctttcaactt cctgaacgct 1740

cagtctactt tcactctggg tgctcagtct ttctcgaacc aggaggtcta catcgaccgt 1800

gtcgagttcg tcccggctga ggtcactttc gaggctgagt acgaccttga gagagcgcag 1860

aaggccgtga acgcgctgtt cacctcgact aaccagttgg gactcaagac aaatgtcact 1920

gactaccaca tcgaccaagt gagtaatctc gttacctacc tgtcggacga gttctgcctg 1980

gatgagaaga gggaactgtc tgagaaagtg aagcacgcaa agcgcctctc ggatgagagg 2040

aacctcctcc aggattccaa cttcaaggac atcaaccggc agccggaaag aggctggggt 2100

ggctccaccg gcattactat ccagggcggc gacgatgtgt tcaaggagaa ctacgtcaca 2160

ctctcaggca ctttcgacga gtgttacccg acgtacctgt accagaagat tgatgagtcc 2220

aagctgaagg cattcacccg ctatcagctt cgcggttaca ttgaagattc ccaggacctt 2280

gagatctacc tcatcaggta caacgccaag cacgagacgg tcaacgtgcc cggcacgggt 2340

tccctgtggc cgttgtctgc ccaatctccc atcgggaagt gcggtgagcc gaaccgttgc 2400

gccccgcacc tggagtggaa ccctgacctg gattgctcct gtagggatgg cgagaagtgc 2460

gctcaccatt cccaccactt ctcgctggac atcgacgtgg ggtgcacgga tctcaacgag 2520

gatctcggag tctgggtcat cttcaagatc aagacgcagg acggacacgc tcgcctcggc 2580

aatctggagt ttcttgagga aaagccgctc gtaggtgagg cgctcgccag ggtaaagcgc 2640

gcggagaaga agtggaggga taagcgtgag aaacttgaat gggagacaaa catcgtctac 2700

aaggaagcaa aggagtcagt tgatgcactg ttcgtcaact ctcagtatga ccagttgcag 2760

gccgatacca acatcgcgat gatccacgcc gctgacaagc gcgttcactc gatccgcgag 2820

gcgtacctcc ctgagctttc cgtcatccca ggcgtcaatg cggccatctt cgaggaactt 2880

gagggccgga tcttcaccgc tttctcgctc tacgacgcta gaaacgtcat caagaacggc 2940

gacttcaaca atggcctgag ctgctggaat gtgaagggtc atgtggatgt cgaagagcag 3000

aacaatcagc gttctgtgct cgtggtcccc gaatgggaag ctgaggtttc ccaagaagtg 3060

agagtgtgcc caggacgcgg ttacattctc cgtgttacgg cgtacaagga gggctatggc 3120

gaaggctgcg tgaccattca tgagatcgag aacaacacgg atgagctgaa gttctcaaac 3180

tgcgtcgagg aagaaatcta ccctaacaac actgtgacct gtaacgacta cacggtgaac 3240

caggaagagt acggcggcgc ttacactagc cgcaaccgtg ggtacaacga ggccccatcg 3300

gtccctgctg actacgcatc cgtctacgag gagaagtctt acactgacgg gagacgggag 3360

aacccctgcg agttcaaccg gggttacagg gattacacgc ctctgcccgt tggctatgtc 3420

actaaggagc tggagtattt cccggagact gataaagtct ggatcgagat cggggagact 3480

gaagggactt tcatcgtcga ttccgtggag ctgcttctga tggaggagtg a 3531

<210> 115
<211> 3501
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC862.

<400> 115
atggacaaca acccgaacat caacgagtgc atcccgtaca actgcctgag caaccctgag 60

gtggaggtgc taggcggtga gcgcatcgag accggctaca cgcctatcga catctccctg 120

agcctgaccc agttcctcct tagcgagttc gtgcccggcg cgggcttcgt gctcggcctc 180

gtggacatca tctggggcat cttcggcccg tcccagtggg acgccttcct cgtgcagatc 240

gagcagctca tcaaccagcg catcgaggag ttcgcccgca accaggccat ctcccgcctg 300

gagggcctct ccaacctcta ccagatctac gccgagtcct tccgcgagtg ggaggccgac 360

ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgacat gaactccgcc 420

ctcaccactg ccatcccgct cctggccgtg cagaactacc aagtgcctct cctgtccgtg 480

tacgtccaag ctgcgaacct ccacctctcc gtgctgcgcg acgtctccgt cttcggccag 540

cgctggggct tcgacgctgc gaccatcaac tcccgctaca acgacctgac ccgcctgatc 600

ggcaactaca ccgactacgc tgtccgctgg tacaacaccg gcctggagcg cgtctggggc 660

cctgactcca gggactgggt ccgctacaac cagttccgcc gggagctgac cctgactgtc 720

ctggacatcg tcgctctgtt ccctaactac gatagccgcc gctaccctat caggactgtc 780

agccagctta ctagggaaat ctacactaac cctgtccttg agaacttcga tggttccttc 840

cgtggtagcg ctcagggtat tgagcggagc attcggagcc cgcaccttat ggacattctt 900

aactctatta ctatctacac ggatgctcac cgtggttact attactggtc tggtcaccag 960

attatggcta gtccggtcgg tttctctggg ccggagttca cgttcccgct ttatgggacg 1020

atgggcaacg cggcaccgca gcaacgtatt gtggcgcagc ttgggcaagg cgtttatcgt 1080

acgctttctt cgacattgta cagacgtccg tttaacattg ggatcaataa ccagcaattg 1140

tctgttctgg atggaacaga atttgcgtat ggaacatctt cgaatttgcc atctgcggtt 1200

tatcgtaagt ctggaacagt tgacagtctc gatgagattc caccacagaa caacaacgtt 1260

ccaccaagac agggattcag tcacagatta agccacgtta gtatgtttag atcaggcttc 1320

tcgaactcat cagtatcaat catacgagca ccaatgttct cgtggataca ccgctcggca 1380

gagttcaaca acatcatcgc ctccgacagc atcacccaga tcccggccgt gaagggcaac 1440

ttcctcttca acggctccgt catctccggc cctggcttca ccggcggcga tctcgtgcgc 1500

tgcacgaatg gttccggcct caccctgtac gtgacgcctg cgcccgacct cacctactcc 1560

aagacctaca agattcgcat tcgctatgcc tccacctccc aagtgcgctt tggcatcgac 1620

ctcggctcct acactcactc catctcctac ttcgacaaga cgatggacaa aggcaacacc 1680

ctcacctaca actccttcaa cctctcctcc gtgtcccgcc ctatcgagat cagcggcggc 1740

aacaagatcg gtgtctcggt cggcggcatc ggctctggcg acgaggtgta catcgacaag 1800

atcgagttca tcccgatgga gcgcgcccag aaggccgtga acgccctctt cactagcact 1860

aaccagctcg gcctcaagac taacgtgacc gactaccaca ttgaccaagt gagcaaccta 1920

gtgacctacc ttagcgacga gttctgcctt gacgagaagc gtgagctgag cgagaaggtg 1980

aagcacgcca agcgcctctc cgacgagcgc aacctcctcc aggactccaa cttcaaggac 2040

atcaaccgcc agcccgagcg cggctggggc ggtagcaccg gcatcaccat ccagggcggt 2100

gacgatgtgt tcaaggagaa ctacgtgacc ctctccggca ccttcgacga gtgctacccg 2160

acctacctct accagaagat cgacgagtcc aagctcaagg cgttcacccg ctaccagctt 2220

cgcggctaca tcgaggactc ccaggatctg gagatctacc tcatccgcta caacgccaag 2280

cacgagaccg tgaacgtgcc cggcaccggc tccctctggc cgctctccgc ccagagccct 2340

atcggcaagt gcggcgagcc caaccgctgc gcgcctcacc tggagtggaa ccctgacctc 2400

gactgctcct gccgcgacgg cgagaagtgc gcccaccata gccaccactt ctctctcgac 2460

atcgacgtgg gctgcaccga cctcaacgag gatctgggcg tgtgggtgat cttcaagatc 2520

aagacccagg acggccacgc caggctgggc aacctggagt tcctggagga gaagcctctg 2580

gtgggtgagg ccctggccag ggtcaagagg gctgagaaga aatggaggga caagagggag 2640

aagctggagt gggagaccaa catcgtgtac aaggaggcta aggagtccgt ggacgctctg 2700

ttcgtcaact ctcagtacga tcagctccag gctgacacca acatcgctat gatccacgct 2760

gcggataaga gggtccactc tatcagggag gcttacctgc ctgagctttc tgtcatccct 2820

ggtgtcaacg cggcaatctt cgaggaactt gagggccgca tcttcactgc gttctcgctt 2880

tacgatgcgc ggaacgtcat taagaacggt gacttcaaca atggtctttc gtgctggaac 2940

gtcaagggtc atgtcgatgt cgaggaacag aacaaccagc ggtcggtcct tgtcgttccc 3000

gagtgggagg ccgaggtctc gcaagaggtc cgggtctgcc ctgggcgcgg gtacattctt 3060

cgtgtcactg cgtacaagga gggctacggc gagggctgcg ttactattca tgagattgag 3120

aacaatacgg atgagcttaa gtttagtaac tgtgttgagg aggagatcta cccgaacaat 3180

acggttacgt gcaatgatta cacggtgaac caggaggaat acggcggagc atacacctca 3240

cgtaatagag ggtacaatga ggcaccgtca gttccggcag attatgcctc agtttatgag 3300

gagaagtcct acacggatgg aagacgcgag aatccatgtg agtttaatag aggataccga 3360

gactacacac cactcccagt tggatacgtt acaaaggagt tggaatactt cccagaaaca 3420

gataaagttt ggatagagat cggagaaaca gaaggaacct tcatcgtgga cagtgtagaa 3480

ctgctgctga tggaagaatg a 3501

<210> 116
<211> 3549
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC1099.

<400> 116
atggacaaca accccaacat caatgagtgc atcccttaca actgcctcag caaccccgag 60

gtggaggtgc tcggcggtga gcggatcgag accgggtaca ctcctatcga catcagcctg 120

tccctgacac agttccttct ctccgagttc gtcccaggcg cgggctttgt cctgggcctt 180

gtggacatca tctggggaat cttcgggcca tctcagtggg acgcgtttct tgtgcagata 240

gagcaactta ttaaccagag gatcgaggag ttcgcgcgca atcaggctat ctcacgcctc 300

gaagggcttt ccaatctgta ccagatctac gcggagtcgt tccgggaatg ggaggcggac 360

cctaccaacc cggcgctgcg cgaggaaatg aggatccagt tcaacgatat gaactcggct 420

ctgaccacgg ccatccctct gctcgccgtg cagaattacc aagtcccgct cctttccgtg 480

tacgtccagg ctgccaacct ccatctgtcg gtcctccgtg acgtgagcgt gttcggtcag 540

cgttggggct tcgatgccgc gactatcaac agccggtaca acgacttgac taggcttatc 600

gggaactaca cagactacgc tgtcagatgg tacaataccg gcctggagag agtgtggggc 660

cccgactccc gcgactgggt gagatacaat cagttccgcc gcgagcttac cctcactgtc 720

ctggacatcg tggccctctt cccaaattac gatagccgcc gttacccgat caggactgtg 780

agtcagttga cccgtgaaat ctacaccaac ccggtcttgg agaactttga cggttccttc 840

cggggttccg cacaaggaat cgagcggagc atccgctctc cacacctgat ggacatcctg 900

aactcaatta ccatctacac ggatgcccac cgaggctact attactggag tggtcaccaa 960

atcatggcca gccctgtggg cttctcgggg ccggagttca cgttcccgct gtacggtacg 1020

atgggcaacg ctgcacctca gcaaaggatc gttgcccaac tgggtcaggg cgtttacaga 1080

accttgtcga gcacactgta ccggaggccg ttcaacatcg gcatcaataa ccagcagctg 1140

agcgttctcg acggcacgga gttcgcctac ggaactagca gcaacctccc gagtgcagtg 1200

taccggaagt ccgggaccgt ggatagcctg gacgaaatcc cacctcagaa caacaacgtc 1260

cctcctcggc aaggcttctc ccaccgcctt agccatgttt ccatgttccg gtcgggattc 1320

agtaactcca gcgtctctat catcagggct cccatgttca gctggattca taggtccgct 1380

gagttcaaca acatcattgc ctccgactca atcactcagc tccctctggt gaaggcgtct 1440

gcgccagtca gtggaaccac cgtactcaag ggccccggct tcaccggcgg cggaatcctt 1500

cgtaggacca ctaacggcac tttcggtact ctgcgtgtca ctgtcaatag cccgctcact 1560

cagcgctacc gtgtaagggt tcgattcgcc agctccggca acttctcaat aaggatcttg 1620

cgcggaaaca catctatcgc ataccagcgc ttcggctcaa cgatgaacag gggccaagag 1680

ctgacttacg agagcttcgt gacctcagag ttcaccacta accagtccga tctacctttc 1740

accttcaccc aggctcagga gaacctcaca attcttgccg aaggcgtgag cactggcagt 1800

gagtatttca tcgaccggat cgagatcata ccagtcaacc cagcacgtga ggctgaggaa 1860

gacctcgaac gcgcgcagaa ggcagtcaac gccctcttta ccagcaccaa tcagttgggt 1920

cttaagacga atgttaccga ttaccatatc gaccaagtta gtaatcttgt tacatacctc 1980

tctgatgagt tctgcttgga cgaaaagcgt gaactctccg agaaggtcaa acacgcgaag 2040

cgcctctctg atgagcgcaa ccttctgcaa gactcgaact tcaaggacat taacaggcaa 2100

ccggaaagag gttggggcgg tagtaccggc attacgatcc agggcggaga cgatgttttc 2160

aaggagaact acgttaccct tagcggcacg tttgacgagt gctaccctac ctatctctat 2220

cagaagatcg acgagagcaa gctcaaggcg tttacccgtt accagctccg tggctacatt 2280

gaggactctc aggatctcga aatctacttg atccggtaca acgcgaagca cgagaccgtc 2340

aacgtccctg ggacagggtc gctgtggccc ctctctgccc aatctccgat tggaaagtgc 2400

ggcgagccaa accggtgtgc tcctcacctg gagtggaatc ccgatctgga ttgctcgtgc 2460

agggacggtg agaagtgtgc acaccattca caccacttct ccctcgacat cgacgtgggc 2520

tgtacggatc tgaacgagga tcttggggtt tgggttatct tcaagattaa gactcaggat 2580

ggacacgcgc gcctcggcaa cctggagttc ctcgaggaaa agcctctcgt gggcgaggcg 2640

ctggcccgcg ttaagcgagc tgagaagaag tggcgggaca aacgcgagaa gctagagtgg 2700

gaaacgaaca tcgtttacaa ggaagccaag gaatctgtgg atgccctctt cgtgaactct 2760

cagtacgacc aactccaggc cgacaccaac atcgccatga tccatgctgc cgacaagcgc 2820

gtgcacagca tccgggaagc gtacctgcct gagctgagcg ttattccggg tgtgaacgct 2880

gcgatcttcg aggaactcga aggcagaatc ttcaccgcat tctccctgta cgatgctagg 2940

aatgttatca agaacgggga cttcaacaat gggctctcct gttggaacgt taagggccat 3000

gtggacgtgg aggaacagaa taaccagagg tctgtcctcg tggtccctga gtgggaggct 3060

gaggtttcac aggaggttcg ggtctgcccg ggtagggggt acattctccg ggttaccgcc 3120

tacaaggagg gctacggcga gggctgtgtg accatacacg agatcgagaa taacaccgat 3180

gagttgaagt ttagcaactg cgtcgaggag gaaatctacc ccaacaatac ggttacctgc 3240

aacgactaca ctgtcaacca ggaggagtat ggcggagctt acacatcccg gaaccgaggc 3300

tacaacgagg ccccatctgt cccggcggac tatgcgtcgg tgtacgagga gaagtcctac 3360

accgatggcc gtagggagaa tccgtgcgag tttaaccgtg gctaccgcga ctacacgcct 3420

ttgcccgtcg gttatgtcac taaggaactg gaatactttc ccgagacgga caaggtgtgg 3480

attgagatcg gtgagactga gggcacgttt atcgttgact cggtggagct gttgctcatg 3540

gaggagtga 3549

<210> 117
<211> 3534
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC1103.

<400> 117
atggacaaca acccgaacat caacgagtgc atcccgtaca actgcctgag caacccggag 60

gttgaggtgc tgggcggtga gcgcatcgag accggctaca cgcctatcga catctccctg 120

agcctgaccc agttcctcct tagcgagttc gtgccgggcg cgggcttcgt gctcggcctc 180

gtggacatca tctggggcat cttcggcccg tcccagtggg acgccttcct cgtgcagatc 240

gagcagctca tcaaccagcg catcgaggag ttcgcccgca accaggccat ctcccgcctg 300

gagggtctgt ccaacctcta ccaaatctac gccgagtcct tccgcgagtg ggaggccgac 360

ccgaccaacc cggccctccg cgaggagatg cggattcagt tcaacgacat gaactccgcc 420

ctcaccaccg ccatcccgct cctggccgtg cagaactacc aagtgccgct cctgtccgtg 480

tacgtgcaag ccgcgaacct ccacctctcc gtgctccgcg acgtctccgt gttcggccag 540

cgctggggct tcgacgccgc gaccatcaac tcccgctaca acgacctcac ccgcctcatc 600

ggcaactaca ccgactacgc cgtgcgctgg tacaacaccg gcctggagcg cgtgtggggc 660

ccggactccc gcgactgggt gcgctacaac cagttccgcc gcgagctgac cctcaccgtg 720

ctcgacatcg tggccctctt cccgaactac gactccaggc gctacccgat ccgcaccgtg 780

tcccagctca cccgcgaaat ctacaccaac ccggtgctgg agaacttcga cggctccttc 840

cggggctccg cccagggcat cgagcgctcc atccgctccc cgcacctcat ggacatcctc 900

aactccatca ccatctacac cgacgcccac aggggctact actactggtc cggccaccag 960

atcatggcca gcccggtggg cttctccggc ccggagttca ccttcccgct ctacggcact 1020

atgggcaacg ccgcgcccca gcaaaggatc gtggcccagc tcggccaggg cgtgtaccgc 1080

accctcagct ctaccctgta caggcggccc ttcaacatcg gcatcaacaa ccagcaactg 1140

agcgtgctgg acggcaccga gttcgcctac ggcaccagca gcaacctgcc cagcgccgtg 1200

taccgcaaga gcggcaccgt ggacagcctg gacgagatcc cgccgcagaa caacaacgtg 1260

ccgccgcgcc agggcttcag ccacaggctg agccacgtga gcatgttccg cagcggcttc 1320

agcaacagct ctgtgagcat catcagggct cccatgttct cctggattca caggagcgcg 1380

gagttcaaca acatcattgc gagcgactct atcacccagc tcccgatggt gaaggcgaac 1440

gcgctgctga gcggcaccag cgtcatcaag ggccctggca gcaccggcgg ggacatcctg 1500

aggcggacgt ctgtgggcag ggtcgggaac ttcaaggtca acgtcaacgg gccgctgacg 1560

cagcgctacc tggtccgcat ccgctacgcg agcacgactg acctggactt ctacgtttac 1620

cggggcggta cgactgtcag caactaccgg ttcaacaaga cgatgaacaa gggcgcgtct 1680

ctgacgtacg acatcttcaa gttcgcgtct ttcagcacgc ccttcacgtt cacgaagacg 1740

caagacgagc tggggatcag tatccagaac ttcagcagcg gcgaggaggt gtacattgac 1800

cggattgagg tcattcccgt cgggacgact tacgaggcgg agacggacct ggagcgggcg 1860

cagaaggcgg tcaacgcgct gttcacgagc acgaaccagc tcggcctcaa gacgaacgtc 1920

actgactacc acattgacca agtctcgaac ctggtcactt acctgtcgga cgagttctgc 1980

ctggacgaga agcgggagct gtcggagaag gtcaagcacg cgaagcggct gtcggacgag 2040

cggaacctgc tccaggactc gaacttcaag gacatcaacc ggcagccaga gcggggctgg 2100

ggcggcagca ccgggattac tatccagggc ggtgacgacg tctttaagga gaactacgtc 2160

actctttcgg gcactttcga cgagtgctac ccaacttacc tttaccagaa gattgacgag 2220

tcgaagctca aggcgtttac tcgctaccaa cttcgcgggt acattgagga cagtcaggat 2280

cttgaaatct accttatccg ctacaacgcg aagcacgaga ctgtcaacgt cccagggact 2340

gggtcgctct ggccactctc ggctcaatcg ccaattggga agtgcggtga gccaaacaga 2400

tgcgctccac accttgagtg gaatccagac cttgactgct cgtgcagaga cggtgagaag 2460

tgcgctcacc atagtcatca cttcagtctt gacatcgacg tcggttgcac tgatctgaat 2520

gaggatcttg gtgtctgggt catcttcaag attaagactc aagatggtca tgctagactc 2580

ggcaacttgg agttcctgga ggagaagcca ttggtcggtg aggctttggc tagagtcaag 2640

cgcgctgaga agaaatggag agataagagg gagaagctgg agtgggagac caacatcgtg 2700

tacaaggagg ctaaagagtc agtcgatgct ctgttcgtca atagccagta cgatcagctc 2760

caagctgaca caaacatcgc tatgattcat gctgccgata agagggtcca tagtatccgt 2820

gaggcttatc tgcctgagtt gtcagtcata cctggtgtga atgctgccat cttcgaggag 2880

ctggagggtc gcatcttcac agcgttcagt ctgtacgatg cacgtaacgt catcaagaac 2940

ggtgacttca acaacggtct gtcctgctgg aatgtgaaag gccacgtgga tgtggaggaa 3000

cagaacaacc agcgttcagt gctcgtcgta cctgagtggg aggcagaggt gtcacaggag 3060

gtgcgtgtgt gccctggacg tggctacatc ctccgtgtta cagcatacaa ggagggatac 3120

ggcgagggat gtgtgaccat ccacgagatc gagaacaaca cagatgaact aaagttcagc 3180

aattgtgttg aggaggaaat ctacccgaac aacacagtga catgcaacga ctacacagtg 3240

aaccaggaag aatacggagg agcatacaca tcacgtaacc gaggatacaa cgaggcacct 3300

tcagttccgg ccgactacgc ctcagtttac gaggagaagt catacacaga tggacgaagg 3360

gagaaccctt gtgagttcaa tcgaggctac cgagactaca cacctttgcc tgtgggctac 3420

gtgacaaagg agctggaata cttccctgaa acagacaaag tatggatcga gatcggagaa 3480

acagaaggaa cattcatcgt ggactccgtg gagctgctgc tgatggagga gtga 3534

<210> 118
<211> 3573
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC845.

<400> 118
atgaccagca accggaagaa cgagaacgag atcatcaacg ccctgagcat cccgaccgtg 60

agcaacccta gcacccagat gaacctgagc cctgacgctc gcatcgagga ctccctctgc 120

gtggctgagg tgaacaacat cgacccgttc gtgtccgcct ccaccgtgca gaccggcatc 180

aacatcgcgg gccgcatcct cggcgtgctc ggcgtgccct ttgcgggcca gctcgcctcc 240

ttctactcct tcctcgtggg agagctgtgg ccctccggcc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catccgccag caagtcaccg agaacacccg caacaccgcc 360

atcgcccgcc tggagggcct gggccgtggc taccgctcct accagcaagc cctggagacc 420

tggctcgaca accgcaacga cgcccgctcc cgctccatca tcctggagcg ctacgtcgcc 480

ctggaactgg acatcaccac tgccatccca ctcttccgca tcaggaacga ggaggtgcct 540

ctgctgatgg tgtacgccca ggctgcgaac ctgcacctgc tgctgctgcg cgacgcaagc 600

ctgtttggct ccgagtgggg tatggcaagc tccgacgtca accagtacta ccaggagcag 660

atccgctaca ccgaggagta cagcaaccac tgcgtccagt ggtacaacac cggtctgaac 720

aatctcagag ggaccaacgc tgagagctgg ctgcgctaca accagttccg gcgggatctg 780

accctaggtg tcctggatct ggtcgctctg ttcccgagct acgataccag gacgtaccct 840

atcaacacct ctgctcagct taccagggag atctacactg atcctatcgg taggactaac 900

gctcctagtg gtttcgccag cactaactgg ttcaacaaca acgcgcctag tttctctgcc 960

atcgaggcgg cgatcttccg gcctcctcac ctcctcgact tcccggagca gcttactatc 1020

tactctgcgt cttcgcggtg gtcttcgact cagcacatga actactgggt tggtcaccgg 1080

cttaacttcc gcccgattgg aggaactctt aacaccagta cgcaaggtct tacgaacaac 1140

acttccatca acccggttac gttgcagttc acgtctcggg acgtttaccg gacggagtcg 1200

aatgctggga cgaacatcct gttcacgaca ccggtgaatg gtgttccgtg ggcacgtttc 1260

aacttcatca acccgcagaa catctacgag cgtggagcaa cgacatactc gcaaccatac 1320

caaggcgttg gcatccaact gtttgactcg gagacggaac tgccaccaga gacgacagaa 1380

cgtccgaatt acgagtcata ctcacacaga ctatcacaca ttggactcat tatcggaaac 1440

acactgagag caccagtgta ctcatggaca catcggtcag cagatcgtac gaacatcatc 1500

gcctccgact ccatcaccca gattccggct gtgaagggca acttcctgtt caacggcagc 1560

gtcatctccg gccctgggtt caccggcggc gatctcgtca gactcaactc ctccggcaac 1620

aacatccaga atcgtggcta cattgaggtt cccatccact tcccatccac ctccacccgc 1680

tatcgcgtgc gcgtgcgcta cgcatccgtg actccgatcc atctcaacgt caactggggc 1740

aactcctcca tcttctccaa caccgtgccc gccaccgcga cctcccttga caacctccag 1800

agcagcgact tcggctactt cgagtctgcc aatgccttca cctcctcctt gggcaacatc 1860

gtaggtgtga ggaacttctc cggcactgcg ggcgtcatca tcgaccgctt cgagttcatc 1920

cctgtgaccg ccactctgga agccgagtac aacctcgagc gtgcccagaa ggcagtgaat 1980

gagctgttca cgagcagcaa tcagatcggg ctcaagaccg acgtgaccga ctaccacatc 2040

gaccaagtca gcaatctggt ggagtgtctc agcgacgagt tctgcctgga cgagaagaag 2100

gaactcagcg agaaggtgaa gcacgccaaa cgtctgtccg acgaacgcaa tctgctgcaa 2160

gatccgaact tcagagggat caacaggcaa ctggaccgtg ggtggcgtgg gtccaccgac 2220

atcaccatcc aaggcggcga cgacgtcttc aaggagaact acgtcactct gctgggcacc 2280

ttcgacgagt gctatccgac ctacctgtac cagaagatcg acgagagcaa gctcaaggcg 2340

tacacccgct accaactcag aggctacatc gaagactccc aggatctgga aatctacctc 2400

attcgctaca acgcgaagca cgagactgtc aacgtgcccg gcactggcag cctgtggccg 2460

ctctccgcgc cgagccctat cggaaagtgt gcgcaccact cccaccactt cagcttggac 2520

atcgacgtgg gttgcacgga cctcaacgaa gacctgggtg tctgggtgat cttcaagatc 2580

aagacgcaag acggacacgc gagactcggc aatctggagt tcctggagga gaagcctctg 2640

gtcggcgaag ctctcgccag ggtgaagcgc gccgagaaga agtggcgcga caagcgggag 2700

aagctggagt gggagacgaa catcgtgtac aaggaggcca aggagtccgt ggacgccctc 2760

tttgtgaaca gccagtacga ccgcctccaa gcagacacca acattgccat gatccacgcc 2820

gccgacaagc gcgtgcactc catcagggag gcgtatctcc ctgaactcag cgtgattcct 2880

ggcgtgaatg ctgccatctt cgaggagctt gagggtcgca tcttcaccgc attcagcctc 2940

tacgacgcca ggaacgtcat caagaacgga gacttcaaca acgggctgtc ctgctggaac 3000

gtcaaaggtc acgttgacgt ggaggagcag aacaaccatc gcagcgtcct cgtcgtccct 3060

gagtgggaag ctgaggtcag ccaagaagtc agggtctgcc ctggtcgcgg ctacatcctc 3120

agggtgaccg cctacaagga agggtatggt gaagggtgtg tcacgatcca tgagattgag 3180

aacaacacgg acgaactcaa gttcagcaac tgcgtcgagg aggaggtcta tccgaacaac 3240

acggtcacct gcaacgacta cactgccact caggaggagt acgagggaac ctacactagc 3300

cgcaaccgtg gctacgatgg agcctacgag tccaactcct ccgttcccgc tgactacgca 3360

agcgcctacg aggagaaggc gtacacggac gggcggcgcg acaacccgtg cgagtccaat 3420

cgtggctacg gtgactacac tccgcttccc gctggttacg tgaccaagga acttgagtac 3480

ttccctgaga ccgacaaggt ctggatcgag ataggtgaga ccgaaggcac gttcatcgtt 3540

gactccgtcg aactgctgct catggaggag tga 3573

<210> 119
<211> 3567
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC846.

<400> 119
atgaccagca accggaagaa cgagaacgag atcatcaacg ccctgagcat cccgaccgtg 60

agcaacccta gcacccagat gaacctgagc cctgacgctc gcatcgagga ctccctctgc 120

gtggctgagg tgaacaacat cgacccgttc gtgtccgcct ccaccgtgca gaccggcatc 180

aacatcgcgg gccgcatcct cggcgtgctc ggcgtgccct ttgcgggcca gctcgcctcc 240

ttctactcct tcctcgtggg agagctgtgg ccctccggcc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catccgccag caagtcaccg agaacacccg caacaccgcc 360

atcgcccgcc tggagggcct gggccgtggc taccgctcct accagcaagc cctggagacc 420

tggctcgaca accgcaacga cgcccgctcc cgctccatca tcctggagcg ctacgtcgcc 480

ctggaactgg acatcaccac tgccatccca ctcttccgca tcaggaacga ggaggtgcct 540

ctgctgatgg tgtacgccca ggctgcgaac ctgcacctgc tgctgctgcg cgacgcaagc 600

ctgtttggct ccgagtgggg tatggcaagc tccgacgtca accagtacta ccaggagcag 660

atccgctaca ccgaggagta cagcaaccac tgcgtccagt ggtacaacac cggtctgaac 720

aatctcagag ggaccaacgc tgagagctgg ctgcgctaca accagttccg gcgggatctg 780

accctaggtg tcctggatct ggtcgctctg ttcccgagct acgataccag gacgtaccct 840

atcaacacct ctgctcagct taccagggag atctacactg atcctatcgg taggactaac 900

gctcctagtg gtttcgccag cactaactgg ttcaacaaca acgcgcctag tttctctgcc 960

atcgaggcgg cgatcttccg gcctcctcac ctcctcgact tcccggagca gcttactatc 1020

tactctgcgt cttcgcggtg gtcttcgact cagcacatga actactgggt tggtcaccgg 1080

cttaacttcc gcccgattgg aggaactctt aacaccagta cgcaaggtct tacgaacaac 1140

acttccatca acccggttac gttgcagttc acgtctcggg acgtttaccg gacggagtcg 1200

aatgctggga cgaacatcct gttcacgaca ccggtgaatg gtgttccgtg ggcacgtttc 1260

aacttcatca acccgcagaa catctacgag cgtggagcaa cgacatactc gcaaccatac 1320

caaggcgttg gcatccaact gtttgactcg gagacggaac tgccaccaga gacgacagaa 1380

cgtccgaatt acgagtcata ctcacacaga ctatcacaca ttggactcat tatcggaaac 1440

acactgagag caccagtgta ctcatggaca catcggtcag cagatcgtac gaacaccatc 1500

ggtcccaacc gcatcactca gatcccgctc gtgaaggcac tgaacctgca cagcggagtg 1560

accgtggtgg gcggccctgg cttcaccggc ggcgacattc ttcgccgcac gaacacaggc 1620

accttcgctg acatgcgtgt gaacatcacc ggcccactca gccaacgcta ccgcgttcgc 1680

atccgctacg ccagcaccac cgatctccag ttcttcaccc gcatcaatgg cacgtctgtg 1740

aaccagggca acttccagcg caccatgaac cgtggcgaca acctcgaatc tggcaacttc 1800

cgcactgcgg gcttctcgac gcccttctcc ttctccaacg cgcagtccac cttcaccttg 1860

ggcactcagg cgttctccaa ccaggaggtg tacatcgaca ggatcgagtt cgtaccagcc 1920

gaagtgacct tcgaggccga gtccgatctc gagcgtgccc agaaggcagt gaatgagctg 1980

ttcacgagca gcaatcagat cgggctcaag accgacgtga ccgactacca catcgaccaa 2040

gtcagcaatc tggtggagtg tctcagcgac gagttctgcc tggacgagaa gaaggaactc 2100

agcgagaagg tgaagcacgc caaacgtctg tccgacgaac gcaatctgct gcaagatccg 2160

aacttcagag ggatcaacag gcaactggac cgtgggtggc gtgggtccac cgacatcacc 2220

atccaaggcg gcgacgacgt cttcaaggag aactacgtca ctctgctggg caccttcgac 2280

gagtgctatc cgacctacct gtaccagaag atcgacgaga gcaagctcaa ggcgtacacc 2340

cgctaccaac tcagaggcta catcgaagac tcccaggatc tggaaatcta cctcattcgc 2400

tacaacgcga agcacgagac tgtcaacgtg cccggcactg gcagcctgtg gccgctctcc 2460

gcgccgagcc ctatcggaaa gtgtgcgcac cactcccacc acttcagctt ggacatcgac 2520

gtgggttgca cggacctcaa cgaagacctg ggtgtctggg tgatcttcaa gatcaagacg 2580

caagacggac acgcgagact cggcaatctg gagttcctgg aggagaagcc tctggtcggc 2640

gaagctctcg ccagggtgaa gcgcgccgag aagaagtggc gcgacaagcg ggagaagctg 2700

gagtgggaga cgaacatcgt gtacaaggag gccaaggagt ccgtggacgc cctctttgtg 2760

aacagccagt acgaccgcct ccaagcagac accaacattg ccatgatcca cgccgccgac 2820

aagcgcgtgc actccatcag ggaggcgtat ctccctgaac tcagcgtgat tcctggcgtg 2880

aatgctgcca tcttcgagga gcttgagggt cgcatcttca ccgcattcag cctctacgac 2940

gccaggaacg tcatcaagaa cggagacttc aacaacgggc tgtcctgctg gaacgtcaaa 3000

ggtcacgttg acgtggagga gcagaacaac catcgcagcg tcctcgtcgt ccctgagtgg 3060

gaagctgagg tcagccaaga agtcagggtc tgccctggtc gcggctacat cctcagggtg 3120

accgcctaca aggaagggta tggtgaaggg tgtgtcacga tccatgagat tgagaacaac 3180

acggacgaac tcaagttcag caactgcgtc gaggaggagg tctatccgaa caacacggtc 3240

acctgcaacg actacactgc cactcaggag gagtacgagg gaacctacac tagccgcaac 3300

cgtggctacg atggagccta cgagtccaac tcctccgttc ccgctgacta cgcaagcgcc 3360

tacgaggaga aggcgtacac ggacgggcgg cgcgacaacc cgtgcgagtc caatcgtggc 3420

tacggtgact acactccgct tcccgctggt tacgtgacca aggaacttga gtacttccct 3480

gagaccgaca aggtctggat cgagataggt gagaccgaag gcacgttcat cgttgactcc 3540

gtcgaactgc tgctcatgga ggagtga 3567

<210> 120
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC858.

<400> 120
atgaccagca accggaagaa cgagaacgag atcatcaacg ccctgagcat cccgaccgtg 60

agcaacccta gcacccagat gaacctgagc cctgacgctc gcatcgagga ctccctctgc 120

gtggctgagg tgaacaacat cgacccgttc gtgtccgcct ccaccgtgca gaccggcatc 180

aacatcgcgg gccgcatcct cggcgtgctc ggcgtgccct ttgcgggcca gctcgcctcc 240

ttctactcct tcctcgtggg agagctgtgg ccctccggcc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catccgccag caagtcaccg agaacacccg caacaccgcc 360

atcgcccgcc tggagggcct gggccgtggc taccgctcct accagcaagc cctggagacc 420

tggctcgaca accgcaacga cgcccgctcc cgctccatca tcctggagcg ctacgtcgcc 480

ctggaactgg acatcaccac tgccatccca ctcttccgca tcaggaacga ggaggtgcct 540

ctgctgatgg tgtacgccca ggctgcgaac ctgcacctgc tgctgctgcg cgacgcaagc 600

ctgtttggct ccgagtgggg tatggcaagc tccgacgtca accagtacta ccaggagcag 660

atccgctaca ccgaggagta cagcaaccac tgcgtccagt ggtacaacac cggtctgaac 720

aatctcagag ggaccaacgc tgagagctgg ctgcgctaca accagttccg gcgggatctg 780

accctaggtg tcctggatct ggtcgctctg ttcccgagct acgataccag gacgtaccct 840

atcaacacct ctgctcagct taccagggag atctacactg atcctatcgg taggactaac 900

gctcctagtg gtttcgccag cactaactgg ttcaacaaca acgcgcctag tttctctgcc 960

atcgaggcgg cgatcttccg gcctcctcac ctcctcgact tcccggagca gcttactatc 1020

tactctgcgt cttcgcggtg gtcttcgact cagcacatga actactgggt tggtcaccgg 1080

cttaacttcc gcccgattgg aggaactctt aacaccagta cgcaaggtct tacgaacaac 1140

acttccatca acccggttac gttgcagttc acgtctcggg acgtttaccg gacggagtcg 1200

aatgctggga cgaacatcct gttcacgaca ccggtgaatg gtgttccgtg ggcacgtttc 1260

aacttcatca acccgcagaa catctacgag cgtggagcaa cgacatactc gcaaccatac 1320

caaggcgttg gcatccaact gtttgactcg gagacggaac tgccaccaga gacgacagaa 1380

cgtccgaatt acgagtcata ctcacacaga ctatcacaca ttggactcat tatcggaaac 1440

acactgagag caccagtgta ctcatggaca catcggtcag cagatcgtac gaacaccatc 1500

ggtccgaacc gcatcactca gatccctatg gtcaaggctt ccgaacttcc tcaaggaacc 1560

accgtggtgc gtggcccagg cttcacaggc ggcgacattc tccgccgcac gaacactggc 1620

ggcttcggcc cgattcgtgt gacagtcaat ggtccgctga cccaacgcta tcgcataggc 1680

tttcgctacg cctccaccgt tgacttcgac ttcttcgtgt cgcgtggtgg tacgaccgtc 1740

aacaacttcc gcttccttcg caccatgaac tcaggagacg agttgaagta cggcaacttc 1800

gtacgccgcg ccttcaccac gccgttcacc ttcacccaga tccaggacat catcaggacc 1860

agcatccaag gtctctctgg caatggcgaa gtgtacatcg acaagatcga gatcatccca 1920

gtgactgcga cctttgaagc agagtacgat ctcgagcgtg cccagaaggc agtgaatgag 1980

ctgttcacga gcagcaatca gatcgggctc aagaccgacg tgaccgacta ccacatcgac 2040

caagtcagca atctggtgga gtgtctcagc gacgagttct gcctggacga gaagaaggaa 2100

ctcagcgaga aggtgaagca cgccaaacgt ctgtccgacg aacgcaatct gctgcaagat 2160

ccgaacttca gagggatcaa caggcaactg gaccgtgggt ggcgtgggtc caccgacatc 2220

accatccaag gcggcgacga cgtcttcaag gagaactacg tcactctgct gggcaccttc 2280

gacgagtgct atccgaccta cctgtaccag aagatcgacg agagcaagct caaggcgtac 2340

acccgctacc aactcagagg ctacatcgaa gactcccagg atctggaaat ctacctcatt 2400

cgctacaacg cgaagcacga gactgtcaac gtgcccggca ctggcagcct gtggccgctc 2460

tccgcgccga gccctatcgg aaagtgtgcg caccactccc accacttcag cttggacatc 2520

gacgtgggtt gcacggacct caacgaagac ctgggtgtct gggtgatctt caagatcaag 2580

acgcaagacg gacacgcgag actcggcaat ctggagttcc tggaggagaa gcctctggtc 2640

ggcgaagctc tcgccagggt gaagcgcgcc gagaagaagt ggcgcgacaa gcgggagaag 2700

ctggagtggg agacgaacat cgtgtacaag gaggccaagg agtccgtgga cgccctcttt 2760

gtgaacagcc agtacgaccg cctccaagca gacaccaaca ttgccatgat ccacgccgcc 2820

gacaagcgcg tgcactccat cagggaggcg tatctccctg aactcagcgt gattcctggc 2880

gtgaatgctg ccatcttcga ggagcttgag ggtcgcatct tcaccgcatt cagcctctac 2940

gacgccagga acgtcatcaa gaacggagac ttcaacaacg ggctgtcctg ctggaacgtc 3000

aaaggtcacg ttgacgtgga ggagcagaac aaccatcgca gcgtcctcgt cgtccctgag 3060

tgggaagctg aggtcagcca agaagtcagg gtctgccctg gtcgcggcta catcctcagg 3120

gtgaccgcct acaaggaagg gtatggtgaa gggtgtgtca cgatccatga gattgagaac 3180

aacacggacg aactcaagtt cagcaactgc gtcgaggagg aggtctatcc gaacaacacg 3240

gtcacctgca acgactacac tgccactcag gaggagtacg agggaaccta cactagccgc 3300

aaccgtggct acgatggagc ctacgagtcc aactcctccg ttcccgctga ctacgcaagc 3360

gcctacgagg agaaggcgta cacggacggg cggcgcgaca acccgtgcga gtccaatcgt 3420

ggctacggtg actacactcc gcttcccgct ggttacgtga ccaaggaact tgagtacttc 3480

cctgagaccg acaaggtctg gatcgagata ggtgagaccg aaggcacgtt catcgttgac 3540

tccgtcgaac tgctgctcat ggaggagtga 3570

<210> 121
<211> 3564
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC866.

<400> 121
atgacgtcca accggaagaa cgagaacgag atcatcaacg cgctgtcaat cccggcggtg 60

tccaaccact cggcccagat gaacctgtcc acggacgcca ggatcgagga cagcctctgc 120

atcgccgagg gcaacaacat tgacccgttc gttagcgcca gcaccgtcca gaccggcatt 180

aacatcgcgg gccgcatcct cggcgttctg ggcgtcccgt tcgcgggcca gatcgcctca 240

ttctacagct tcctggtggg agagctgtgg ccgcgtgggc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catcaggcag caagtcaccg agaacacacg agacaccgca 360

ttggcgcggc tccagggcct gggcaactca ttccgcgcgt accagcagag cctggaggac 420

tggctggaga accgagacga cgcgcgcact cgctccgtcc tgtacacgca gtacatcgcc 480

cttgagctgg acttcctcaa cgcgatgcca ctgttcgcca tccggaacca ggaggtgccg 540

ctcctcatgg tgtacgctca ggctgccaac ctccacctgc tgctgctgcg ggacgcatcc 600

ctgttcgggt cggagttcgg cctgacctcc caggagattc agcgctacta cgagcggcaa 660

gtggagaaga caagggagta cagtgattat tgcgcccgct ggtacaacac cgggttgaac 720

aacttgcgag gcaccaacgc ggagtcatgg ctccgctaca atcagttccg gcgagacctc 780

acgctgggcg tgctcgacct cgtggccctc ttcccttcgt acgacacgcg cgtgtaccca 840

atgaacacca gcgcccagct cacgcgcgaa atctacacgg accctatcgg tcgtacaaat 900

gcaccgtcgg gcttcgcctc gacaaactgg tttaacaata acgcgccttc tttctcagca 960

atcgaggccg ctgtgattcg gcctccgcat ctgctcgatt tcccggagca gcttactatc 1020

ttctctgttc tcagccgctg gagcaacaca cagtacatga actattgggt cgggcaccgt 1080

ctggagagcc gcactattcg aggctcactg tcaaccagca cacacgggaa caccaacacc 1140

agcatcaacc cggtcaccct gcaattcact tcccgcgatg tgtacaggac ggagagtttc 1200

gcgggcatta acattctgct cactactccg gtcaacggtg tcccgtgggc ccgcttcaac 1260

tggaggaatc cgctgaacag cctgcgtggc tcccttctct acacaatcgg ttacacgggt 1320

gtcggcaccc agctcttcga ctctgaaacc gagctgccac ctgaaactac agagcgacct 1380

aactacgagt cctacagtca ccgtctgagc aacatccgcc taatcagcgg gaacactctt 1440

cgagcgccgg tgtactcctg gacccacaga tcagcagacc ggaccaacac tatctcgtcc 1500

gactccatca cccagatccc tctcgtgaag gctcacaccc ttcagagcgg gacgaccgtg 1560

gtgcgcgggc cgggcttcac cggtggcgac atcctgcgac gtacttccgg cggtccattc 1620

gcatacacaa tcgtgaacat caacggacag ctcccgcagc gttaccgggc caggataaga 1680

tacgccagca caacgaacct ccggatctac gtgaccgtgg ctggcgaacg gatctttgct 1740

ggtcagttca ataagacaat ggacacgggc gacccgctga cgttccaatc gttctcgtac 1800

gcgaccatca atacggcatt cacgtttccg atgtcccagt cctccttcac agtcggtgct 1860

gatacgttca gctccgggaa cgaggtgtac atcgaccgct ttgaactaat tccggtcacc 1920

gcgacgttcg aggctgagta cgatttggag cgcgcgcaga aggcagtcaa tgagctgttc 1980

acttccagca accagatcgg cctcaagacc gatgtgaccg actaccacat cgaccaagtg 2040

tcgaacctgg tggagtgcct gagtgacgag ttctgcttgg acgagaagaa ggagctgtct 2100

gagaaggtca aacatgctaa gaggctgtca gacgagcgta acctccttca agatcctaac 2160

ttccgtggca tcaaccgcca actggatcgc gggtggcgag gcagtacaga catcaccatc 2220

cagggaggcg atgacgtctt caaggagaac tacgttacgc tgctgggcac gttcgacgaa 2280

tgctacccga cttacctgta ccagaagatc gacgagagca agttgaaggc ttacacccgt 2340

taccagctca gaggttacat tgaggactcc caggatctgg aaatctacct gatccgctac 2400

aacgctaagc acgagaccgt gaatgtgccc ggcactggtt ccctgtggcc gctgtcggct 2460

ccgtcgccca tcggcaagtg cgcccaccac tcccaccact tcagtctgga cattgacgtc 2520

ggctgcaccg acctaaatga ggacttgggt gtctgggtta tcttcaagat caagacccag 2580

gacgggcacg ctcgtcttgg caacttggag tttctggagg agaagccgct ggtcggcgag 2640

gccctggccc gagtaaagcg ggccgagaag aagtggcgcg ataagcggga gaagctggag 2700

tgggagacca acatcgtgta caaggaggca aaggagagcg tggacgccct cttcgttaat 2760

tcgcagtacg accggttgca agctgacaca aacatcgcta tgatccacgc cgccgacaag 2820

agggtccact ccatcaggga ggcttacctg ccggaattgt cggtgatccc tggcgtgaat 2880

gctgccatct ttgaggagct ggagggacgg atcttcacag cgttctccct gtacgatgct 2940

aggaacgtca tcaagaacgg cgacttcaac aacggtctga gctgctggaa cgtgaagggc 3000

catgtggacg tggaggagca gaacaaccac cgcagcgtgc tggtcgtgcc cgagtgggaa 3060

gcggaggtca gccaggaggt gcgtgtctgc cctggacgcg gctacatcct gcgggtgacc 3120

gcgtacaagg agggctacgg cgaaggatgc gttaccatcc acgagattga gaacaatacg 3180

gacgagctta agtttagcaa ctgcgtggag gaggaggtgt acccgaataa cacggtgacc 3240

tgcaacgact acacggccac ccaagaggaa tacgaaggca cctacaccag ccgcaaccgt 3300

ggctacgacg gtgcctacga gagcaacagc tccgtgcccg ctgactacgc ttcggcctac 3360

gaggagaagg cgtacactga cggcagacgc gacaaccctt gcgagagcaa cagaggctac 3420

ggtgactaca cgccgctccc ggctggctac gtcaccaagg aactggagta cttcccggag 3480

accgacaagg tgtggataga gataggcgag accgagggca ctttcatcgt cgatagcgtc 3540

gagcttcttc tgatggagga gtga 3564

<210> 122
<211> 3525
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC838.

<400> 122
atggaggaga acaaccagaa ccagtgcatc ccgtacaact gcctgagcaa ccctgaggag 60

gtgctcctag acggtgagcg catctccacc ggcaactcta gcatcgacat ctccctgagc 120

ctagtgcagt tcctagtgag caacttcgtg cccggcggtg gcttcctcgt gggcctcatc 180

gacttcgtgt ggggcatcgt gggcccgtcc cagtgggacg ccttcctcgt gcagatcgag 240

cagctcatca acgagcgcat cgccgagttc gcccgcaacg ccgctatcgc caacctggag 300

ggcctcggca acaatttcaa catctacgtg gaggcgttca aggagtggga ggaagacccg 360

aacaatcccg ccacccgcac ccgcgtgatc gaccgcttcc gcatcctcga cggcctcctg 420

gagcgcgaca tcccgtcctt cgccatctcc ggcttcgagg tgccgctcct gtccgtgtac 480

gcccaggccg ctaacctcca cctcgccatc ctccgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactatcaa cgtgaacgag aactacaacc gcctcatccg ccacatcgac 600

gagtacgctg accactgcgc caacacctac aacaggggcc tcaacaatct gcccaagtcc 660

acctaccagg actggatcac ctacaacagg ctgaggcgcg acctgaccct gaccgtgctg 720

gacatcgctg cgttcttccc gaactacgac aacaggcgct accccatcca gagcgtcggc 780

cagctcacca gggagatcta caccgatccg ctcatcacct tcaaccctca gctccagtcc 840

gtcgctcagc tccctacctt caacgtcatg gagtccaacg ctatccgcac ccctcacctg 900

ttcgatgtcc tgaacaatct taccatcttc accgattggt tctccgtcgg caggaacttc 960

tactggggcg gtcacagggt cattagcaac aggatcggcg gtggcaacat taccagccct 1020

atctacggca gggaggctaa ccaggagcct ccgaggagct tcactttcaa cggtcctgtc 1080

ttccggactc ttagcaaccc tactttccgg cctcttcagc aaccttggcc cgcgccgcca 1140

ttcaaccttc ggggtgtcga gggtgtcgag ttcagcactc cgcttaacag cttcacttac 1200

cggggtagag gcactgtcga tagccttact gaacttccgc cagaagataa cagcgtcccg 1260

ccacgtgaag ggtacagcca ccgtctttgt cacgcgactt tcgtccagcg tagtgggact 1320

ccgttcctta cgacagggcc ggtctttagt tggacgcatc gctctgcgac ggacagaaac 1380

attatcgcgt ctgacagtat tacgcaaatt ccggcggtta agggcaactt cctcttcaat 1440

gggtctgtta tttctgggcc agggtttacg ggcggagatt tggttcgttt gaacagcagc 1500

gggaacaaca tccagaacag aggctacatt gaagtgccaa tccacttccc atctacgagc 1560

acgcgctatc gcgttagagt tcgatacgca tcggttacac caatccacct gaatgttaat 1620

tggggaaatt catcaatctt ctcgaacaca gttccagcaa cagcaacatc attagacaat 1680

ttgcaatcat cagatttcgg atacttcgaa tcggcaaacg cattcacatc gtcgttgggc 1740

aacatcgtag gagtacgcaa cttctcgggc accgccggcg tgatcattga ccgcttcgag 1800

tttatcccgg tcactgctac gctggaggcg gaatacaacc ttgagagagc gcagaaggcc 1860

gtgaacgcgc tgttcacctc gactaaccag ttgggactca agacaaatgt cactgactac 1920

cacatcgacc aagtgagtaa tctcgttacc tacctgtcgg acgagttctg cctggatgag 1980

aagagggaac tgtctgagaa agtgaagcac gcaaagcgcc tctcggatga gaggaacctc 2040

ctccaggatt ccaacttcaa ggacatcaac cggcagccgg aaagaggctg gggtggctcc 2100

accggcatta ctatccaggg cggcgacgat gtgttcaagg agaactacgt cacactctca 2160

ggcactttcg acgagtgtta cccgacgtac ctgtaccaga agattgatga gtccaagctg 2220

aaggcattca cccgctatca gcttcgcggt tacattgaag attcccagga ccttgagatc 2280

tacctcatca ggtacaacgc caagcacgag acggtcaacg tgcccggcac gggttccctg 2340

tggccgttgt ctgcccaatc tcccatcggg aagtgcggtg agccgaaccg ttgcgccccg 2400

cacctggagt ggaaccctga cctggattgc tcctgtaggg atggcgagaa gtgcgctcac 2460

cattcccacc acttctcgct ggacatcgac gtggggtgca cggatctcaa cgaggatctc 2520

ggagtctggg tcatcttcaa gatcaagacg caggacggac acgctcgcct cggcaatctg 2580

gagtttcttg aggaaaagcc gctcgtaggt gaggcgctcg ccagggtaaa gcgcgcggag 2640

aagaagtgga gggataagcg tgagaaactt gaatgggaga caaacatcgt ctacaaggaa 2700

gcaaaggagt cagttgatgc actgttcgtc aactctcagt atgaccagtt gcaggccgat 2760

accaacatcg cgatgatcca cgccgctgac aagcgcgttc actcgatccg cgaggcgtac 2820

ctccctgagc tttccgtcat cccaggcgtc aatgcggcca tcttcgagga acttgagggc 2880

cggatcttca ccgctttctc gctctacgac gctagaaacg tcatcaagaa cggcgacttc 2940

aacaatggcc tgagctgctg gaatgtgaag ggtcatgtgg atgtcgaaga gcagaacaat 3000

cagcgttctg tgctcgtggt ccccgaatgg gaagctgagg tttcccaaga agtgagagtg 3060

tgcccaggac gcggttacat tctccgtgtt acggcgtaca aggagggcta tggcgaaggc 3120

tgcgtgacca ttcatgagat cgagaacaac acggatgagc tgaagttctc aaactgcgtc 3180

gaggaagaaa tctaccctaa caacactgtg acctgtaacg actacacggt gaaccaggaa 3240

gagtacggcg gcgcttacac tagccgcaac cgtgggtaca acgaggcccc atcggtccct 3300

gctgactacg catccgtcta cgaggagaag tcttacactg acgggagacg ggagaacccc 3360

tgcgagttca accggggtta cagggattac acgcctctgc ccgttggcta tgtcactaag 3420

gagctggagt atttcccgga gactgataaa gtctggatcg agatcgggga gactgaaggg 3480

actttcatcg tcgattccgt ggagctgctt ctgatggagg agtga 3525

<210> 123
<211> 3519
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC841.

<400> 123
atggaggaga acaaccagaa ccagtgcatc ccgtacaact gcctgagcaa ccctgaggag 60

gtgctcctag acggtgagcg catctccacc ggcaactcta gcatcgacat ctccctgagc 120

ctagtgcagt tcctagtgag caacttcgtg cccggcggtg gcttcctcgt gggcctcatc 180

gacttcgtgt ggggcatcgt gggcccgtcc cagtgggacg ccttcctcgt gcagatcgag 240

cagctcatca acgagcgcat cgccgagttc gcccgcaacg ccgctatcgc caacctggag 300

ggcctcggca acaatttcaa catctacgtg gaggcgttca aggagtggga ggaagacccg 360

aacaatcccg ccacccgcac ccgcgtgatc gaccgcttcc gcatcctcga cggcctcctg 420

gagcgcgaca tcccgtcctt cgccatctcc ggcttcgagg tgccgctcct gtccgtgtac 480

gcccaggccg ctaacctcca cctcgccatc ctccgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactatcaa cgtgaacgag aactacaacc gcctcatccg ccacatcgac 600

gagtacgctg accactgcgc caacacctac aacaggggcc tcaacaatct gcccaagtcc 660

acctaccagg actggatcac ctacaacagg ctgaggcgcg acctgaccct gaccgtgctg 720

gacatcgctg cgttcttccc gaactacgac aacaggcgct accccatcca gagcgtcggc 780

cagctcacca gggagatcta caccgatccg ctcatcacct tcaaccctca gctccagtcc 840

gtcgctcagc tccctacctt caacgtcatg gagtccaacg ctatccgcac ccctcacctg 900

ttcgatgtcc tgaacaatct taccatcttc accgattggt tctccgtcgg caggaacttc 960

tactggggcg gtcacagggt cattagcaac aggatcggcg gtggcaacat taccagccct 1020

atctacggca gggaggctaa ccaggagcct ccgaggagct tcactttcaa cggtcctgtc 1080

ttccggactc ttagcaaccc tactttccgg cctcttcagc aaccttggcc cgcgccgcca 1140

ttcaaccttc ggggtgtcga gggtgtcgag ttcagcactc cgcttaacag cttcacttac 1200

cggggtagag gcactgtcga tagccttact gaacttccgc cagaagataa cagcgtcccg 1260

ccacgtgaag ggtacagcca ccgtctttgt cacgcgactt tcgtccagcg tagtgggact 1320

ccgttcctta cgacagggcc ggtctttagt tggacgcatc gctctgcgac ggacagaaac 1380

accatcgacc cggagcgcat caaccagatc cctctcgtga agggcttccg cgtgtggggc 1440

ggcacctccg tcatcaccgg ccctggcttc accggcggcg acattcttcg ccgcacgaac 1500

acaggcacct tcgctgacat gcgtgtgaac atcaccggcc cactcagcca acgctaccgc 1560

gttcgcatcc gctacgccag caccaccgat ctccagttct tcacccgcat caatggcacg 1620

tctgtgaacc agggcaactt ccagcgcacc atgaaccgtg gcgacaacct cgaatctggc 1680

aacttccgca ctgcgggctt ctcgacgccc ttctccttct ccaacgcgca gtccaccttc 1740

accttgggca ctcaggcgtt ctccaaccag gaggtgtaca tcgacaggat cgagttcgta 1800

ccagccgaag tgaccttcga ggccgagtcc gatctcgagc gcgcccagaa ggccgtgaac 1860

gccctcttca ctagcactaa ccagctcggc ctcaagacta acgtgaccga ctaccacatt 1920

gaccaagtga gcaacctagt gacctacctt agcgacgagt tctgccttga cgagaagcgt 1980

gagctgagcg agaaggtgaa gcacgccaag cgcctctccg acgagcgcaa cctcctccag 2040

gactccaact tcaaggacat caaccgccag cccgagcgcg gctggggcgg tagcaccggc 2100

atcaccatcc agggcggtga cgatgtgttc aaggagaact acgtgaccct ctccggcacc 2160

ttcgacgagt gctacccgac ctacctctac cagaagatcg acgagtccaa gctcaaggcg 2220

ttcacccgct accagcttcg cggctacatc gaggactccc aggatctgga gatctacctc 2280

atccgctaca acgccaagca cgagaccgtg aacgtgcccg gcaccggctc cctctggccg 2340

ctctccgccc agagccctat cggcaagtgc ggcgagccca accgctgcgc gcctcacctg 2400

gagtggaacc ctgacctcga ctgctcctgc cgcgacggcg agaagtgcgc ccaccatagc 2460

caccacttct ctctcgacat cgacgtgggc tgcaccgacc tcaacgagga tctgggcgtg 2520

tgggtgatct tcaagatcaa gacccaggac ggccacgcca ggctgggcaa cctggagttc 2580

ctggaggaga agcctctggt gggtgaggcc ctggccaggg tcaagagggc tgagaagaaa 2640

tggagggaca agagggagaa gctggagtgg gagaccaaca tcgtgtacaa ggaggctaag 2700

gagtccgtgg acgctctgtt cgtcaactct cagtacgatc agctccaggc tgacaccaac 2760

atcgctatga tccacgctgc ggataagagg gtccactcta tcagggaggc ttacctgcct 2820

gagctttctg tcatccctgg tgtcaacgcg gcaatcttcg aggaacttga gggccgcatc 2880

ttcactgcgt tctcgcttta cgatgcgcgg aacgtcatta agaacggtga cttcaacaat 2940

ggtctttcgt gctggaacgt caagggtcat gtcgatgtcg aggaacagaa caaccagcgg 3000

tcggtccttg tcgttcccga gtgggaggcc gaggtctcgc aagaggtccg ggtctgccct 3060

gggcgcgggt acattcttcg tgtcactgcg tacaaggagg gctacggcga gggctgcgtt 3120

actattcatg agattgagaa caatacggat gagcttaagt ttagtaactg tgttgaggag 3180

gagatctacc cgaacaatac ggttacgtgc aatgattaca cggtgaacca ggaggaatac 3240

ggcggagcat acacctcacg taatagaggg tacaatgagg caccgtcagt tccggcagat 3300

tatgcctcag tttatgagga gaagtcctac acggatggaa gacgcgagaa tccatgtgag 3360

tttaatagag gataccgaga ctacacacca ctcccagttg gatacgttac aaaggagttg 3420

gaatacttcc cagaaacaga taaagtttgg atagagatcg gagaaacaga aggaaccttc 3480

atcgtggaca gtgtagaact gctgctgatg gaagaatga 3519

<210> 124
<211> 3516
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC842.

<400> 124
atggctgaga acaacatcca gaaccagtgc gtgccctaca actgcctgaa caaccctgag 60

gtggagatcc tgaacgagga gcgtagcacc ggtaggctcc cgctagacat ctccctgagc 120

ctgacccgct tcctccttag cgagttcgtg cccggcgtgg gcgtggcctt cggcctcttc 180

gacctcatct ggggcttcat cacgccgtcc gactggtccc tcttcctcct ccagatcgag 240

cagctcatcg agcagcgcat cgagaccctg gagcgcaacc gcgccatcac cactctgcgc 300

ggcctcgccg actcctacga aatctacatc gaggccctcc gcgagtggga ggccaacccg 360

aacaatgccc agctccgcga ggacgtgcgg attcgcttcg ccaacaccga cgatgccctc 420

atcaccgcca tcaacaactt caccctcacc tccttcgaga tcccgctgct cagcgtgtac 480

gtgcaagccg ctaacctcca cctctccctc ctgcgcgacg ccgtgagctt cggccagggc 540

tggggcctcg acatcgccac cgtgaacaat cactacaacc gcctcatcaa cctcatccac 600

cgctacacca agcactgcct cgacacctac aaccagggcc tggagaacct gcgcggcacc 660

aacacccgcc agtgggcccg cttcaaccag ttccgccgcg acctcaccct caccgtgctc 720

gacatcgtgg ccctgttccc taactacgac gtccgcacct acccgatcca gacctccagc 780

cagctcacga gggaaatcta cacctccagc gtgatcgagg acagcccggt gtccgccaac 840

atcccgaacg gcttcaaccg cgccgagttc ggcgtgcgcc cgcctcacct catggacttc 900

atgaactccc tcttcgtcac tgccgagacc gtgcgctccc agaccgtgtg gggcggtcac 960

ctcgtgtcca gccgcaacac cgctggcaac cgcatcaact tcccgtccta cggcgtgttc 1020

aaccctggcg gtgccatctg gatcgccgac gaggacccga ggcccttcta ccgcaccctg 1080

tccgaccctg tgttcgtgcg cggcggtttc ggcaaccctc actacgtgct gggcctgagg 1140

ggcgtggcct tccagcaaac cggcaccaac cacacccgca ccttccgcaa ctccggcacc 1200

atcgactccc tggacgagat ccctcctcag gacaactccg gcgccccttg gaacgactac 1260

tcccacgtgc tgaaccacgt gaccttcgtg aggtggcctg gcgagatcag cggcagcgac 1320

tcctggaggg cccctatgtt ctcctggacc cacaggagcg ccaccccgac caacaccatc 1380

gaccctgagc gcatcaccca gatccctctg gtgaaggccc acactctcca gagcggcact 1440

actgtcgttc gcggccctgg cttcactggc ggtgacatcc tgaggcggac tagcggcggt 1500

cctttcgctt acactatcgt caacatcaac ggccagctcc ctcagcgcta cagggcccgc 1560

atccgctacg ccagcactac taacctgagg atctacgtca ctgtcgctgg cgagcgcatc 1620

ttcgccgggc agttcaacaa gacgatggac actggtgacc cgctgacttt ccagtctttc 1680

tcttacgcta ctatcaacac tgctttcact ttcccgatga gtcagtcttc gttcactgtc 1740

ggtgctgaca ctttctcttc gggcaacgag gtgtacatcg accggttcga gctgatcccg 1800

gtcactgcta ctttcgaggc tgagtacgac ctggagcggg ctcagaaggc tgtcaacgct 1860

ctgttcactt ctatcaacca gatcggtatt aagacggacg tcacggacta ccacattgac 1920

caagtcagca acctggtgga ctgcctgtct gacgagttct gcctggacga gaagcgggag 1980

ctgtctgaga aggtcaagca cgctaagcgg ctgtctgacg agcggaacct gctccaagac 2040

ccgaacttta agggtattaa ccgccagctc gaccgtggtt ggcgcggtag tacggacatt 2100

acgattcagc gcggtgacga tgtcttcaag gagaactacg tcacgctgcc gggtaccttc 2160

gacgagtgct acccgacgta cctttaccag aagattgatg agtctaagct caaggcgttt 2220

acgcgctacc agcttcgtgg ttacattgag gatagtcagg atcttgaaat ctaccttatt 2280

cgttacaacg ctaagcacga gacggtcaac gtgccgggta cgggctctct ttggccgctt 2340

tctgctcagt ctccgattgg gaagtgcggt gagccgaacc gttgcgcgcc gcaccttgag 2400

tggaacccgg atcttgattg ctcgtgccgt gatggcgaga agtgcgcgca ccatagtcat 2460

cacttcagtc ttgacattga tgtcgggtgc acggatctta acgaggatct tggcgtgtgg 2520

gtgatcttca agattaagac gcaagatggg cacgcgcgtc ttgggaacct tgagtttctt 2580

gaggagaagc cacttgtcgg cgaggcgctt gcgcgtgtca agcgtgcgga gaagaaatgg 2640

cgtgataagc gtgagaagtt ggagtgggag acgaacattg tgtacaagga ggcgaaggag 2700

tcggtcgatg cgttgttcgt caacagtcag tacgatcagt tgcaagcgga tacgaacatc 2760

gccatgattc atgcggcaga taagcgtgtc cattcgatca gggaggcgta cctgccagag 2820

ttgtcggtca ttccaggcgt caatgcggca atcttcgagg aattggaggg cagaatcttc 2880

acggcattct cgctgtacga tgcaaggaat gtgattaaga atggcgattt caacaacggg 2940

ctgtcgtgct ggaatgttaa gggccatgtt gatgttgagg aacagaacaa ccagagatca 3000

gtgttggtcg taccagagtg ggaggcagag gtttcacaag aggtgagagt ctgcccaggg 3060

cgcggctaca tcttgagagt taccgcctac aaggagggat acggcgaggg atgtgtgaca 3120

attcacgaaa tcgagaacaa cacagacgag ctaaagttct caaactgtgt cgaggaagaa 3180

atctacccga ataacacagt cacatgcaat gattacacag tgaaccagga agaatacgga 3240

ggagcataca catcacgcaa tagaggatac aatgaagcac caagcgttcc agcagactac 3300

gcctcagtgt acgaggagaa gtcatacaca gacggacgaa gggagaatcc atgtgagttc 3360

aatcgaggat accgagatta cacaccactt ccagttggat acgttaccaa ggaactcgaa 3420

tacttcccag aaacagacaa agtgtggatc gaaatcggcg aaaccgaagg cacattcatc 3480

gtggactcag tcgaactgct gctgatggag gagtga 3516

<210> 125
<211> 3444
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC850.

<400> 125
atggagaaca acatccagaa ccagtgcgtg ccctacaact gcctgaacaa ccctgaggtt 60

gagatcctga acgaggagcg tagcaccggt aggctcccgc tagacatctc cctgagcctg 120

acccgcttcc tccttagtga gttcgtgccc ggcgtgggcg tggccttcgg cctcttcgac 180

ctcatctggg gcttcatcac tccttccgac tggtccctct tcctccttca gattgagcaa 240

ctgatcgagc agcgcatcga gacccttgag cgcaaccgcg ccatcaccac tctcagaggt 300

ctcgccgact cctacgaaat ctacatcgag gcactccgtg agtgggaggc caacccgaac 360

aatgcccagc tccgcgagga cgtgaggatc agattcgcca acaccgacga tgccctcatc 420

accgccatca acaatttcac cctcacctcc ttcgagatcc ctcttctgtc tgtgtacgtt 480

caagctgcta accttcacct ttccctcctg cgcgacgccg tgagcttcgg ccagggctgg 540

ggcctcgaca tcgccaccgt gaacaatcac tacaaccgcc tcatcaacct catccaccgc 600

tacaccaagc actgccttga cacctacaac cagggccttg agaacctccg tggcaccaac 660

acccgccagt gggcccgctt caaccagttc cgcagagacc tcaccctcac cgtgctcgac 720

atcgtggcac tcttcccaaa ctacgacgtg cgtacctacc ctatccagac ctccagccag 780

ctcaccaggg aaatctacac ctccagcgtg atcgaggact ctcctgtgtc cgccaacatc 840

cctaacggct tcaaccgcgc cgagttcggc gtgcgccctc ctcacctcat ggacttcatg 900

aactccctct tcgtcactgc cgagaccgtg cgctcccaga ccgtgtgggg cggtcacctc 960

gtgtccagcc gtaacaccgc tggcaacagg atcaacttcc cgtcctacgg cgtgttcaac 1020

ccaggcggtg ccatctggat cgccgatgaa gaccctcgtc ctttctaccg taccctgtcc 1080

gaccctgtgt tcgtgcgtgg cggtttcggc aaccctcact acgtgctggg cctgcgtggc 1140

gtggccttcc agcaaaccgg caccaaccac accaggacgt tccgtaactc cggcaccatc 1200

gacagtcttg acgagatccc tccgcaagac aactccggtg caccttggaa cgactactcc 1260

cacgtgctga accacgtgac cttcgtgagg tggcctggcg aaatctccgg ctccgactcc 1320

tggagggctc ctatgttcag ttggacccac aggagcgcta cgcctaccaa caccatcgat 1380

cctgaacgca tcactcagat ccctctcgtc aaggctcaca ccctccagtc aggcaccacc 1440

gtggtgcgcg gccctggctt caccggcggc gacatccttc gccgcacgaa caccggcacc 1500

tttgcagaca tgagggtcaa catcaccggt cctctctccc aacgctatcg cgtccgcatt 1560

cgctatgcct cgaccaccga tctacagttc ttcacccgga tcaatggcac gtctgtgaac 1620

cagggcaact tccaacgcac gatgaaccgt ggtgacaacc tggagtccgg gaacttccga 1680

actgctggct tctccacgcc cttcagcttc tccaatgccc agtccacctt cactcttggc 1740

acccaagcgt tcagcaacca ggaagtgtac atcgaccgca ttgagtttgt tcccgctgag 1800

gtgaccttcg aggccgagtc tgatctcgag cgtgcccaga aggcagtgaa tgagctgttc 1860

acgagcagca atcagatcgg gctcaagacc gacgtgaccg actaccacat cgaccaagtc 1920

agcaatctgg tggagtgtct cagcgacgag ttctgcctgg acgagaagaa ggaactcagc 1980

gagaaggtga agcacgccaa acgtctgtcc gacgaacgca atctgctgca agatccgaac 2040

ttcagaggga tcaacaggca actggaccgt gggtggcgtg ggtccaccga catcaccatc 2100

caaggcggcg acgacgtctt caaggagaac tacgtcactc tgctgggcac cttcgacgag 2160

tgctatccga cctacctgta ccagaagatc gacgagagca agctcaaggc gtacacccgc 2220

taccaactca gaggctacat cgaagactcc caggatctgg aaatctacct cattcgctac 2280

aacgcgaagc acgagactgt caacgtgccc ggcactggca gcctgtggcc gctctccgcg 2340

ccgagcccta tcggaaagtg tgcgcaccac tcccaccact tcagcttgga catcgacgtg 2400

ggttgcacgg acctcaacga agacctgggt gtctgggtga tcttcaagat caagacgcaa 2460

gacggacacg cgagactcgg caatctggag ttcctggagg agaagcctct ggtcggcgaa 2520

gctctcgcca gggtgaagcg cgccgagaag aagtggcgcg acaagcggga gaagctggag 2580

tgggagacga acatcgtgta caaggaggcc aaggagtccg tggacgccct ctttgtgaac 2640

agccagtacg accgcctcca agcagacacc aacattgcca tgatccacgc cgccgacaag 2700

cgcgtgcact ccatcaggga ggcgtatctc cctgaactca gcgtgattcc tggcgtgaat 2760

gctgccatct tcgaggagct tgagggtcgc atcttcaccg cattcagcct ctacgacgcc 2820

aggaacgtca tcaagaacgg agacttcaac aacgggctgt cctgctggaa cgtcaaaggt 2880

cacgttgacg tggaggagca gaacaaccat cgcagcgtcc tcgtcgtccc tgagtgggaa 2940

gctgaggtca gccaagaagt cagggtctgc cctggtcgcg gctacatcct cagggtgacc 3000

gcctacaagg aagggtatgg tgaagggtgt gtcacgatcc atgagattga gaacaacacg 3060

gacgaactca agttcagcaa ctgcgtcgag gaggaggtct atccgaacaa cacggtcacc 3120

tgcaacgact acactgccac tcaggaggag tacgagggaa cctacactag ccgcaaccgt 3180

ggctacgatg gagcctacga gtccaactcc tccgttcccg ctgactacgc aagcgcctac 3240

gaggagaagg cgtacacgga cgggcggcgc gacaacccgt gcgagtccaa tcgtggctac 3300

ggtgactaca ctccgcttcc cgctggttac gtgaccaagg aacttgagta cttccctgag 3360

accgacaagg tctggatcga gataggtgag accgaaggca cgttcatcgt tgactccgtc 3420

gaactgctgc tcatggagga gtga 3444

<210> 126
<211> 3447
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC859.

<400> 126
atggagaaca acatccagaa ccagtgcgtg ccctacaact gcctgaacaa ccctgaggtt 60

gagatcctga acgaggagcg tagcaccggt aggctcccgc tagacatctc cctgagcctg 120

acccgcttcc tccttagtga gttcgtgccc ggcgtgggcg tggccttcgg cctcttcgac 180

ctcatctggg gcttcatcac tccttccgac tggtccctct tcctccttca gattgagcaa 240

ctgatcgagc agcgcatcga gacccttgag cgcaaccgcg ccatcaccac tctcagaggt 300

ctcgccgact cctacgaaat ctacatcgag gcactccgtg agtgggaggc caacccgaac 360

aatgcccagc tccgcgagga cgtgaggatc agattcgcca acaccgacga tgccctcatc 420

accgccatca acaatttcac cctcacctcc ttcgagatcc ctcttctgtc tgtgtacgtt 480

caagctgcta accttcacct ttccctcctg cgcgacgccg tgagcttcgg ccagggctgg 540

ggcctcgaca tcgccaccgt gaacaatcac tacaaccgcc tcatcaacct catccaccgc 600

tacaccaagc actgccttga cacctacaac cagggccttg agaacctccg tggcaccaac 660

acccgccagt gggcccgctt caaccagttc cgcagagacc tcaccctcac cgtgctcgac 720

atcgtggcac tcttcccaaa ctacgacgtg cgtacctacc ctatccagac ctccagccag 780

ctcaccaggg aaatctacac ctccagcgtg atcgaggact ctcctgtgtc cgccaacatc 840

cctaacggct tcaaccgcgc cgagttcggc gtgcgccctc ctcacctcat ggacttcatg 900

aactccctct tcgtcactgc cgagaccgtg cgctcccaga ccgtgtgggg cggtcacctc 960

gtgtccagcc gtaacaccgc tggcaacagg atcaacttcc cgtcctacgg cgtgttcaac 1020

ccaggcggtg ccatctggat cgccgatgaa gaccctcgtc ctttctaccg taccctgtcc 1080

gaccctgtgt tcgtgcgtgg cggtttcggc aaccctcact acgtgctggg cctgcgtggc 1140

gtggccttcc agcaaaccgg caccaaccac accaggacgt tccgtaactc cggcaccatc 1200

gacagtcttg acgagatccc tccgcaagac aactccggtg caccttggaa cgactactcc 1260

cacgtgctga accacgtgac cttcgtgagg tggcctggcg aaatctccgg ctccgactcc 1320

tggagggctc ctatgttcag ttggacccac aggagcgcta cgcctaccaa cgaggtccgt 1380

gtgtcccgca tcacccagct ccctatggtg aaggctcaca cccttcacgc tggtgccaca 1440

gtagtcagag gtccgggctt cactggcggc gacattcttc gccgcaccac ctccggctcc 1500

ttcggcgaca tgaggatcac caacttcagc tcctcctcca gtcgctatcg cgttcgtatt 1560

cgctacgcct ccaccaccga cctccagttc ttcctgtccg ttggcggcac gcctgtgaat 1620

gtggcagact tcccgaagac catcgaccgt ggcgagaacc tggagtacgg cagctttcgc 1680

actgctggct tcaccactcc gttcagcttc gtctcctcca ccaacaactt cacactaggc 1740

gtgcagtctg tctcctccgg caacgagatc ttcgtggatc ggatcgagtt cgtaccagcc 1800

gacgcgacct tcgaagccga gtacgatctc gagcgtgccc agaaggcagt gaatgagctg 1860

ttcacgagca gcaatcagat cgggctcaag accgacgtga ccgactacca catcgaccaa 1920

gtcagcaatc tggtggagtg tctcagcgac gagttctgcc tggacgagaa gaaggaactc 1980

agcgagaagg tgaagcacgc caaacgtctg tccgacgaac gcaatctgct gcaagatccg 2040

aacttcagag ggatcaacag gcaactggac cgtgggtggc gtgggtccac cgacatcacc 2100

atccaaggcg gcgacgacgt cttcaaggag aactacgtca ctctgctggg caccttcgac 2160

gagtgctatc cgacctacct gtaccagaag atcgacgaga gcaagctcaa ggcgtacacc 2220

cgctaccaac tcagaggcta catcgaagac tcccaggatc tggaaatcta cctcattcgc 2280

tacaacgcga agcacgagac tgtcaacgtg cccggcactg gcagcctgtg gccgctctcc 2340

gcgccgagcc ctatcggaaa gtgtgcgcac cactcccacc acttcagctt ggacatcgac 2400

gtgggttgca cggacctcaa cgaagacctg ggtgtctggg tgatcttcaa gatcaagacg 2460

caagacggac acgcgagact cggcaatctg gagttcctgg aggagaagcc tctggtcggc 2520

gaagctctcg ccagggtgaa gcgcgccgag aagaagtggc gcgacaagcg ggagaagctg 2580

gagtgggaga cgaacatcgt gtacaaggag gccaaggagt ccgtggacgc cctctttgtg 2640

aacagccagt acgaccgcct ccaagcagac accaacattg ccatgatcca cgccgccgac 2700

aagcgcgtgc actccatcag ggaggcgtat ctccctgaac tcagcgtgat tcctggcgtg 2760

aatgctgcca tcttcgagga gcttgagggt cgcatcttca ccgcattcag cctctacgac 2820

gccaggaacg tcatcaagaa cggagacttc aacaacgggc tgtcctgctg gaacgtcaaa 2880

ggtcacgttg acgtggagga gcagaacaac catcgcagcg tcctcgtcgt ccctgagtgg 2940

gaagctgagg tcagccaaga agtcagggtc tgccctggtc gcggctacat cctcagggtg 3000

accgcctaca aggaagggta tggtgaaggg tgtgtcacga tccatgagat tgagaacaac 3060

acggacgaac tcaagttcag caactgcgtc gaggaggagg tctatccgaa caacacggtc 3120

acctgcaacg actacactgc cactcaggag gagtacgagg gaacctacac tagccgcaac 3180

cgtggctacg atggagccta cgagtccaac tcctccgttc ccgctgacta cgcaagcgcc 3240

tacgaggaga aggcgtacac ggacgggcgg cgcgacaacc cgtgcgagtc caatcgtggc 3300

tacggtgact acactccgct tcccgctggt tacgtgacca aggaacttga gtacttccct 3360

gagaccgaca aggtctggat cgagataggt gagaccgaag gcacgttcat cgttgactcc 3420

gtcgaactgc tgctcatgga ggagtga 3447

<210> 127
<211> 3612
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC861.

<400> 127
atgaagctca agaaccctga caagcaccag tccttctcta gcaacgcgaa ggtggacaag 60

atcagcactg actccctcaa gaacgagact gacatcgagc ttcagaacat caaccacgag 120

gactgcctca agatcagcga gtacgagaac gtggaaccct tcgtgagcgc ctccaccatc 180

cagaccggca tctccattgc cgggaagatc ctcggcaccc tcggcgtacc cttcgcgggc 240

caagtggcct ccctctactc cttcatcctc ggcgaactgt ggcccaaagg caagaaccag 300

tgggagatct tcatggagca cgtggaggag atcatcaacc agaagatcag cacctatgcc 360

cgcaacaagg ccctcaccga cctcaagggc ctcggcgacg ccctcgccgt gtaccacgag 420

tccctggagt cctgggtcgg caaccgcaag aacacccgcg cccgctccgt ggtgaagtcc 480

cagtacatcg ccctggagct gatgttcgtg cagaagctcc cgtccttcgc cgtgagcggc 540

gaggaggtgc cgctcctgcc catctacgcc caggccgcta acctccacct cctgctcctg 600

cgcgacgcct ccatcttcgg caaggagtgg ggcctctcca gctccgagat cagcaccttc 660

tacaaccgcc aagtggaacg cgctggcgac tactccgatc actgcgtgaa gtggtacagc 720

accggcctca acaacctgcg cggcaccaac gctgagagct gggtgcgcta caaccagttc 780

cgcaaggaca tgaccctgat ggtcctggat ctggtcgcac tgttcccgag ctacgacact 840

ctggtgtacc cgatcaagac caccagccag ctcaccaggg aggtgtacac cgacgctatc 900

ggcactgtcc atcctaatgc gagcttcgcc agcaccacct ggtacaacaa caacgctcct 960

tccttcagca ccatcgagag cgctgtggtg aggaatcctc acctgcttga cttcctggag 1020

caagtcacca tctacagcct gctttctagg tggtccaaca cccagtacat gaacatgtgg 1080

ggcggtcaca ggctggagtt ccgcaccatc ggcggtatgc tgaacacctc tactcaggga 1140

agtaccaaca cttccatcaa ccctgtcact ctgcctttca cttctaggga cgtttacagg 1200

actgagtccc tggctggact taacctgttc cttactcagc ctgtcaacgg tgtcccgagg 1260

gtggacttcc actggaagtt cgtcactcat cccatcgcgt ctgacaactt ctactaccct 1320

ggttacgcgg gtatcggtac tcagctccag gacagtgaga atgagttgcc gccggagacg 1380

accggtcagc cgaactacga gtcttacagt cacaggcttt ctcacatcgg tctcatctct 1440

gcgtctcatg tcaaagcgct tgtgtactcg tggacgcatc gctcagcgga tcgcacgaat 1500

accatcgacc cagagcgcat caaccagatc ccgctcgtga agggcttccg cgtgtggggc 1560

ggcacctccg tcatcaccgg tccgggcttc accggcggcg acatcctccg ccgcaacacc 1620

ttcggcgact tcgtgtcact ccaagtgaac atcaacagcc cgatcaccca gcgctatcgc 1680

ctccgcttcc gctacgcctc ctcccgcgac gctagagtga tcgtgctcac cggagcggcg 1740

tccacaggcg taggcggcca agtgtctgtg aacatgccgc tccagaagac tatggagatt 1800

ggtgagaacc tcacctctcg caccttccgc tacaccgact tctccaatcc gttctccttc 1860

agagccaacc cagacatcat cggcatctcc gagcagcctc tctttggcgc tggctccatc 1920

tcctccggcg agctgtacat cgacaagatt gagatcatcc ttgccgacgc caccttcgaa 1980

gctgagtccg atctcgagcg tgcccagaag gcagtgaatg agctgttcac gagcagcaat 2040

cagatcgggc tcaagaccga cgtgaccgac taccacatcg accaagtcag caatctggtg 2100

gagtgtctca gcgacgagtt ctgcctggac gagaagaagg aactcagcga gaaggtgaag 2160

cacgccaaac gtctgtccga cgaacgcaat ctgctgcaag atccgaactt cagagggatc 2220

aacaggcaac tggaccgtgg gtggcgtggg tccaccgaca tcaccatcca aggcggcgac 2280

gacgtcttca aggagaacta cgtcactctg ctgggcacct tcgacgagtg ctatccgacc 2340

tacctgtacc agaagatcga cgagagcaag ctcaaggcgt acacccgcta ccaactcaga 2400

ggctacatcg aagactccca ggatctggaa atctacctca ttcgctacaa cgcgaagcac 2460

gagactgtca acgtgcccgg cactggcagc ctgtggccgc tctccgcgcc gagccctatc 2520

ggaaagtgtg cgcaccactc ccaccacttc agcttggaca tcgacgtggg ttgcacggac 2580

ctcaacgaag acctgggtgt ctgggtgatc ttcaagatca agacgcaaga cggacacgcg 2640

agactcggca atctggagtt cctggaggag aagcctctgg tcggcgaagc tctcgccagg 2700

gtgaagcgcg ccgagaagaa gtggcgcgac aagcgggaga agctggagtg ggagacgaac 2760

atcgtgtaca aggaggccaa ggagtccgtg gacgccctct ttgtgaacag ccagtacgac 2820

cgcctccaag cagacaccaa cattgccatg atccacgccg ccgacaagcg cgtgcactcc 2880

atcagggagg cgtatctccc tgaactcagc gtgattcctg gcgtgaatgc tgccatcttc 2940

gaggagcttg agggtcgcat cttcaccgca ttcagcctct acgacgccag gaacgtcatc 3000

aagaacggag acttcaacaa cgggctgtcc tgctggaacg tcaaaggtca cgttgacgtg 3060

gaggagcaga acaaccatcg cagcgtcctc gtcgtccctg agtgggaagc tgaggtcagc 3120

caagaagtca gggtctgccc tggtcgcggc tacatcctca gggtgaccgc ctacaaggaa 3180

gggtatggtg aagggtgtgt cacgatccat gagattgaga acaacacgga cgaactcaag 3240

ttcagcaact gcgtcgagga ggaggtctat ccgaacaaca cggtcacctg caacgactac 3300

actgccactc aggaggagta cgagggaacc tacactagcc gcaaccgtgg ctacgatgga 3360

gcctacgagt ccaactcctc cgttcccgct gactacgcaa gcgcctacga ggagaaggcg 3420

tacacggacg ggcggcgcga caacccgtgc gagtccaatc gtggctacgg tgactacact 3480

ccgcttcccg ctggttacgt gaccaaggaa cttgagtact tccctgagac cgacaaggtc 3540

tggatcgaga taggtgagac cgaaggcacg ttcatcgttg actccgtcga actgctgctc 3600

atggaggagt ga 3612

<210> 128
<211> 3432
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC848.

<400> 128
atggagatca acaaccagaa gcagtgcatc ccgtacaact gcctgagcaa ccctgaggag 60

gtgctcctag acggtgagcg catcctccct gacatcgacc cgctggaggt gagcctgagc 120

ctcctccagt tcctgctgaa caacttcgtg cccggcggtg gcttcatctc cgggctcgtg 180

gacaagatct ggggcgcgct ccgcccgtcc gagtgggatc tcttcctcgc ccagatcgag 240

cgcctcatcg accagcgcat cgaggccacc gtgcgcgcca aggccatcac cgagctggag 300

ggcctgggcc gcaactacca aatctacgcc gaggcgttca aggagtggga gtccgacccg 360

gacaacgagg ccgctaagtc ccgcgtgatc gaccgcttcc gcatcctcga cggcctcatc 420

gaggccaaca tcccgtcctt ccgcatcatt ggcttcgagg tgccgctgct gtccgtgtac 480

gtgcaagccg ctaacctcca cctcgccctc ctgcgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactaagaa cgtgaacgac atctacaacc gccagatccg cgagatccac 600

gaatactcca accactgcgt ggacacctac aacaccgagc tggagcgcct cggcttccgc 660

tccatcgccc agtggcgcat ctacaaccag ttccgccgcg agctgaccct gaccgtcctg 720

gacatcgtgg ccctgttccc gaactacgac agccgcctct acccgatcca gaccttctcc 780

cagctcaccc gcgagatcgt gaccagcccg gtgtccgagt tctactatgg cgtcatcaac 840

tccggcaaca tcattggcac cctcaccgag cagcaaatcc gccgcccgca cctgatggac 900

ttcttcaact ccatgattat gtacacctcc gacaaccgcc gcgagcacta ctggtccggc 960

ctggagatga ccgcctactt caccggcttc gcgggcgcgc aagtgtcctt cccgctcgtg 1020

ggcaccaggg gcgagtccgc gccgcctctc accgtgaggt ccgtgaacga cggcatctac 1080

cgcatcctca gcgcgccgtt ctacagcgcg ccgttcctgg gcaccatcgt gctgggcagc 1140

cgtggcgaga agttcgactt cgccctgaac aacatcagcc ctccgcctag caccatctac 1200

aggcaccctg gcaccgtgga cagcctggtg agcatccctc ctcaagacaa cagcgtgcct 1260

ccgcacaggg gcagctctca caggctgagc cacgtgacca tgagggccag ctctcctatc 1320

ttccactgga cccacaggag cgccaccact accaacacca tcaaccctaa cgctatcatt 1380

cagatcccgc tggtgaagtc caccaacctc ggctccggca cctccgtggt gaagggtccg 1440

ggcttcactg gcggcgacat cctccgccgc accagccctg gccaaatctc caccctccgt 1500

gtgaacatca ctgctccact gtcccaacgc tatcgcgtac ggattcgtta cgcatccacc 1560

accaacctcc agttccacac ctccattgac ggccgaccga tcaaccaggg caacttctcc 1620

gctacgatgt cctccgggtc caatctccag tctgggagct ttcgcactgt cggcttcacc 1680

actccgttca acttctccaa tggctcctcc gtgttcacac tttcggcgca cgtcttcaac 1740

tctggcaacg aagtgtacat cgacaggatc gagtttgttc ccgcagaggt gaccttcgaa 1800

gccgagtacg atctcgagcg tgcccagaag gcagtgaatg agctgttcac gagcagcaat 1860

cagatcgggc tcaagaccga cgtgaccgac taccacatcg accaagtcag caatctggtg 1920

gagtgtctca gcgacgagtt ctgcctggac gagaagaagg aactcagcga gaaggtgaag 1980

cacgccaaac gtctgtccga cgaacgcaat ctgctgcaag atccgaactt cagagggatc 2040

aacaggcaac tggaccgtgg gtggcgtggg tccaccgaca tcaccatcca aggcggcgac 2100

gacgtcttca aggagaacta cgtcactctg ctgggcacct tcgacgagtg ctatccgacc 2160

tacctgtacc agaagatcga cgagagcaag ctcaaggcgt acacccgcta ccaactcaga 2220

ggctacatcg aagactccca ggatctggaa atctacctca ttcgctacaa cgcgaagcac 2280

gagactgtca acgtgcccgg cactggcagc ctgtggccgc tctccgcgcc gagccctatc 2340

ggaaagtgtg cgcaccactc ccaccacttc agcttggaca tcgacgtggg ttgcacggac 2400

ctcaacgaag acctgggtgt ctgggtgatc ttcaagatca agacgcaaga cggacacgcg 2460

agactcggca atctggagtt cctggaggag aagcctctgg tcggcgaagc tctcgccagg 2520

gtgaagcgcg ccgagaagaa gtggcgcgac aagcgggaga agctggagtg ggagacgaac 2580

atcgtgtaca aggaggccaa ggagtccgtg gacgccctct ttgtgaacag ccagtacgac 2640

cgcctccaag cagacaccaa cattgccatg atccacgccg ccgacaagcg cgtgcactcc 2700

atcagggagg cgtatctccc tgaactcagc gtgattcctg gcgtgaatgc tgccatcttc 2760

gaggagcttg agggtcgcat cttcaccgca ttcagcctct acgacgccag gaacgtcatc 2820

aagaacggag acttcaacaa cgggctgtcc tgctggaacg tcaaaggtca cgttgacgtg 2880

gaggagcaga acaaccatcg cagcgtcctc gtcgtccctg agtgggaagc tgaggtcagc 2940

caagaagtca gggtctgccc tggtcgcggc tacatcctca gggtgaccgc ctacaaggaa 3000

gggtatggtg aagggtgtgt cacgatccat gagattgaga acaacacgga cgaactcaag 3060

ttcagcaact gcgtcgagga ggaggtctat ccgaacaaca cggtcacctg caacgactac 3120

actgccactc aggaggagta cgagggaacc tacactagcc gcaaccgtgg ctacgatgga 3180

gcctacgagt ccaactcctc cgttcccgct gactacgcaa gcgcctacga ggagaaggcg 3240

tacacggacg ggcggcgcga caacccgtgc gagtccaatc gtggctacgg tgactacact 3300

ccgcttcccg ctggttacgt gaccaaggaa cttgagtact tccctgagac cgacaaggtc 3360

tggatcgaga taggtgagac cgaaggcacg ttcatcgttg actccgtcga actgctgctc 3420

atggaggagt ga 3432

<210> 129
<211> 3429
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC849.

<400> 129
atggagatca acaaccagaa gcagtgcatc ccgtacaact gcctgagcaa ccctgaggag 60

gtgctcctag acggtgagcg catcctccct gacatcgacc cgctggaggt gagcctgagc 120

ctcctccagt tcctgctgaa caacttcgtg cccggcggtg gcttcatctc cgggctcgtg 180

gacaagatct ggggcgcgct ccgcccgtcc gagtgggatc tcttcctcgc ccagatcgag 240

cgcctcatcg accagcgcat cgaggccacc gtgcgcgcca aggccatcac cgagctggag 300

ggcctgggcc gcaactacca aatctacgcc gaggcgttca aggagtggga gtccgacccg 360

gacaacgagg ccgctaagtc ccgcgtgatc gaccgcttcc gcatcctcga cggcctcatc 420

gaggccaaca tcccgtcctt ccgcatcatt ggcttcgagg tgccgctgct gtccgtgtac 480

gtgcaagccg ctaacctcca cctcgccctc ctgcgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactaagaa cgtgaacgac atctacaacc gccagatccg cgagatccac 600

gaatactcca accactgcgt ggacacctac aacaccgagc tggagcgcct cggcttccgc 660

tccatcgccc agtggcgcat ctacaaccag ttccgccgcg agctgaccct gaccgtcctg 720

gacatcgtgg ccctgttccc gaactacgac agccgcctct acccgatcca gaccttctcc 780

cagctcaccc gcgagatcgt gaccagcccg gtgtccgagt tctactatgg cgtcatcaac 840

tccggcaaca tcattggcac cctcaccgag cagcaaatcc gccgcccgca cctgatggac 900

ttcttcaact ccatgattat gtacacctcc gacaaccgcc gcgagcacta ctggtccggc 960

ctggagatga ccgcctactt caccggcttc gcgggcgcgc aagtgtcctt cccgctcgtg 1020

ggcaccaggg gcgagtccgc gccgcctctc accgtgaggt ccgtgaacga cggcatctac 1080

cgcatcctca gcgcgccgtt ctacagcgcg ccgttcctgg gcaccatcgt gctgggcagc 1140

cgtggcgaga agttcgactt cgccctgaac aacatcagcc ctccgcctag caccatctac 1200

aggcaccctg gcaccgtgga cagcctggtg agcatccctc ctcaagacaa cagcgtgcct 1260

ccgcacaggg gcagctctca caggctgagc cacgtgacca tgagggccag ctctcctatc 1320

ttccactgga cccacaggag cgccaccact accaacacca tcaaccctaa cgctatcatt 1380

cagatcccgc tggtgaaggc gttcaacctg cacagcggtg ccaccgtggt gcgcggacct 1440

ggcttcaccg gcggcgacat ccttcgccgc accaacaccg gcacctttgg tgacattcgc 1500

ctcaacatca acgtgccact gtcccaacgc tatcgcgttc gtattcgtta cgcctccacc 1560

accgatcttc agttcttcac tcgcatcaat ggcaccaccg tgaacattgg gaacttctct 1620

cgaaccatga accgtggtga caacctggag taccgcagct tccggactgc tggcttcagc 1680

acgccgttca acttcctcaa cgcccagtcc accttcactc taggtgcaca gtccttctcc 1740

aaccaggaag tgtacatcga ccgcgttgag tttgtgcccg ctgaggtcac gttcgaagcg 1800

gagtacgacc tcgagcgtgc ccagaaggca gtgaatgagc tgttcacgag cagcaatcag 1860

atcgggctca agaccgacgt gaccgactac cacatcgacc aagtcagcaa tctggtggag 1920

tgtctcagcg acgagttctg cctggacgag aagaaggaac tcagcgagaa ggtgaagcac 1980

gccaaacgtc tgtccgacga acgcaatctg ctgcaagatc cgaacttcag agggatcaac 2040

aggcaactgg accgtgggtg gcgtgggtcc accgacatca ccatccaagg cggcgacgac 2100

gtcttcaagg agaactacgt cactctgctg ggcaccttcg acgagtgcta tccgacctac 2160

ctgtaccaga agatcgacga gagcaagctc aaggcgtaca cccgctacca actcagaggc 2220

tacatcgaag actcccagga tctggaaatc tacctcattc gctacaacgc gaagcacgag 2280

actgtcaacg tgcccggcac tggcagcctg tggccgctct ccgcgccgag ccctatcgga 2340

aagtgtgcgc accactccca ccacttcagc ttggacatcg acgtgggttg cacggacctc 2400

aacgaagacc tgggtgtctg ggtgatcttc aagatcaaga cgcaagacgg acacgcgaga 2460

ctcggcaatc tggagttcct ggaggagaag cctctggtcg gcgaagctct cgccagggtg 2520

aagcgcgccg agaagaagtg gcgcgacaag cgggagaagc tggagtggga gacgaacatc 2580

gtgtacaagg aggccaagga gtccgtggac gccctctttg tgaacagcca gtacgaccgc 2640

ctccaagcag acaccaacat tgccatgatc cacgccgccg acaagcgcgt gcactccatc 2700

agggaggcgt atctccctga actcagcgtg attcctggcg tgaatgctgc catcttcgag 2760

gagcttgagg gtcgcatctt caccgcattc agcctctacg acgccaggaa cgtcatcaag 2820

aacggagact tcaacaacgg gctgtcctgc tggaacgtca aaggtcacgt tgacgtggag 2880

gagcagaaca accatcgcag cgtcctcgtc gtccctgagt gggaagctga ggtcagccaa 2940

gaagtcaggg tctgccctgg tcgcggctac atcctcaggg tgaccgccta caaggaaggg 3000

tatggtgaag ggtgtgtcac gatccatgag attgagaaca acacggacga actcaagttc 3060

agcaactgcg tcgaggagga ggtctatccg aacaacacgg tcacctgcaa cgactacact 3120

gccactcagg aggagtacga gggaacctac actagccgca accgtggcta cgatggagcc 3180

tacgagtcca actcctccgt tcccgctgac tacgcaagcg cctacgagga gaaggcgtac 3240

acggacgggc ggcgcgacaa cccgtgcgag tccaatcgtg gctacggtga ctacactccg 3300

cttcccgctg gttacgtgac caaggaactt gagtacttcc ctgagaccga caaggtctgg 3360

atcgagatag gtgagaccga aggcacgttc atcgttgact ccgtcgaact gctgctcatg 3420

gaggagtga 3429

<210> 130
<211> 3570
<212> ADN
<213> Nhân tạo

<220>
<223> Trình tự ADN tổng hợp để biểu hiện ở thực vật ghi mã TIC847.

<400> 130
atgaccagca accggaagaa cgagaacgag atcatcaacg ccctgagcat cccgaccgtg 60

agcaacccta gcacccagat gaacctgagc cctgacgctc gcatcgagga ctccctctgc 120

gtggctgagg tgaacaacat cgacccgttc gtgtccgcct ccaccgtgca gaccggcatc 180

aacatcgcgg gccgcatcct cggcgtgctc ggcgtgccct ttgcgggcca gctcgcctcc 240

ttctactcct tcctcgtggg agagctgtgg ccctccggcc gcgacccgtg ggagatcttc 300

ctggagcacg tggagcagct catccgccag caagtcaccg agaacacccg caacaccgcc 360

atcgcccgcc tggagggcct gggccgtggc taccgctcct accagcaagc cctggagacc 420

tggctcgaca accgcaacga cgcccgctcc cgctccatca tcctggagcg ctacgtcgcc 480

ctggaactgg acatcaccac tgccatccca ctcttccgca tcaggaacga ggaggtgcct 540

ctgctgatgg tgtacgccca ggctgcgaac ctgcacctgc tgctgctgcg cgacgcaagc 600

ctgtttggct ccgagtgggg tatggcaagc tccgacgtca accagtacta ccaggagcag 660

atccgctaca ccgaggagta cagcaaccac tgcgtccagt ggtacaacac cggtctgaac 720

aatctcagag ggaccaacgc tgagagctgg ctgcgctaca accagttccg gcgggatctg 780

accctaggtg tcctggatct ggtcgctctg ttcccgagct acgataccag gacgtaccct 840

atcaacacct ctgctcagct taccagggag atctacactg atcctatcgg taggactaac 900

gctcctagtg gtttcgccag cactaactgg ttcaacaaca acgcgcctag tttctctgcc 960

atcgaggcgg cgatcttccg gcctcctcac ctcctcgact tcccggagca gcttactatc 1020

tactctgcgt cttcgcggtg gtcttcgact cagcacatga actactgggt tggtcaccgg 1080

cttaacttcc gcccgattgg aggaactctt aacaccagta cgcaaggtct tacgaacaac 1140

acttccatca acccggttac gttgcagttc acgtctcggg acgtttaccg gacggagtcg 1200

aatgctggga cgaacatcct gttcacgaca ccggtgaatg gtgttccgtg ggcacgtttc 1260

aacttcatca acccgcagaa catctacgag cgtggagcaa cgacatactc gcaaccatac 1320

caaggcgttg gcatccaact gtttgactcg gagacggaac tgccaccaga gacgacagaa 1380

cgtccgaatt acgagtcata ctcacacaga ctatcacaca ttggactcat tatcggaaac 1440

acactgagag caccagtgta ctcatggaca catcggtcag cagatcgtac gaacaccatc 1500

ggtcctaacc gtatcaccca gattcctctc gtgaaggccc acaccctcca gtccggcacc 1560

accgtggtga agggaccggg cttcaccggc ggcgacatcc tccgccgcac ctctggcggc 1620

ccgttcgcgt tctccaacgt gaacctcgac ttcaacctct cccaacgcta cagagcacgc 1680

atacgctacg cctccaccac caacctacgc atctacgtga ccgtggctgg cgaacgcatc 1740

ttcgctgggc agtttgacaa gactatggac gctggcgcac ctcttacctt ccagtccttc 1800

tcctacgcca ccatcaacac cgcgttcacc ttcccagagc gctcctcctc cttgaccgtg 1860

ggtgccgaca ccttcagcag cggcaatgag gtgtacgtgg atcgcttcga actcatccca 1920

gtgaccgcca ccttcgaggc cgagtcggat ctcgagcgtg cccagaaggc agtgaatgag 1980

ctgttcacga gcagcaatca gatcgggctc aagaccgacg tgaccgacta ccacatcgac 2040

caagtcagca atctggtgga gtgtctcagc gacgagttct gcctggacga gaagaaggaa 2100

ctcagcgaga aggtgaagca cgccaaacgt ctgtccgacg aacgcaatct gctgcaagat 2160

ccgaacttca gagggatcaa caggcaactg gaccgtgggt ggcgtgggtc caccgacatc 2220

accatccaag gcggcgacga cgtcttcaag gagaactacg tcactctgct gggcaccttc 2280

gacgagtgct atccgaccta cctgtaccag aagatcgacg agagcaagct caaggcgtac 2340

acccgctacc aactcagagg ctacatcgaa gactcccagg atctggaaat ctacctcatt 2400

cgctacaacg cgaagcacga gactgtcaac gtgcccggca ctggcagcct gtggccgctc 2460

tccgcgccga gccctatcgg aaagtgtgcg caccactccc accacttcag cttggacatc 2520

gacgtgggtt gcacggacct caacgaagac ctgggtgtct gggtgatctt caagatcaag 2580

acgcaagacg gacacgcgag actcggcaat ctggagttcc tggaggagaa gcctctggtc 2640

ggcgaagctc tcgccagggt gaagcgcgcc gagaagaagt ggcgcgacaa gcgggagaag 2700

ctggagtggg agacgaacat cgtgtacaag gaggccaagg agtccgtgga cgccctcttt 2760

gtgaacagcc agtacgaccg cctccaagca gacaccaaca ttgccatgat ccacgccgcc 2820

gacaagcgcg tgcactccat cagggaggcg tatctccctg aactcagcgt gattcctggc 2880

gtgaatgctg ccatcttcga ggagcttgag ggtcgcatct tcaccgcatt cagcctctac 2940

gacgccagga acgtcatcaa gaacggagac ttcaacaacg ggctgtcctg ctggaacgtc 3000

aaaggtcacg ttgacgtgga ggagcagaac aaccatcgca gcgtcctcgt cgtccctgag 3060

tgggaagctg aggtcagcca agaagtcagg gtctgccctg gtcgcggcta catcctcagg 3120

gtgaccgcct acaaggaagg gtatggtgaa gggtgtgtca cgatccatga gattgagaac 3180

aacacggacg aactcaagtt cagcaactgc gtcgaggagg aggtctatcc gaacaacacg 3240

gtcacctgca acgactacac tgccactcag gaggagtacg agggaaccta cactagccgc 3300

aaccgtggct acgatggagc ctacgagtcc aactcctccg ttcccgctga ctacgcaagc 3360

gcctacgagg agaaggcgta cacggacggg cggcgcgaca acccgtgcga gtccaatcgt 3420

ggctacggtg actacactcc gcttcccgct ggttacgtga ccaaggaact tgagtacttc 3480

cctgagaccg acaaggtctg gatcgagata ggtgagaccg aaggcacgtt catcgttgac 3540

tccgtcgaac tgctgctcat ggaggagtga 3570