



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

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



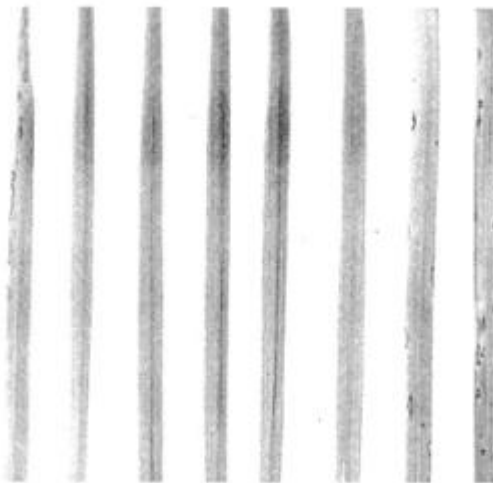
1-0028375

(51)⁷ C12N 15/29; C07K 14/415; C12N 15/82; (13) B
A01H 5/00; C12N 15/10

-
- (21) 1-2017-00426 (22) 21/12/2015
(86) PCT/CN2015/098148 21/12/2015 (87) WO 2016/101859 30/06/2016
(30) 201410822942.8 22/12/2014 CN
(45) 25/05/2021 398 (43) 25/10/2017 355A
(73) Plant Protection Research Institute of Guangdong Academy of Agricultural Sciences (CN)
No.7 Jinying Road, Tianhe District, Guangzhou City, Guangdong Province 510640, P. R. China
(72) Jing SU (CN); Xiaoyuan ZHU (CN); Shen CHEN (CN); Wenjuan WANG (CN); Congying WANG (CN); Liexian Zeng (CN); Jianyuan YANG (CN).
(74) Công ty TNHH Sáng chế ACTIP (ACTIP PATENT LIMITED)
-

(54) GEN KHÁNG BỆNH ĐẠO ÔN Pi50, PHƯƠNG PHÁP TẠO DÒNG GEN VÀ TRÌNH TỰ PROTEIN ĐƯỢC MÃ HÓA BỞI GEN NÀY

(57) Sáng chế đề cập đến gen kháng bệnh đạo ôn Pi50, phương pháp tạo dòng, và trình tự protein được mã hóa bởi gen này. Trình tự nucleotit của gen kháng bệnh đạo ôn Pi50 được thể hiện dưới dạng trình tự SEQ ID NO.1. Gen kháng bệnh đạo ôn Pi50 có thể thu được bằng phương pháp tổng hợp hóa học hoặc thu được bằng cách thiết kế các cặp mồi bằng cách sử dụng ADN thuộc hệ gen của giống lúa EBZ làm khuôn, sau đó thực hiện khuếch đại gen bằng phản ứng chuỗi trùng hợp (PCR - Polymerase Chain Reaction). Gen kháng bệnh đạo ôn Pi50 có hiệu quả kháng bệnh nổi trội, có thể được sử dụng để chọn/tạo giống lúa có khả năng kháng bệnh đạo ôn.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến gen kháng bệnh đạo ôn, cụ thể là gen kháng bệnh đạo ôn Pi50, phương pháp tạo dòng, và trình tự protein được mã hóa bởi gen này.

Tình trạng kỹ thuật của sáng chế

Lúa gạo là một trong số các loại cây trồng thực phẩm quan trọng nhất, hơn một nửa dân số trên thế giới sử dụng gạo làm nguồn lương thực chính, tuy nhiên, bệnh đạo ôn gây ra bởi chủng nấm gây bệnh đạo ôn (*Magnapothae oryzae*) đã ảnh hưởng mạnh đến quá trình sản xuất lúa an toàn, và làm giảm nghiêm trọng sản lượng lúa hàng năm. Từ quan điểm xây dựng một nền nông nghiệp phát triển bền vững, việc tạo giống và sử dụng các giống lúa có khả năng kháng bệnh là phương pháp thân thiện với môi trường, an toàn và hiệu quả nhất để phòng ngừa và điều trị bệnh đạo ôn. Tuy nhiên, do các chủng *Magnaporthae oryzae* đa dạng và biến đổi không ngừng, cùng với sự hiểu biết chưa đầy đủ và ứng dụng chưa hiệu quả các gen kháng, dẫn đến chu kỳ kháng bệnh của các giống kháng đơn khá ngắn (thường kéo dài từ 2 đến 4 năm), đây là vấn đề gây nhức nhối nhất với các nhà khoa học chọn giống (Zheng Zhao, v.v., 2009). Do đó, việc phát triển, tạo dòng và sử dụng hợp lý các gen kháng bệnh có phổ rộng trở thành một phương thức quan trọng để thu được các giống kháng bệnh có phổ rộng và tác dụng kháng bệnh lâu dài, đây là một vấn đề quan trọng cần được ưu tiên giải quyết trong nghiên cứu nông nghiệp.

Đến nay, có hơn 80 gen kháng bệnh đạo ôn đã được xác định (theo Ballini et al., 2008; E Zhiguo, et al., 2009; Jiang et al., 2012), và có ít nhất 22 gen kháng bệnh đạo ôn đã được tách dòng (theo Wang et al., 1999; Qu et al., 2006; Hayashi et al., 2010; Jiang et al., 2012; Hua et al., 2012; Das et al., 2012). Các nghiên cứu đã chỉ ra rằng hầu hết các gen kháng bệnh đạo ôn đều là các protein có cấu trúc có chứa vùng vị trí liên kết nucleotit – vùng lặp lại giàu leucin (NBS-LRR), tức là, có vùng vị trí liên kết nucleotit (NBS) tại đầu N và vùng lặp lại giàu leucin (LRR) tại đầu C. Vùng NBS có thể liên quan đến việc dẫn truyền tín hiệu (theo Hammond-Kosack & Jones 1997),

vùng NBS có hoạt tính kết hợp ATP (adenosine triphosphate) hoặc GTP (Guanosine 5'-Triphosphate) với hydrolaza, có thể hoạt hóa enzym kinaza hoặc các protein G (Guanyl-Nucleotide Binding Proteins – Protein liên kết với nucleotit Guanyl) (theo Traut et al., 1994); chức năng chủ yếu của vùng LRR liên quan đến tương tác protein-protein (theo Jones & Jones, 1996; Dangl & Jones, 2001; Takken & Tameling, 2009; Bernoux, 2011), vùng này được cho là vị trí tương tác trực tiếp giữa các sản phẩm gen kháng với sản phẩm hoặc sản phẩm gián tiếp được mã hóa bởi gen không độc tính (gen Avr) có khả năng gây bệnh (theo Bent, 1996; Baker et al., 1997; Jia et al. 2000). Thử nghiệm lai nấm men hai pha chỉ ra rằng sự kết hợp giữa hợp chất kích thích được mã hóa bởi vùng AVR-Pita và vùng giàu leucin (Leucine Rich Domain – LRD) của thụ thể Pita có thể hoạt hóa phản ứng phòng vệ có trung gian là Pita trong tế bào, điều này trực tiếp chứng minh rằng vùng LRR có thể là vùng nhận biết của nấm gây bệnh.

Gen chứa vùng NBS-LRR có xu hướng sắp xếp theo nhóm, thông thường các gen trong cùng một nhóm có sự tương đồng cao, và duy trì nguyên tắc đồng di truyền với các gen có chức năng trong cùng con đường kháng (Michelmore & Meyers, 1998; Leister, 2004). Nhóm gen Pi2/9 nằm trên cánh dài của nhiễm sắc thể số 6 của lúa là vùng nóng trong nghiên cứu của các nhà khoa học, các gen có khả năng kháng bệnh đạo ôn đã được nghiên cứu chuyên sâu và báo cáo từ phạm vi nghiên cứu này bao gồm Pi2, Pi9, Pigm(t), Pi-z và Piz-t, các gen này được bố trí tại nhóm gen chứa từ 2-9 vùng NBS-LRR. Người ta đã chứng minh rằng các gen chứa một vùng NBS-LRR trong gen Pi2, Pi9 và Pi-z kiểm soát tác dụng kháng bệnh đạo ôn phổ rộng (Zhou, 2006, 2007; Dai, 2010), trong đó chỉ có 8 gốc axit amin khác nhau trong vùng LRR của gen Pi2 và Piz-t dẫn đến các kết quả và khả năng kháng bệnh khác nhau (Qu et al., 2006; Zhou et al., 2006). Nhiều gen kháng bệnh chứa vùng NBS-LRR khác trong phạm vi nghiên cứu này vẫn tiếp tục được phát triển và ứng dụng, và việc ứng dụng nhóm gen này có triển vọng cao.

Kỹ thuật tách dòng và sự am hiểu đầy đủ về các gen kháng trong phạm vi nghiên cứu này có thể góp phần kiểm soát tốt hơn và giảm các thiệt hại gây ra bởi chủng nấm *Magnaporthe oryzae*, và hiểu rõ hơn về cơ chế phân nhóm của gen kháng bệnh đạo ôn. Đồng thời, có thể mở rộng phổ kháng của cây trồng hoặc nhân giống nhằm đích chính

xác hơn bằng cách biến đổi gen hoặc chuyển gen tại các vị trí then chốt.

Bản chất kỹ thuật của sáng chế

Mục tiêu chính của sáng chế là khắc phục các nhược điểm và thiếu sót trong tình trạng kỹ thuật trước đó, và đề xuất gen kháng bệnh đạo ôn Pi50.

Mục tiêu khác của sáng chế là đề xuất phương pháp tạo dòng gen kháng bệnh đạo ôn Pi50.

Mục tiêu khác nữa của sáng chế là đề xuất protein được mã hóa bởi gen kháng bệnh đạo ôn Pi50.

Các mục tiêu của sáng chế đạt được thông qua các giải pháp kỹ thuật như sau:

gen kháng bệnh đạo ôn Pi50 có trình tự nucleotit như được thể hiện trong trình tự SEQ ID NO.1;

gen kháng bệnh đạo ôn Pi50 có chiều dài gồm 3.099 nucleotit, và có thể mã hóa cho protein gồm 1.032 axit amin;

phương pháp tạo dòng gen kháng bệnh đạo ôn Pi50, trong đó gen Pi 50 có thể thu được bằng phương pháp tổng hợp hóa học, hoặc thu được bằng cách thiết kế cặp mồi, nhờ sử dụng ADN của hệ gen thuộc giống lúa EBZ làm khuôn, sau đó thực hiện khuếch đại bằng phản ứng chuỗi trùng hợp (PCR - Polymerase Chain Reaction);

trình tự protein được mã hóa bởi gen kháng bệnh đạo ôn Pi 50 được thể hiện trong trình tự SEQ ID NO.2.

Trình tự protein được mã hóa bởi gen kháng bệnh đạo ôn Pi50 có tất cả 1.032 axit amin, bao gồm hai vùng chính: vùng vị trí liên kết nucleotit (NBS) và vùng lặp lại giàu leucin (LRR), trong đó vùng NBS (gồm 170-462 axit amin) chứa vùng có enzym kinaza AAA được bảo tồn, và vùng LRR (gồm 578-1032 axit amin) ở đầu C, có hàm lượng leucin chiếm 14,8% (67 leucin trong tổng số 454 axit amin).

Sáng chế có các lợi thế và hiệu quả đạt được so với tình trạng kỹ thuật trước đó như sau: gen kháng bệnh đạo ôn Pi50 được đề xuất theo sáng chế có thể được chuyển vào cây trồng dễ bị bệnh, góp phần tạo nên cây trồng mới có khả năng kháng bệnh. Cụ thể là, nhiều gen có khả năng kháng bệnh có thể được chuyển vào cây trồng bằng kỹ thuật chuyển gen mà không gặp các vấn đề về việc liên kết các gen ác tính trong hệ

gen thường xảy ra ở các kỹ thuật nhân giống truyền thống, và kỹ thuật chuyển gen này có thể rút ngắn thời gian nhân giống. Việc tách dòng gen có khả năng kháng bệnh là điều kiện tiên quyết để giải quyết vấn đề còn tồn tại trong kỹ thuật nhân giống truyền thống, đó là khó chuyển gen kháng bệnh giữa các loài thực vật. Sáng chế có thể còn đề xuất hoặc ứng dụng các cây trồng chuyển gen có khả năng kháng bệnh thu được dựa trên các đoạn ADN nêu trên và các hạt giống của nó, và các cây trồng được chuyển gen theo sáng chế hoặc tái tổ hợp dựa trên các gen này, hoặc các hạt giống của nó. Gen theo sáng chế có thể được chuyển vào trong các cây trồng khác nhờ lai hữu tính.

Mô tả vắn tắt các hình vẽ

Fig.1 là biểu đồ so sánh thể hiện phổ kháng của các alen thuộc nhóm Pi/2/9 có khả năng kháng lại các chủng nấm đạo ôn *Magnaporthe oryzae* khác nhau;

Fig.2 là biểu đồ thể hiện chiến lược tách dòng gen kháng bệnh đạo ôn Pi50;

Fig.3 là sơ đồ điện di trên gel agarosa thể hiện dấu hiệu nhận biết của các thể biến nạp bằng phương pháp PCR, trong đó các dải 1 và 12 là chỉ thị có trọng lượng phân tử 1 kb ladder, dải 2 là sản phẩm của quá trình PCR thu được bằng cách sử dụng plasmit pCambia-Pi50-N4 làm khuôn, dải 3 là sản phẩm của quá trình PCR thu được bằng cách sử dụng ADN của hệ gen thuộc giống Nipponbare chứa thụ thể biến nạp làm khuôn, dải 4 là sản phẩm của quá trình PCR thu được bằng cách sử dụng ADN của hệ gen thuộc giống lúa EBZ làm khuôn, các dải 5-11 là các sản phẩm của quá trình PCR thu được bằng cách sử dụng các ADN của hệ gen thuộc các thể biến nạp T_0 khác nhau làm khuôn; và

Fig.4 là sơ đồ thể hiện kết quả cấy ghép sau 7 ngày của giống lúa đối chứng và các thể biến nạp dương tính pCambia-Pi50 bởi chủng nấm gây bệnh đạo ôn *Magnaporthe oryzae* TY19, trong đó, từ trái sang phải, lá thứ nhất là giống dễ mắc bệnh Nipponbare, lá thứ hai là giống kháng bệnh đạo ôn EBZ, các lá từ thứ 3 đến thứ 6 là các cá thể được cấy chuyển dương tính pCambia-Pi50, lá thứ 7 là chủng kháng IRBLzt-T, và lá thứ 8 là chủng kháng IRBLz5-CA.

Mô tả chi tiết sáng chế

Sau đây, sáng chế được mô tả chi tiết kết hợp với các ví dụ và hình vẽ, tuy nhiên, các phương án theo sáng chế không nhằm giới hạn phạm vi của sáng chế. Trừ khi có các

chỉ định cụ thể trong tài liệu này, các phương pháp kỹ thuật được sử dụng trong phần ví dụ là các phương pháp thông thường được biết đến bởi những người có trình độ trung bình trong lĩnh vực kỹ thuật này. Theo các ví dụ của sáng chế, các đặc tính kháng, các đặc điểm về phân lập, tách dòng và chức năng, và định dạng của gen Pi50 đã được bộc lộ, và gen Pi50 đã phân lập được liên kết với vectơ thích hợp, sau đó chuyển gen vào trong cây trồng để tạo cho cây trồng có khả năng kháng bệnh nhất định.

Các công việc sơ bộ được thực hiện trong phòng thí nghiệm cụ thể như sau (tham khảo Zhu, et al. The identification of Pi50(t), a new member of the rice blast resistance Pi2/Pi9 multigene family. Theor Appl Genet, 2012(124):1295-1304):

(1) Đặc tính kháng của gen kháng Pi50

Để so sánh và xác định phổ kháng của gen Pi50 và 4 alen (Pi2, Pi9, Piz, Piz-t) thuộc vùng Pi2/9, tiến hành so sánh và phân tích phổ kháng của các giống mang 5 alen nêu trên (EBZ, IRBLz5-CA, C101A51, IRBL9-w, IRBLz-Fu, IRBLz-T và Toride, theo thứ tự) kháng lại tổng cộng 523 mẫu bệnh đạo ôn, trong đó có 85 mẫu bệnh đạo ôn phân lập được tại tỉnh Guangdong (Quảng Đông) và tỉnh Guangxi (Quảng Tây), 40 mẫu bệnh đạo ôn phân lập được tại tỉnh Fujian (Phúc Kiến), 41 mẫu bệnh đạo ôn phân lập được tại tỉnh Hunan (Hồ Nam), 60 mẫu bệnh đạo ôn phân lập được tại tỉnh Guizhou (Quý Châu), 66 mẫu bệnh đạo ôn phân lập được tại tỉnh Sichuan (Tứ Xuyên), 108 mẫu bệnh đạo ôn phân lập được tại tỉnh Liaoning (Liêu Ninh), 60 mẫu bệnh đạo ôn phân lập được tại tỉnh Jilin (Cát Lâm), và 63 mẫu bệnh đạo ôn phân lập được tại tỉnh Heilongjiang (Hắc Long Giang). Kết quả chỉ ra rằng gen Pi50 có khả năng kháng cao đối với hầu hết các mẫu bệnh đạo ôn phân lập được tại tỉnh Guangdong, tỉnh Guangxi, tỉnh Sichuan, tỉnh Fujian, tỉnh Heilongjiang, và tỉnh Jilin (với phổ kháng lên đến 97,7%), điều này chỉ ra rằng gen Pi50 có thể được sử dụng trong các khu vực nêu trên (Fig.1).

(2) Chiến lược tách dòng gen kháng bệnh đạo ôn Pi50

Các nghiên cứu xác định vị trí trước đó trong phòng thí nghiệm đã chỉ ra rằng gen Pi50 được định vị trong các vị trí giống nhau của hệ gen trên các gen nhóm Pi2/9; các đặc tính chỉ ra rằng các vùng NBS-LRR trong gen Pi50 được bố trí trong nhóm cũng có tính tương tự cao với các gen nhóm Pi2/9; và việc phân tích phổ kháng tiếp tục chỉ

ra rằng gen Pi50, tương tự với 4 alen thuộc nhóm Pi2/9, có phản ứng kháng tương tự nhau kháng lại hầu hết các mẫu bệnh đạo ôn thử nghiệm (theo Zhu et al., 2012); như được kết luận từ các kết quả tóm tắt từ ba khía cạnh trên đây, thì gen Pi50 cũng là một alen thuộc nhóm Pi2/9. Cần lưu ý rằng, gen Pi50 được định vị trong nhóm gen Pi2/9 khác với các gen Pi2, Pi9 và Piz-t ở chủng loài cụ thể tương ứng, trong đó gen Pi50 có thể chống lại sự xâm lấn của các mầm bệnh tương thích khác của nhóm Pi2/9, do đó biểu thị phổ kháng rộng hơn (Fig.1). Các vị trí đóng khung trong hình vẽ này thể hiện các mầm bệnh tương thích của mỗi kiểu gen tương ứng của nó, với các giá trị số biểu thị các mức độ nhạy cảm khác nhau, các số từ 1-3 thể hiện khả năng kháng cao, các số từ 4-6 thể hiện mức độ nhạy cảm vừa phải, và các số từ 7-9 thể hiện mức độ nhạy cảm cao với bệnh đạo ôn. Theo đó, sáng chế đã xây dựng được bản đồ vật lý điện tử của vùng đích Pi50. Điều này có nghĩa là, các chỉ thị liên kết với hoặc đồng tách biệt khỏi Pi50 được tích hợp với các vị trí vật lý bằng các cách thức tin sinh học trên cơ sở thông tin trình tự của giống lúa đối chứng Nipponbare, để xây dựng bản đồ các đoạn gen kế tiếp (contig map) của các vùng gen đích và bản đồ tích hợp vật lý điện tử (Fig.2).

Sau đó, bằng phương pháp khuếch đại gen nhờ phản ứng chuỗi trùng hợp “Long Range PCR” tiến hành trên nhiễm sắc thể, và kiểm tra các trình tự tham khảo của các kiểu gen tương đồng và các kiểu gen tương tự, chủ đơn thu được trình tự gen từ chủng 28zhan có khả năng kháng bệnh rất cao (giống lúa Er-Ba-Zhan (EBZ), ở vùng trồng lúa phía nam Trung Quốc, đã bộc lộ trong tài liệu "Zhu, et al. The identification of Pi50(t), a new member of the rice blast resistance *Pi2/Pi9* multigene family. Theor Appl Genet, 2012(124): 1295-1304"), trình tự gen này bao gồm hơn 100kb trong vùng vị trí gen Pi50, trong đó vùng này chứa hơn 9 gen mã hóa cho vùng NBS-LRR tương tự như protein theo dự đoán của FGenesh (dữ liệu chưa được công bố bởi chủ đơn). Theo dữ liệu này, các công việc dưới đây đã được thực hiện theo sáng chế.

Ví dụ thực hiện sáng chế

Ví dụ 1: Phân tích và xác định gen kháng bệnh đạo ôn Pi50

1. Mô tả và phân tích trình tự của các gen có khả năng chứa gen kháng bệnh đạo ôn Pi50

Để xác định các gen có khả năng chứa gen Pi50, tác giả sáng chế thực hiện dự đoán, mô tả, phân tích và xác định gen trên vùng đích bằng cách sử dụng trình tự tham khảo của giống lúa Nipponbare, và hai phần mềm dự đoán gen RiceGAAS (<http://ricegaas.dna.affrc.go.jp>) và FGENESH từ Softberry (<http://www.softberry.com>). Ở giống lúa Nipponbare, các gen có khả năng kháng trong các vùng Pi50 là 7 mẫu gen mã hóa vùng vị trí liên kết nucleotit và vùng lặp lại giàu leucin (NBS-LRR) (Pi501-NP, Pi502-NP, Pi503-NP, Pi504-NP, Pi505-NP, Pi506-NP và Pi507-NP), trong đó Pi501-NP và Pi507-NP được bảo tồn cao trong tất cả các giống đã báo cáo có chứa họ Pi2/9, hầu như không có bất kỳ sự khác biệt nào. Tuy nhiên, 5 gen còn lại (Pi502-NP, Pi503-NP, Pi504-NP, Pi505-NP và Pi506-NP) có sự khác nhau giữa các mẫu, chúng là các gen giả do cấu trúc không hoàn chỉnh của chúng dựa trên việc phân tích về tính toàn vẹn của gen và sự có mặt/vắng mặt của dấu hiệu chỉ thị cụ thể. Bằng cách tiếp tục so sánh với các cấu trúc của các gen thuộc nhóm Pi2/9, các tác giả sáng chế đã phát hiện ra rằng có một đoạn khuyết có chiều dài khoảng 20 kb trong giống lúa Nipponbare (giữa Pi502-NP và Pi503-NP), và đoạn khuyết này tương ứng chính xác với đoạn chức năng của gen Pi2/z-t. Điều này có thể giải thích cho câu hỏi tại sao giống lúa Nipponbare dễ mắc bệnh, liên quan đến các chủng kháng thuộc nhóm Pi2/9. Tương tự, có 8-9 gen có thể chứa vùng NBS-LRR ở vị trí gen Pi50 thuộc hệ gen của giống lúa EBZ, trong đó có 5 trường hợp dự đoán tương ứng với giống lúa Nipponbare cũng có cấu trúc không hoàn chỉnh. Do đó, có thể suy ra rằng gen kháng có khả năng chứa Pi50 có thể hiện diện trong đoạn gen bị khuyết của giống lúa dễ mắc bệnh Nipponbare.

Các phân tích tiếp theo chỉ ra rằng gen có trong giống lúa EBZ, và không có trong giống lúa Nipponbare, có cấu trúc bảo tồn vùng NBS-LRR hoàn chỉnh và có thể được biểu hiện, được đặt tên là Pi50-N4 theo vị trí của nó trong nhóm gen, trong đó Pi50-N4 có độ dài trình tự là 7063bp trong hệ gen, như được thể hiện bằng trình tự SEQ ID NO.3.

2. Thu cADN của gen Pi50-N4

Quá trình phiên mã ngược được thực hiện bằng kỹ thuật Realtime PCR (RT-PCR), sử dụng enzym phiên mã ngược SuperScript III (từ công ty Invitrogen, Mỹ) và

Oligo(dT)18, với hệ thống phiên mã ngược cũng như các điều kiện phản ứng theo chỉ dẫn đối với enzym này.

Sau đó thực hiện phản ứng khuếch đại PCR sử dụng các cặp mồi như sau:

mồi ngược F1: 5'- ATGGCGGAGACGGTGCTGAGC-3';

mồi xuôi R1: 5'- TCAGCCAGCTTGAGCTGTGC-3'.

Trình tự cADN thu được của Pi50- N4 có 3.099bp.

3. Phân tích trình tự

Bằng cách so sánh trình tự gen (SEQ ID NO.3) với trình tự cADN (SEQ ID NO.1) của Pi50-N4, kết quả cho thấy rằng trình tự gen của Pi50-N4 chứa hai intron, với exon thứ nhất ở vị trí 1-116, exon thứ hai ở vị trí 3953-6904 và exon thứ ba ở vị trí 7033-7063. Trình tự cDNA của gen được bảo tồn cao với alen đã tách dòng Pi2 (tham khảo GenBank: DQ352453.1, ở mục sản xuất gen “NBS-LRR type R protein, Nbs4-Pi”) và Pi-zt (tham khảo GenBank: DQ352040.1) trong nhóm Pi2/9, với độ tương đồng là 98% hoặc cao hơn. Mặc dù có nhiều liên kết trình tự, nhưng có thể thấy rằng có sự khác biệt giữa các gen tập trung chủ yếu ở vùng LRR gần đầu 3', chỉ có 8 axit amin khác nhau giữa Pi2 và Pi-z, tức là, 8 axit amin khác nhau này dẫn đến hiệu quả kháng bệnh khác nhau giữa hai gen (Zhou et al., 2006. The Eight Amino-Acid Differences Within Three Leucine-Rich Repeats Between Pi2 and Piz-t Resistance Proteins Determine the Resistance Specificity to Rice blast bacteria. MPMI, 19: 1216–1228), và còn có 10 axit amin khác nhau giữa Pi50-N4 (trình tự axit amin của protein đã mã hóa được thể hiện bằng trình tự SEQ ID NO.2) và hai gen nêu trên trong vùng này.

Ví dụ 2. Thử nghiệm hỗ trợ di truyền của gen kháng bệnh đạo ôn Pi50 và xác định thể biến nạp

1. Thu gen Pi50-N4

Gen di truyền chứa Pi50-N4 được thu bằng phương pháp khuếch đại nhờ sử dụng hệ gen của giống lúa EBZ làm khuôn và sử dụng các cặp mồi như sau:

mồi xuôi F2: 5'- ACGAATTCGAGCTCGGTACC

CTTGACATCCAAACCGCACC -3';

mỗi ngược R2: 5'- AGTGCCAAGCTTCGGCGCGCC
TCAGCCAGCTTGAGCTGTGC -3'.

Hệ thống phản ứng PCR như sau: enzym có độ tin cậy cao KOD *plus neo*: 0,4μl, mẫu gen của giống lúa EBZ: 1,2μl (100pg~50ng), Dimetyl sulfoxit (DMSO): 0,5μl, mỗi xuôi F2 (10 μM): 0,6μl, mỗi ngược R2 (10μM): 0,6μl, dNTP (2mM): 4μl, 2 lần đệm, mỗi lần 10μl, nước cất được cất hai lần với lượng lên đến 20μl.

Các điều kiện về chu trình nhiệt độ của quá trình PCR là: 94°C trong 2 phút; 98°C trong 10 giây, 56°C trong 30 giây, 68°C trong 9 phút, và 68°C trong thời gian mở rộng với gradient mở rộng 10 giây trên mỗi vòng cho chu kì chạy 5 vòng; 98°C trong 10 giây, 68°C trong 10 phút, 68°C trong thời gian mở rộng với gradient mở rộng 10 giây trên mỗi vòng cho chu kì chạy 32 vòng; 68°C trong thời gian mở rộng 15 phút; bảo quản ở nhiệt độ 12°C.

Các sản phẩm PCR thu được được chạy điện di trên gel agarosa 0,8%, và gen di truyền chứa Pi50-N4 thu được bằng cách sử dụng kit tách chiết gel ADN. ADN được định lượng gần đúng bằng kỹ thuật điện di.

2. Xây dựng và biến nạp vector biểu hiện

(1) Sự tuyến tính hóa vector: vector pCAMBIA1300AscI được phân cắt hai lần bằng enzym giới hạn endonucleaza Kpn I và Asc I (theo phương pháp được đề xuất trong tài liệu Liu Xinqiong, et. al, “Binary Vector Construction of Candidate Genes of the Blast Resistance Gene Pi36”, Tạp chí Khoa học nông nghiệp Hubei, tập 2, 2007) trong đó: đệm phân cắt enzym (Cutsmart): 2μl, mỗi loại enzym Kpn I và Asc I: 0,5μl, gen Pi50-N4: 3μl, bổ sung nước được cất hai lần đến thể tích 20μl; ủ ở 37°C qua đêm để phân hủy enzym; điện di trên gel agarosa 0,8%, khôi phục ADN, định lượng gần đúng ADN bằng kỹ thuật điện di.

(2) Phản ứng In-fusion: gen Pi50-N4 thu được từ bước (1): 350 ng, vector pCambia 1300AscI đã tuyến tính hóa: 200 ng, 5×In-Fusion HD 2μl, bổ sung nước chung cất hai lần lên đến thể tích 10μl. Thực hiện phản ứng theo hướng dẫn sử dụng

của bộ tách dòng In-Fusion HD để thu được chủng *Escherichia coli* JM109 được biến nạp với vector tái tổ hợp pCambia-Pi50.

(3) phát triển và nuôi cấy chủng *E.coli* JM109 được biến nạp bằng vector tái tổ hợp pCambia-Pi50, tách chiết plasmit pCambia-Pi50, giải trình tự, và nhận biết gen di truyền Pi50-N4 mở rộng về phía đầu 5' như được thể hiện trong trình tự SEQ ID NO.4.

(4) theo phương pháp nêu trong "Hiei, et al. 1994 Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J. 6:271–282", plasmit pCambia-Pi50 được biến nạp vào trong các tế bào khả biến EHA105 của vi khuẩn *agrobacterium*, sau đó được chuyển vào giống lúa dễ bị mắc bệnh, Nipponbare (có sẵn ở thư viện tài nguyên chất mầm quốc tế thuộc viện nghiên cứu lúa quốc tế (<http://www.irri.org>)), được sàng lọc bằng hygromycin B, để thu được các thể biến nạp.

Xác định các thể biến nạp thu được: ADN của các thể biến nạp (dải) được tách chiết bằng kit tách chiết ADN thực vật. Thử nghiệm PCR được thực hiện trên các mẫu ADN tách chiết được bằng cách sử dụng một cặp mồi gồm mồi đặc trưng F3 ở đầu 3' của gen và mồi R3 ở vị trí tạo dòng của vector. Các kết quả cho thấy rằng đã thu được 28 thực vật chuyển gen dương tính, và ngoài ra, 19 thực vật là thực vật chuyển gen âm tính, và điện di đồ định danh một phần các thể biến nạp được thể hiện trên Fig.3.

mồi xuôi F3: 5'-GGTATATGGATAGCAGAAGG-3';

mồi ngược R3: 5'- GCGATTAAGTTGGGTAACGC-3'.

Thử nghiệm cấy nguyên bào nhân tạo được thực hiện trên cây con chứa các dòng chuyển gen Pi50 và giống đối chứng (theo phương pháp được đề xuất trong Zhang Lei et al., "Resistances of Four Sets of Rice Varieties to 19 Strains of *Magnaporthe grisea* in Jiangsu Area", Jiangsu Journal of Agricultural Science, issue 5, 2010) bằng cách sử dụng chủng TY 19 không độc hại với Pi50 (Zhu Xiaoyuan, et al., "Race identification and pathogenicity test of the blast fungus causing the resistance breakdown of hybrid rice Tianyou 998", Guangdong Agricultural Science, 2008, 12: 84 -86). Các kết quả được thể hiện trong bảng 1 và trên Fig.4.

Bảng 1: Các kết quả phân tích của các thử nghiệm hỗ trợ di truyền

Gen thí nghiệm			Số lượng cây T ₀ và các kiểu phản ứng với bệnh (R:S) ^a		Bổ trợ chức năng
Sản phẩm thiết kế (ADN)	Giống thụ thể	Số cây thử nghiệm	Số cây dương tính với chuyển gen	Số cây âm tính với chuyển gen	Tỉ lệ thành công (%) ^b
Pi50-N4	Nipponbare	47	28 (25:3)	19 (2:17)	89,29%

^a kiểu phản ứng kháng bệnh hoặc kiểu phản ứng dễ mắc bệnh của mỗi loại cây T₀ sau khi được ghép chủng Pi50-avirulent TY 19. R, kháng bệnh; S, dễ mắc bệnh.

^b Tỉ lệ thành công bổ trợ chức năng (%): tỉ lệ phần trăm của số cây có khả năng kháng bệnh R so với số cây dương tính với chuyển gen.

Từ kết quả nêu trong bảng 1, có thể thấy rằng có 25 cây trong số 28 dương tính với chuyển gen biểu thị khả năng kháng bệnh. Các lá trong nhóm đối chứng và một phần các cây chuyển gen sau khi ghép được thể hiện trên Fig.4. Trên Fig.4, từ trái sang phải, lá thứ nhất là giống dễ bị mắc bệnh Nipponbare (Nip), lá thứ 2 là giống lúa EBZ có khả năng kháng lại bệnh đạo ôn, các lá từ thứ 3 đến thứ 6 là các cá thể dương tính với chuyển gen pCambia-Pi50, lá thứ 7 là giống có khả năng kháng IRBLzt-T (các giống của thể cho Piz-t, được bộc lộ trong tài liệu "Zhou et al., 2006. 8 axit amin khác nhau trong ba vùng lặp lại giàu leucin giữa các protein có khả năng kháng Pi2 và Piz-t xác định khả năng kháng đặc hiệu với vi khuẩn gây bệnh đạo ôn. MPMI, 19: 1216–1228"), lá thứ 8 là giống có khả năng kháng IRBLz5-CA (các giống của thể cho Pi2, bộc lộ trong tài liệu "WANG Wen-juan, etc., "Pathogenicity analysis of the rice blast fungus isolated from the blast panicles of Yuejingsimiao 2" Guangdong Agricultural Science, 2012 (23): 59 -61"). Như được thể hiện trên Fig.4, tất cả các giống Nipponbare, IRBLzt-T và IRBLz5-CA đều xuất hiện vết tan bệnh, và thể hiện khả năng nhạy cảm với chủng TY 19; tuy nhiên, các cá thể dương tính với chuyển gen EBZ và pCambia-Pi50 không xuất hiện vết tan bệnh, và thể hiện khả năng kháng lại bệnh.

Một phần của các chủng chuyển gen T₁ thu được tiếp tục được chọn lọc, và thực hiện phân tích đồng phân ly trên kiểu hình có khả năng kháng bệnh/kiểu gen có khả năng kháng bệnh. Các kết quả chỉ ra rằng kiểu hình có khả năng kháng bệnh và kiểu gen

có khả năng kháng bệnh được phân ly hoàn toàn đồng thời, tức là, các cấu trúc biến đổi gen có thể được di truyền ổn định và biểu hiện ở thế hệ con cháu chuyển gen, điều này cho thấy gen Pi50 có thể được sử dụng để chuyển vào giống lúa để mắc bệnh, và sau đó một loạt các tiến trình nhân giống như tự lai, sàng lọc, chọn lọc, v.v., sẽ tạo ra các chủng có khả năng kháng bệnh và được ứng dụng trong sản xuất.

Các ví dụ trên đây là các phương án ưu tiên của sáng chế, nhưng các phương án này không nhằm giới hạn sáng chế, các thay đổi, sửa đổi, thay thế, kết hợp, và đơn giản hóa bất kỳ đều là các cách thức thay thế tương đương, không vượt ra khỏi tinh thần và nguyên lý của sáng chế, và đều thuộc phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Gen kháng bệnh đạo ôn Pi50, khác biệt ở chỗ trình tự nucleotit của gen kháng bệnh đạo ôn Pi50 được thể hiện dưới dạng trình tự SEQ ID NO.1.
2. Phương pháp tạo dòng gen kháng bệnh đạo ôn Pi50 theo điểm 1, khác biệt ở chỗ gen kháng bệnh đạo ôn Pi50 có thể thu được bằng phương pháp tổng hợp hóa học hoặc thu được bằng cách thiết kế các cặp mồi bằng cách sử dụng ADN thuộc hệ gen của giống lúa EBZ làm khuôn, sau đó thực hiện khuếch đại gen bằng phản ứng chuỗi trùng hợp (PCR - Polymerase Chain Reaction).
3. Trình tự protein được mã hóa bởi gen kháng bệnh đạo ôn Pi50 theo điểm 1, như được thể hiện trong trình tự SEQ ID NO.2.

Chủng	Nguồn gốc	Kiểu phản ứng							
		Ni1-e1 Pi50(t)	Ni1-e2 Pi50(t)	IRBL25-CA Pi2	C101A51 Pi2+Pi6	IRBL9-W Pi9	IRBL2-Fu Pi2	IRBL2t-T Pi2-t	Toride Pi2-t+Pi5h
08-2045a	Guangdong	8	8	1	1	1	1	1	1
W08-037a	Guangdong	8	8	1	1	1	3	3	3
10-026a	Guangdong	1	1	7	6	1	3	1	1
SC0602	Sichuan	1	1	8	7	1	7	1	1
05-006a	Guangdong	1	1	1	1	8	7	1	2
W06-	Guangxi	1	1	1	1	6	3	1	1
05-161a	Guangdong	1	1	2	1	1	8	1	1
HL107-80-1	Heilongjiang	1	1	1	1	1	8	1	1
FJ0601	Fujian	1	1	1	1	1	1	7	7
JL0807	Jilin	1	1	1	2	1	1	7	7
phổ kháng (523 chủng)		97,7	97,7	92,2	92,3	93,7	68,4	54,5	66,5

Fig.1

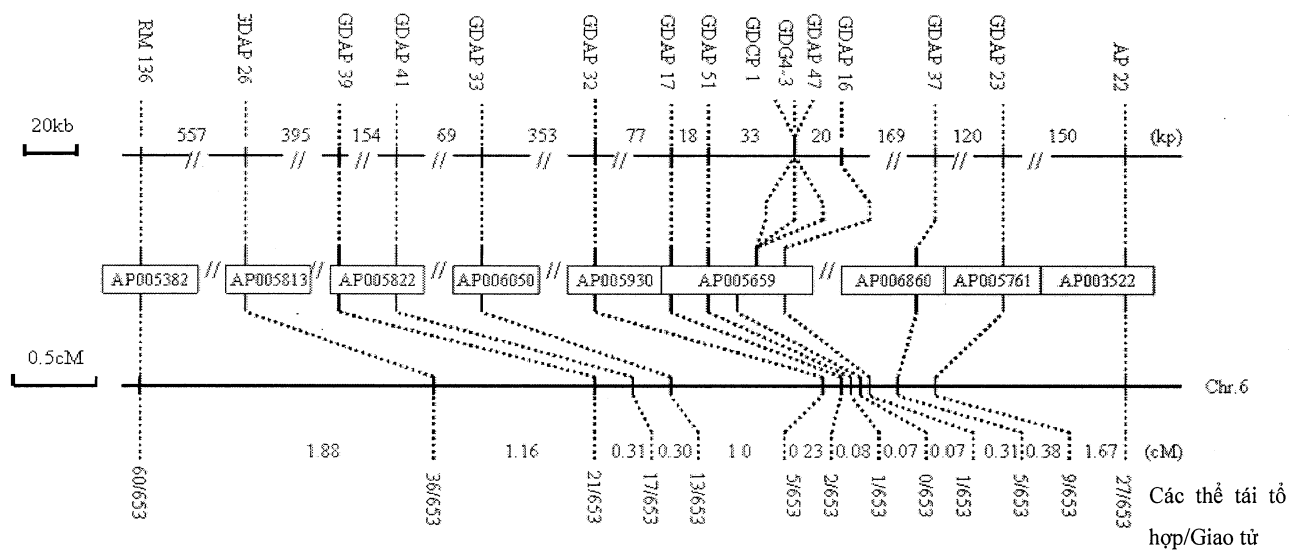


Fig.2

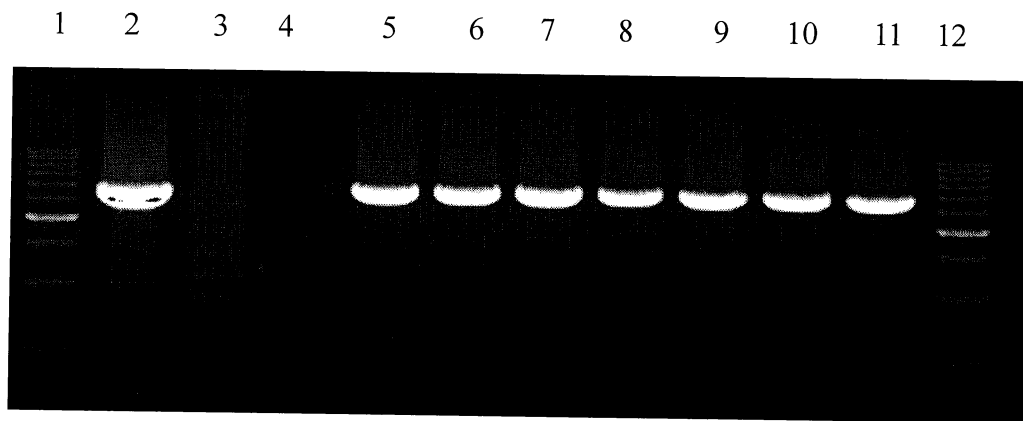


Fig.3

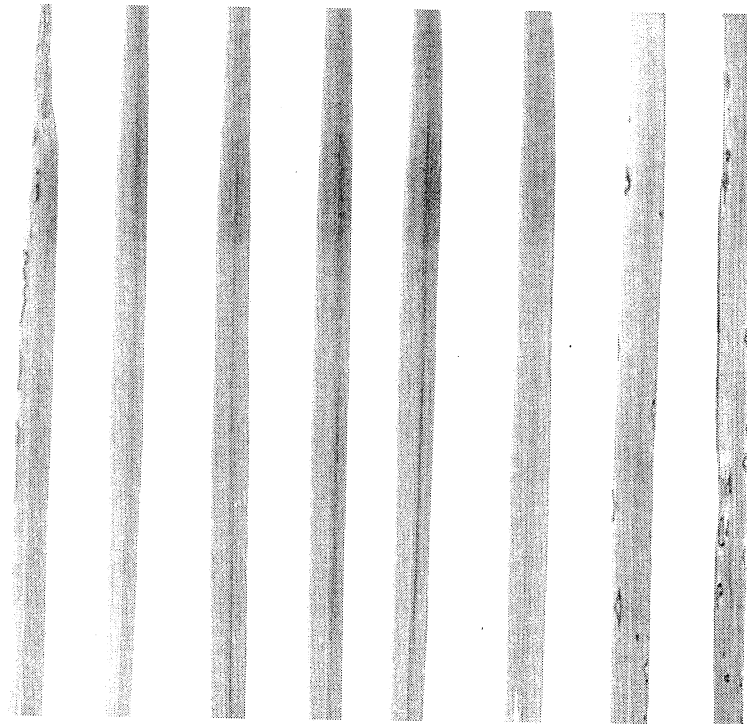


Fig.4

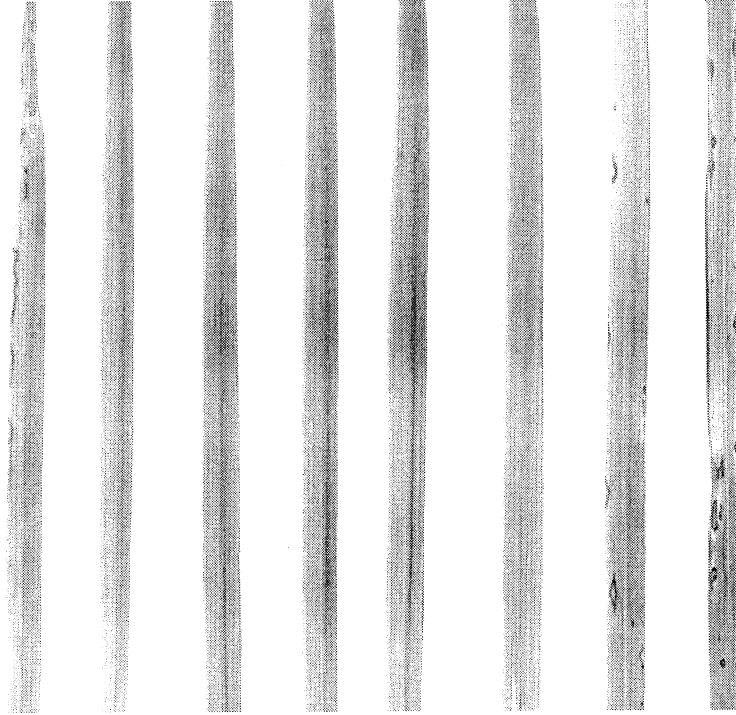


Fig.5

DANH MỤC TRÌNH TỰ

- <110> Plant Protection Research Institute of Guangdong Academy of Agricultural science
- <120> Gen kháng bệnh đạo ôn Pi50, phương pháp tạo dòng gen này và trình tự protein được mã hóa bởi gen này
- <130> 1
- <160> 10
- <170> PatentIn version 3.2
- <210> 1
- <211> 3099
- <212> ADN
- <213> lúa
- <400> 1
- | | |
|---|------|
| atggcgagaga cgggtgctgag catggcgagg tcgctgggtg gcagcgccat cagcaaggcc | 60 |
| gcctccgctg ctgccgacga gaccagcctc ctgctgggcg tcgagaaaga catctggtat | 120 |
| atcaaagatg agctaaaaac gatgcaagca ttccttagag ctgctgaact tatgaaaaag | 180 |
| aaagatgaac tattaaggt ttgggcagag caaatacgtg acctgtcata tgacattgaa | 240 |
| gattcccttg atgaatttaa ggtccatatt gaaagccaaa ccctatttcg tcagttgggtg | 300 |
| aaactcagag aacgccaccg aattgctatc cgtatccaca accttaaadc aagagttgaa | 360 |
| gaagtgaagta gcaggaacac acgctacagt ttagtcaagc ctatttcctc tggcacagag | 420 |
| attgacatgg attcctatgc agaagacatt cgtaatcagt cagctcacia tgtggatgaa | 480 |
| gctgagcttg ttgggttttc tgactccaag aaaaggttgc ttgaaatgat cgataccaat | 540 |
| gctaattgat gtccggccaa agtaattctgt gttgttggga tgggtgggtt aggcaagaca | 600 |
| gctctttcga ggaagatctt tgaaagcgaa gaagacatta ggaagaactt cccttgcaat | 660 |
| gcttgatta cagtgtcaca atcatttcac aggattgagc tactcaaaga tatgatacgc | 720 |
| caacttcttg gccccagttc tctggatcaa ctcttgcaag aattgcaagg gaaggtgggtg | 780 |
| gtgcaagtac atcatctttc tgagtacctg atagaagagc tcaaggagaa gaggtacttt | 840 |
| gttgttctag atgatctatg gattttacat gattggaatt ggataaatga aattgcattt | 900 |
| cctaagaaca ataagaaggg cagtcgaata gtaataacca ctcggaatgt tgatcttgcg | 960 |
| gagaagtgtg ccacagcctc actggtgtac caccttgatt tcttgcatat gaacgatgcc | 1020 |
| ataacattgc tactgagaaa aacaaataaa aatcatgaag acatggaatc aaataaaaaat | 1080 |
| atgcaaaaaga tgggtgaacg aattgtaaat aaatgtggtc gtctaccatt agcaataactt | 1140 |
| acaataggag ctgtgcttgc aactaaacag gtgtcagaat gggagaaaatt ctatgaacaa | 1200 |
| cttccttcag aactagaaat aaaccaagc ctggaagctt tgaggagaat ggtgacccta | 1260 |
| ggttacaacc acctaccatc ccattctgaaa ccattgctttt tgtatctaag tatcttttct | 1320 |

gaggattttg aaatcaaaaag gaatcgtcta gtaggtatat ggatagcaga aggggttggtt 1380
 agaccaaagg ttgggatgac gactaaggat gtcggagaaa gttactttta tgagctaatac 1440
 aaccgaagta tgattcaacg atcaagagtg ggcatagcag gaaaaattaa gacttgtcga 1500
 attcatgata tcatccgtga tatcacagtt tcaatctcga gacaggaaaaa ttttgtatta 1560
 ttaccaatgg gagatggctc tgatttagtt caggaaaaca ctgccacat agcattccat 1620
 gggagtatgt cctgcaaaac aggattggat tggagcatta ttcgatcatt agctatTTTT 1680
 ggtgacagac ccaagagtct agcacatgca gtttgtccag atcaattgag gatgttacgg 1740
 gtcttggatc ttgaagatgt gacattctta atcactcaaa aagatttcga ccgtattgca 1800
 ttgttgtgcc acttgaaata cttgagtatt ggatattcgt catccatata ttcacttccc 1860
 agatccattg gtaaaactaca gggcctacaa actttgaaca tgccgagcac atacattgca 1920
 gcactaccaa gtgagatcag taaactccaa tgtctgcata ctcttcgttg tataggacag 1980
 tttcattatg acaactttag tctaaaccac ccaatgaagt gcataactaa cacaatatgc 2040
 ctgcctaaag tattcacacc tttagttagt cgcgatgatc gtgcaaaaca aattgctgaa 2100
 ttgcacatgg ccaccaaagg ttgctggctc gaatcattcg gtgtgaagggt acccaaagga 2160
 ataggtaagt tgcgagactt gcaggttcta gagtatgtag atatcaagcg gaccagtagt 2220
 agagcaatca aagagctggg gcagtttaagc aagttgagga aattaggtgt gataacaaaa 2280
 ggctcgacaa aggaaaaatg taagatactt tatgcagcca ttgagaagct ctcttccttc 2340
 caatatctct atgtgaatgc tgcgttatta tcagatattg aaacacttga gtgcctagat 2400
 tctatttcat ctctcctcc cctactaagt acactcaggt tgaatggaag tcttgaagag 2460
 atgcctaact ggattgagca gctcactcac ctgaagaagt tctacttacg gaggagcaaa 2520
 ctaaaggaag gtaaaacat gctgatactt ggggcactgc ccaacctcat gttcctttct 2580
 ctttatcata attcttatct tggggagaag ctagtattca aaacgggagc attcccaaat 2640
 cttagaacac tttgtattta cgaattggat cagctaagag agatcagatt tgaggacggc 2700
 agctcacccc tgttggaaaa gatagaaata ggcaagtgc ggttggaatc tgggattatt 2760
 ggtatcattc accttccaaa gctcaaggag attccaatta catacggaag taaagtggct 2820
 gggcttggtc agctggaggg agaagtgaac gcacacccaa atcgccccgt gctgctaattg 2880
 gacagtgacc gaaggtatca cgacctgggg gctgaagccg aaggatcttc tatagaagtg 2940
 caaacagcag atcctgttcc tgatgccgaa ggatcagtca ctgtagcagt ggaagcaacg 3000
 gatccccctc ccgagcagga gggagagagc tcgcagtcgc aggtgatcac gttgacgacg 3060
 aatgatagcg aagagatagc cacagctcaa gctggctga 3099

<210> 2

<211> 1032

<212> PRT

<213> lúa

<400> 2

Met Ala Glu Thr Val Leu Ser Met Ala Arg Ser Leu Val Gly Ser Ala
 1 5 10 15
 Ile Ser Lys Ala Ala Ser Ala Ala Ala Asp Glu Thr Ser Leu Leu Leu
 20 25 30
 Gly Val Glu Lys Asp Ile Trp Tyr Ile Lys Asp Glu Leu Lys Thr Met
 35 40 45
 Gln Ala Phe Leu Arg Ala Ala Glu Leu Met Lys Lys Lys Asp Glu Leu
 50 55 60
 Leu Lys Val Trp Ala Glu Gln Ile Arg Asp Leu Ser Tyr Asp Ile Glu
 65 70 75 80
 Asp Ser Leu Asp Glu Phe Lys Val His Ile Glu Ser Gln Thr Leu Phe
 85 90 95
 Arg Gln Leu Val Lys Leu Arg Glu Arg His Arg Ile Ala Ile Arg Ile
 100 105 110
 His Asn Leu Lys Ser Arg Val Glu Glu Val Ser Ser Arg Asn Thr Arg
 115 120 125
 Tyr Ser Leu Val Lys Pro Ile Ser Ser Gly Thr Glu Ile Asp Met Asp
 130 135 140
 Ser Tyr Ala Glu Asp Ile Arg Asn Gln Ser Ala His Asn Val Asp Glu
 145 150 155 160
 Ala Glu Leu Val Gly Phe Ser Asp Ser Lys Lys Arg Leu Leu Glu Met
 165 170 175
 Ile Asp Thr Asn Ala Asn Asp Gly Pro Ala Lys Val Ile Cys Val Val
 180 185 190
 Gly Met Gly Gly Leu Gly Lys Thr Ala Leu Ser Arg Lys Ile Phe Glu
 195 200 205
 Ser Glu Glu Asp Ile Arg Lys Asn Phe Pro Cys Asn Ala Trp Ile Thr
 210 215 220
 Val Ser Gln Ser Phe His Arg Ile Glu Leu Leu Lys Asp Met Ile Arg
 225 230 235 240
 Gln Leu Leu Gly Pro Ser Ser Leu Asp Gln Leu Leu Gln Glu Leu Gln
 245 250 255
 Gly Lys Val Val Val Gln Val His His Leu Ser Glu Tyr Leu Ile Glu
 260 265 270
 Glu Leu Lys Glu Lys Arg Tyr Phe Val Val Leu Asp Asp Leu Trp Ile
 275 280 285

Leu His Asp Trp Asn Trp Ile Asn Glu Ile Ala Phe Pro Lys Asn Asn
 290 295 300
 Lys Lys Gly Ser Arg Ile Val Ile Thr Thr Arg Asn Val Asp Leu Ala
 305 310 315 320
 Glu Lys Cys Ala Thr Ala Ser Leu Val Tyr His Leu Asp Phe Leu Gln
 325 330 335
 Met Asn Asp Ala Ile Thr Leu Leu Leu Arg Lys Thr Asn Lys Asn His
 340 345 350
 Glu Asp Met Glu Ser Asn Lys Asn Met Gln Lys Met Val Glu Arg Ile
 355 360 365
 Val Asn Lys Cys Gly Arg Leu Pro Leu Ala Ile Leu Thr Ile Gly Ala
 370 375 380
 Val Leu Ala Thr Lys Gln Val Ser Glu Trp Glu Lys Phe Tyr Glu Gln
 385 390 395 400
 Leu Pro Ser Glu Leu Glu Ile Asn Pro Ser Leu Glu Ala Leu Arg Arg
 405 410 415
 Met Val Thr Leu Gly Tyr Asn His Leu Pro Ser His Leu Lys Pro Cys
 420 425 430
 Phe Leu Tyr Leu Ser Ile Phe Pro Glu Asp Phe Glu Ile Lys Arg Asn
 435 440 445
 Arg Leu Val Gly Ile Trp Ile Ala Glu Gly Leu Val Arg Pro Lys Val
 450 455 460
 Gly Met Thr Thr Lys Asp Val Gly Glu Ser Tyr Phe Tyr Glu Leu Ile
 465 470 475 480
 Asn Arg Ser Met Ile Gln Arg Ser Arg Val Gly Ile Ala Gly Lys Ile
 485 490 495
 Lys Thr Cys Arg Ile His Asp Ile Ile Arg Asp Ile Thr Val Ser Ile
 500 505 510
 Ser Arg Gln Glu Asn Phe Val Leu Leu Pro Met Gly Asp Gly Ser Asp
 515 520 525
 Leu Val Gln Glu Asn Thr Arg His Ile Ala Phe His Gly Ser Met Ser
 530 535 540
 Cys Lys Thr Gly Leu Asp Trp Ser Ile Ile Arg Ser Leu Ala Ile Phe
 545 550 555 560
 Gly Asp Arg Pro Lys Ser Leu Ala His Ala Val Cys Pro Asp Gln Leu
 565 570 575

Arg Met Leu Arg Val Leu Asp Leu Glu Asp Val Thr Phe Leu Ile Thr
 580 585 590
 Gln Lys Asp Phe Asp Arg Ile Ala Leu Leu Cys His Leu Lys Tyr Leu
 595 600 605
 Ser Ile Gly Tyr Ser Ser Ser Ile Tyr Ser Leu Pro Arg Ser Ile Gly
 610 615 620
 Lys Leu Gln Gly Leu Gln Thr Leu Asn Met Pro Ser Thr Tyr Ile Ala
 625 630 635 640
 Ala Leu Pro Ser Glu Ile Ser Lys Leu Gln Cys Leu His Thr Leu Arg
 645 650 655
 Cys Ile Gly Gln Phe His Tyr Asp Asn Phe Ser Leu Asn His Pro Met
 660 665 670
 Lys Cys Ile Thr Asn Thr Ile Cys Leu Pro Lys Val Phe Thr Pro Leu
 675 680 685
 Val Ser Arg Asp Asp Arg Ala Lys Gln Ile Ala Glu Leu His Met Ala
 690 695 700
 Thr Lys Ser Cys Trp Ser Glu Ser Phe Gly Val Lys Val Pro Lys Gly
 705 710 715 720
 Ile Gly Lys Leu Arg Asp Leu Gln Val Leu Glu Tyr Val Asp Ile Lys
 725 730 735
 Arg Thr Ser Ser Arg Ala Ile Lys Glu Leu Gly Gln Leu Ser Lys Leu
 740 745 750
 Arg Lys Leu Gly Val Ile Thr Lys Gly Ser Thr Lys Glu Lys Cys Lys
 755 760 765
 Ile Leu Tyr Ala Ala Ile Glu Lys Leu Ser Ser Leu Gln Tyr Leu Tyr
 770 775 780
 Val Asn Ala Ala Leu Leu Ser Asp Ile Glu Thr Leu Glu Cys Leu Asp
 785 790 795 800
 Ser Ile Ser Ser Pro Pro Pro Leu Leu Ser Thr Leu Arg Leu Asn Gly
 805 810 815
 Ser Leu Glu Glu Met Pro Asn Trp Ile Glu Gln Leu Thr His Leu Lys
 820 825 830
 Lys Phe Tyr Leu Arg Arg Ser Lys Leu Lys Glu Gly Lys Thr Met Leu
 835 840 845
 Ile Leu Gly Ala Leu Pro Asn Leu Met Phe Leu Ser Leu Tyr His Asn
 850 855 860

Ser Tyr Leu Gly Glu Lys Leu Val Phe Lys Thr Gly Ala Phe Pro Asn
 865 870 875 880
 Leu Arg Thr Leu Cys Ile Tyr Glu Leu Asp Gln Leu Arg Glu Ile Arg
 885 890 895
 Phe Glu Asp Gly Ser Ser Pro Leu Leu Glu Lys Ile Glu Ile Gly Lys
 900 905 910
 Cys Arg Leu Glu Ser Gly Ile Ile Gly Ile Ile His Leu Pro Lys Leu
 915 920 925
 Lys Glu Ile Pro Ile Thr Tyr Gly Ser Lys Val Ala Gly Leu Gly Gln
 930 935 940
 Leu Glu Gly Glu Val Asn Ala His Pro Asn Arg Pro Val Leu Leu Met
 945 950 955 960
 Asp Ser Asp Arg Arg Tyr His Asp Leu Gly Ala Glu Ala Glu Gly Ser
 965 970 975
 Ser Ile Glu Val Gln Thr Ala Asp Pro Val Pro Asp Ala Glu Gly Ser
 980 985 990
 Val Thr Val Ala Val Glu Ala Thr Asp Pro Leu Pro Glu Gln Glu Gly
 995 1000 1005
 Glu Ser Ser Gln Ser Gln Val Ile Thr Leu Thr Thr Asn Asp Ser
 1010 1015 1020
 Glu Glu Ile Gly Thr Ala Gln Ala Gly
 1025 1030

<210> 3

<211> 7063

<212> ADN

<213> lúa

<220>

<221> exon thứ nhất

<222> (1)..(116)

<220>

<221> exon thứ hai

<222> (3953)..(6904)

<220>

<221> exon thứ ba

<222> (7033)..(7063)

<400> 3

atggcggaga	cggtgctgag	catggcgagg	tcgctggtgg	gcagcgccat	cagcaaggcc	60
gcctccgctg	ctgccgacga	gaccagcctc	ctgctgggcg	tcgagaaaga	catctggtac	120
gtactgctg	actctcgta	atttattctg	tagatgctca	ggaatcagca	actatttgt	180
tgatttccat	cgtagcatat	cgattttgtt	ggccaccaat	tctaatacggc	cggaacaagc	240
tagtcactaa	atctggcaaa	tcgatcagct	gctgagtgc	caaacatgca	tgttattctt	300
tttttttttg	gttatatgtt	aagcaacaaa	gcccttggt	aagatatgca	tggcaaatga	360
actaatatcg	acatacgtaa	agcggaggac	ccctcgttcc	atgcgtgggt	gactcgagcg	420
gtgacaaatc	ctagcacctc	cacctccttg	gatggcctgt	ggtgacgctt	tcggccccga	480
gtttcccttg	aatacatcat	ctacaagggtg	ctattaatgt	ctagtcacat	catttacgag	540
gcgctattat	tgcctagtct	gccgaagat	agtttagaca	acactcttgg	atggcgggtg	600
ccttcgcca	gtgatgtcca	agagcccggtg	gatgtttaat	tgtttagaca	tggtgttggg	660
tggtgcacta	gtgggcctga	tgggccagtt	gtaggtccag	tggtaaccaa	tcattgcttag	720
caatagccgg	atgcccggt	tgggtgctt	tcttttttcg	gtgtcgacgc	atggtagtat	780
ttacttttcc	tgtttttcct	gattatagca	tcctaggcta	tactcttcta	atttattcat	840
gctatattaa	tattaaaact	tgggtatggtt	tgtttcattc	aagacccttg	gtggtcaaag	900
gcttggttgg	ttcaagttca	ttcctagcct	taccaacttt	ttggcaatag	caagaaatgg	960
tcattgaaaa	aaaaaggcaa	aaattggcta	ggcctacagt	ttatttccta	gcaaagttat	1020
acttttagcat	tccactaagc	caaataattc	ggcaatgcca	ttttcttata	tacatgccaa	1080
atatatggct	aataatttgg	cattaattac	tcttattttt	tttggcaaaa	ttgatcaaaa	1140
gttcacattt	ttagctctat	agtatcaaaa	gttatctatt	cactttaata	gaacgaaagt	1200
ttactcggtt	cgttttttag	cactaccgtc	tcttttctct	tgatttgccg	tcaattttga	1260
ccggcagtc	tacccccagg	agacattgag	cagcagcccg	tgatccccct	ctctcgccgc	1320
cggtgacgct	gtggtggcat	cgttcctgct	gcgggcagaa	taagtctggc	gtcatcgccc	1380
tatcgctgg	agctgcaacc	accactaccg	ggcccatcga	tcgtctagag	cgttatccac	1440
cctgcctgcc	ccattacttg	cagctccggc	tgggtcagaa	cctctccatg	cctgacaaat	1500
tggttcaaga	ttgtcgctgt	ccggccagcg	cttgaatttt	cagaatatgc	catcgaatac	1560
gcgctgcttt	taagatatgc	tacccgattc	gtgctaattt	tagaatacgc	catcggaaca	1620
cgaattttct	tcgttccgtg	ccactccgtc	tctcgagtc	agtcgtgccg	tcgtcatccg	1680
tccgcccagc	actgtcgctg	tcagtccgcc	accgtgcct	gactgcccgt	tcagctgcgc	1740
cgtcgctccg	ccgtcgccgc	catcgccgtc	gtccactgtc	gcgcccgcac	ctgcaccctg	1800
gtcaggcgcg	ccgtcgctccg	tccgctgtca	ccatcatcgt	cgtccactgc	ggcgagcgca	1860
gacggctgtg	gacggatgat	tgcgcggcaa	gcgcaagcgg	cgggtggacgg	acgacggggc	1920
agtgcacgcg	agcacggtag	ccgatggact	gacgttggcg	gcattgggag	atggacgacg	1980
acggcatgat	tgacgtgggg	aacggaatgt	cacggawcgg	agaaaattcg	cactccgggtg	2040
gcgtatatcc	cgaaaatagc	acgaattggg	tggcatatcc	taaaaacagt	gcctattcgg	2100

tgccatattc	taaaaattct	cggtcagcat	aatccccatc	aatccccaat	ccctcaacag	2160
ttgggttaat	attcctggag	atgtgttcgg	ttgttttaggt	tgaagttctc	cacttcacct	2220
ccatgagtag	atgcacctct	acacgtacgt	tcttaatgag	tttgtttgtt	ctatcctccg	2280
cttgggttct	attttggttg	ttccgatctg	atgtgatctg	gagcgggggtc	gatcttccac	2340
gacggcgaga	gagacgttgt	tcgggctgct	cgatttggtt	caactgttta	ggtcgaaggg	2400
aggggtagaa	ttgcaattca	agtgcattgt	cagtcaattt	gggtcaaaat	taacatcaaa	2460
ctgggataaa	gagacgacag	tgccaaaatt	ggtaacggga	aaactttgag	ttctattaaa	2520
gtgaaccagt	aactttcggt	gctatagaat	aaaaacgtaa	acttttgatg	ctccttgggt	2580
gagcttggtg	caaaccacac	agacgtaaaa	taaacactat	cctgaatcaa	gtctactaag	2640
ttccattgaa	ctcaaccagg	atacgtacac	ttcctcttag	aagatgtctt	gttttcactt	2700
tgtacaattt	tttctattgt	aaatttggtg	cctcgttgta	cctaggtaca	agaggtacca	2760
tgagatacca	aattttacac	taaaattttg	gtacctcatg	gtacctcctc	aacgaccgta	2820
gaattgctct	taattttaatt	taaaaaaaaa	cataatattt	ttaaagcata	ttatggaaat	2880
tttagtaatt	attacttttg	taatatatga	gttacgggtg	tactcgagat	atcctaaatt	2940
gcttggagat	gaataattac	aaggtatatc	aaagatgagt	tgaaaataat	gcaggcattc	3000
cttagagctg	cagaagttat	gaaaaagaaa	gacgaactat	taaagggttg	ggtagagcaa	3060
atacgtgacc	tgtcatatga	cattgaagat	tccttgatg	aatttaagggt	ccatgttgaa	3120
agtcaaacc	tatttcgtca	gttggtgaaa	cttagatagc	gtcaccggat	cgctagcccg	3180
tggatgttta	gttgtttgca	catgggtgctg	gatgggtgct	tcattgggtc	gttgtaggtc	3240
tggtagcaac	cagtcattgt	tagaaatagc	cggatcagtg	cacgggtgct	ggactttact	3300
tgggtggtct	tgcagcgcta	tcgacatgtg	gtgggtgtgt	tttttttttc	cggattacaa	3360
tctcataggg	ctacactcta	gttattttgc	tgctatatta	atatgaaaac	ttgggtatgt	3420
tcgtttctta	tagaaaaaaa	cctagttgat	caagggttag	ttttcttcaa	gtgcattcct	3480
aatcttagct	tctttttttt	ttgcaatggc	aagaattgtt	cattaaaaaa	attgataaaa	3540
attggctagg	cctacgtttt	gtttcttacc	aaagtgttac	tttaacaata	aactaaggca	3600
aatatttcgg	caatgccatt	ttcttgtcta	cagaccaa	atatggctaa	attttggcat	3660
aaccattttt	ttgtttgctt	ggttgagctt	ggtacaaacc	aaacagaccc	aaaataaaca	3720
gtgtcatgaa	tcacgtctac	taaattcctt	tgaactgaac	tagaatatag	ttgctcttaa	3780
aagatttctt	gatttcactc	ggtaccattt	actagtacaa	acttagattt	aatttttaaa	3840
aataaaatca	taatattggt	attatggaaa	atttagtcat	agtacttttg	taatatatga	3900
gatgggttat	acttgagata	tcctaaattg	ctttaagatg	aataattgct	aggtatatca	3960
aagatgagct	aaaaacgatg	caagcattcc	ttagagctgc	tgaacttatg	aaaaagaaag	4020
atgaactatt	aaagggttgg	gcagagcaaa	tacgtgacct	gtcatatgac	attgaagatt	4080
cccttgatga	atttaagggtc	catattgaaa	gccaaaccct	atctcgtcag	ttgggtgaaac	4140
tcagagaacg	ccaccgaatt	gctatccgta	tccacaacct	taaatcaaga	gttgaagaag	4200
tgagtagcag	gaacacacgc	tacagtttag	tcaagcctat	ttcctctggc	acagagattg	4260

acatggattc ctatgcagaa gacattcgta atcagtcagc tcacaatgtg gatgaagctg 4320
 agcttggttg gttttctgac tccaagaaaa ggttgcttga aatgatcgat accaatgcta 4380
 atgatgggtcc ggccaaagta atctgtgttg ttgggatggg tggtttaggc aagacagctc 4440
 tttcgaggaa gatctttgaa agcgaagaag acattaggaa gaacttccct tgcaatgctt 4500
 ggattacagt gtcacaatca tttcacagga ttgagctact caaagatatg atacgccaac 4560
 ttcttggtccc cagttctctg gatcaactct tgcaagaatt gcaaggggaag gtgggtggtgc 4620
 aagtacatca tctttctgag tacctgatag aagagctcaa ggagaagagg tactttgttg 4680
 ttctagatga tctatggatt ttacatgatt ggaattggat aaatgaaatt gcatttccta 4740
 agaacaataa gaagggcagt cgaatagtaa taaccactcg gaatgttgat cttgcggaga 4800
 agtgtgccac agcctcactg gtgtaccacc ttgatttctt gcagatgaac gatgccataa 4860
 cattgctact gagaaaaaca aataaaaatc atgaagacat ggaatcaaatt aaaaatatgc 4920
 aaaagatggt tgaacgaatt gtaaataaat gtggtcgtct accattagca atacttacia 4980
 taggagctgt gcttgcaact aaacaggtgt cagaatggga gaaattctat gaacaacttc 5040
 cttcagaact agaaataaac ccaagcctgg aagctttgag gagaatggtg accctaggtt 5100
 acaaccacct accatcccat ctgaaacct gctttttgta tctaagtatc tttcctgagg 5160
 attttgaaat caaaaggaat cgtctagtag gtatatggat agcagaaggg ttggttagac 5220
 caaaggttg gatgacgact aaggatgtcg gagaaagtta cttttatgag ctaatcaacc 5280
 gaagtatgat tcaacgatca agagtgggca tagcaggaaa aattaagact tgtcgaattc 5340
 atgatatcat ccgtgatatc acagtttcaa tctcgagaca ggaaaatttt gtattattac 5400
 caatgggaga tggctctgat ttagttcagg aaaacactcg ccacatagca ttccatggga 5460
 gtatgtcctg caaaacagga ttggattgga gcattattcg atcattagct atttttggtg 5520
 acagacccaa gagtctagca catgcagttt gtccagatca attgaggatg ttacgggtct 5580
 tggatcttga agatgtgaca ttcttaatca ctcaaaaaga tttcgaccgt attgcattgt 5640
 tgtgccactt gaaatacttg agtattggat attcgtcatc catatattca cttcccagat 5700
 ccattggtaa actacagggc ctacaaactt tgaacatgcc gagcacatac attgcagcac 5760
 taccaagtga gatcagtaaa ctccaatgtc tgcatactct tcgttgata ggacagtttc 5820
 attatgacaa ctttagtcta aaccacccaa tgaagtgc atactaacaca atatgcctgc 5880
 ctaaagtatt cacaccttta gttagtcgag atgatcgtgc aaaacaaatt gctgaattgc 5940
 acatggccac caaaagttgc tggctctgaat cattcgggtg gaaggtaccc aaaggaatag 6000
 gtaagttgcg agacttgag gttctagagt atgtagatat caagcggacc agtagtagag 6060
 caatcaaaga gctggggcag ttaagcaagt tgaggaaatt aggtgtgata acaaaaggct 6120
 cgacaaagga aaaatgtaag atactttatg cagccattga gaagctctct tccctccaat 6180
 atctctatgt gaatgctgcg ttattatcag atattgaaac acttgagtgc ctagattcta 6240
 tttcatctcc tcctccccta ctaagtacac tcaggttgaa tggaagtctt gaagagatgc 6300
 ctaactggat tgagcagctc actcacctga agaagttcta cttacggagg agcaaaactaa 6360
 aggaaggtaa aaccatgctg atacttgggg cactgcccaa cctcatgttc ctttctcttt 6420

```

atcataattc ttatcttggg gagaagctag tattcaaaac gggagcattc ccaaattcta 6480
gaacactttg tatttacgaa ttggatcagc taagagagat cagatttgag gacggcagct 6540
caccctgtt ggaaaagata gaaataggca agtgcagggtt ggaatctggg attattggta 6600
tcattcacct tccaaagctc aaggagattc caattacata cggaagtaaa gtggctgggc 6660
ttggtcagct ggagggagaa gtgaacgcac acccaaatcg ccccgctgctg ctaatggaca 6720
gtgaccgaag gtatcacgac ctgggggctg aagccgaagg atcttctata gaagtgcaaa 6780
cagcagatcc tgttcctgat gccgaaggat cagtcactgt agcagtggaa gcaacggatc 6840
cccttcccgga gcaggaggga gagagctcgc agtcgcagggt gatcacgttg acgacgaatg 6900
ataggtcagt cactccctag atggcagctt aattaacttg tttctaattc tcttcttgtt 6960
cagtattagc catcaggtga gggcgatgat ttcaactcac ttttcatctc tctcgttttc 7020
ttaacctgac agcgaagaga taggcacagc tcaagctggc tga 7063

```

<210> 4

<211> 7238

<212> ADN

<213> lúa

<220>

<221> exon thứ nhất

<222> (176)..(291)

<220>

<221> exon thứ hai

<222> (4128)..(7079)

<220>

<221> exon thứ ba

<222> (7208)..(7238)

<400> 4

```

cttgacatcc aaaccgcacc taactctatg ataagcatcc tctctcagat tgttcagtgc 60
aaaagctacc aatactgctc cgagagccag aaaaaagcgc tggtcgccta gtgctatctt 120
ctatgcagtc gtgagattgg ttgctctaag cttgaaggga gagtcgaacg agtccatggc 180
ggagacggtg ctgagcatgg cgaggctcgt ggtgggcagc gccatcagca aggccgcctc 240
cgctgctgcc gacgagacca gcctcctgct gggcgctcag aaagacatct ggtacgtact 300
gcgtgactct cgttaattta ttctgtagat gctcaggaat cagcaactat tgtgttgatt 360
tccatcgtag catatcgatt ttgttgcca ccaattctaa tcggccggaa caagctagtc 420
actaaatctg gcaaatcgat cagctgctga gtgcacaaac atgcatgtta ttcttttttt 480
tttgggttat atgttaagca acaaagcccc ttggtaagat atgcatggca aatgaactaa 540
tatcgacata cgtaaagcgg aggaccctc gttccatgcg tgggtgactc gagcgggtgac 600

```

aaatcctagc	acctccacct	ccttggaagg	cctgtggtga	cgctttcggc	cccagatttc	660
ccttgaatac	atcatctaca	aggtgctatt	aatgtctagt	cacatcattt	acgaggcgct	720
attattgcct	agtctgcccg	aagatagttt	agacaacact	cttggaaggc	ggtgtccttc	780
gcccagtgat	gtccaagagc	ccgtggaagt	ttaattgttt	agacatggtg	ttgggtggtg	840
cactagtggg	cctgatgggc	cagttgtagg	tccagtggta	accaatcatg	cttagcaata	900
gccggatgcc	cggattggtg	cttggtcttt	tttcggtgtc	gacgcattgt	agtatttact	960
tttcctgttt	ttcctgatta	tagcatccta	ggctatactc	ttctaattta	ttcatgctat	1020
attaatatta	aaacttggtg	tggtttggtt	cattcaagac	ccttggtggt	caaaggcttg	1080
tttggttcaa	gttcattcct	agccttacca	actttttggc	aatagcaaga	aatggtcatt	1140
gaaaaaaaa	ggcaaaaatt	ggctaggcct	acagtttatt	tcctagcaaa	gttatacttt	1200
agcattccac	taagccaaat	aattcggaac	tgccattttc	ttatctacat	gccaaatata	1260
tggctaatat	tttggcatta	attactctta	ttttttttgg	caaaattgat	caaaagttca	1320
catttttagc	tctatagtat	caaaagttat	ctattcactt	taatagaacg	aaagtttact	1380
cggttccggt	tttagcacta	ccgtctcttt	tctcttgatt	tgccgtcaat	tttgaccggc	1440
agtcctaccc	ccaggagaca	ttgagcagca	gcccgtgatc	cccctctctc	gccgccggtg	1500
acgctgtggt	ggcatcggtc	ctgctgcggg	cagaataagt	ctggcgatcat	cgccctatcg	1560
cctggagctg	caaccaccac	taccggggcc	atcgatcgtc	tagagcggtta	tccaccctgc	1620
ctgccccatt	acttgcagct	ccggctgggt	cagaacctct	ccatgcctga	caaattgggt	1680
caagattgtc	gctgtccggc	cagcgcttga	attttcagaa	tatgccatcg	aatacgcgct	1740
gcttttaaga	tatgctaccc	gattcggtgt	aatttttagaa	tacgccatcg	gaacacgaat	1800
tttcttcgtt	ccgtgccact	ccgtctctcg	gagtcagtcg	tgccgtcgtc	atccgtccgc	1860
ccagcactgt	cgtcgtcagt	ccgccacccg	tgccgtgactg	cccgttcagc	tgccgccgtcg	1920
tccgtccgtc	gccgccatcg	ccgtcggtcc	ctgtcgcgcc	cgcacctgca	cccgtgtcag	1980
gcgcgccgtc	gtccgtccgc	gttcaccatc	atcgctcgcc	actgcggcga	gcgcagacgg	2040
ctgtggacgg	atgattgcgc	ggcaagcgca	agcggcggtg	gacggacgac	ggggcagtgc	2100
acgcgagcac	ggtagccgat	ggactgacgt	tggcggcatt	gggagatgga	cgacgacggc	2160
atgattgacg	tggggaacgg	aatgtcacgg	awcggagaaa	attcgcactc	cggtggcgta	2220
tattccgaaa	atagcacgaa	ttgggtggca	tatcctaaaa	acagtgccta	ttcggtggca	2280
tattctaaaa	attctcggtc	agcataatcc	ccatcaatcc	ccaatccctc	aacagttggg	2340
ttaatatatt	tggagatgtg	ttcggttggt	taggttgaag	ttctccactt	cacctccatg	2400
agtacatgca	cctctacacg	tacgttctta	atgagtttgt	ttgttctatc	ctccgcttgg	2460
gttctatatt	gttggttccg	atctgatttg	atctggagcg	gggtcgatct	tccacgacgg	2520
cgagagagac	gttggttcggg	ctgctcgatt	tggttcaact	gtttaggtcg	aaggaggggg	2580
tagaattgca	attcaagtgc	atggtcagtc	aatttggtgc	aaaattaaca	tcaaaactggg	2640
ataaagagac	gacagtgcc	aaattggtaa	cgggaaaact	ttgagttcta	ttaaagttaa	2700
ccagtaactt	tcgttgctat	agaataaaaa	cgtaaaacttt	tgatgctcct	tggttgagct	2760

tggtacaaac	caaacagacg	taaaataaac	actatcctga	atcaagtcta	ctaagttcca	2820
ttgaactcaa	ccaggatacg	tacacttcct	cttagaagat	gtcttggttt	cactttgtac	2880
aatTTTTtct	attgtaaatt	tggtacctcg	ttgtacctag	gtacaagagg	taccatgaga	2940
taccaaattt	tactactaaaa	ttttggtacc	tcattggtacc	tcctcaacga	ccgtagaatt	3000
gctcttaatt	taattttaaaa	aaaaacataa	tattttttaa	gcatattatg	gaaatttttag	3060
taattattac	ttttgtaata	tatgagttac	ggttatactc	gagatatacct	aaattgcttg	3120
gagatgaata	attacaaggt	atatcaaaga	tgagttgaaa	ataatgcagg	cattccttag	3180
agctgcagaa	gttatgaaaa	agaaagacga	actattaaag	gtttgggtag	agcaaatacg	3240
tgacctgtca	tatgacattg	aagattccct	tgatgaattt	aagggtccatg	ttgaaagtca	3300
aaccctatTT	cgtcagttgg	tgaaacttag	atagcgtcac	cggatcgcta	gcccggtggat	3360
gttttagttgt	ttgcacatgg	tgctggatgg	tgcgctcatg	gtcttggttg	aggtctggta	3420
ccaaccagtc	atgcttagaa	atagccggat	cagtgacagg	tgctaggact	ttacttggtg	3480
gtctgtgcag	cgctatcgac	atgtggtggg	gtgctTTTT	ttttccggat	tacaatctca	3540
tagggctaca	ctctagttat	tttgctgcta	tattaatatg	aaaacttggt	atggttcggt	3600
tcttatagaa	aaaaacctag	ttgatcaagg	gctagttttc	ttcaagtgca	ttcctaatact	3660
tagcttcttt	tttttttgca	atggcaagaa	ttgttcatta	aaaaaattga	taaaaattgg	3720
ctaggcctac	gttttgtttc	ttaccaaagt	tgtactttta	caataaacta	aggcaaatat	3780
ttcggcaatg	ccattttctt	gtctacagac	caaatatatg	gctaaatttt	ggcataacca	3840
tttttttggt	tgcttggttg	agcttggtac	aaaccaaaca	gacccaaaat	aaacagtgtc	3900
atgaatcacg	tctactaaat	tcctttgaac	tgaactagaa	tatagttgct	cttaaaagat	3960
ttcttgattt	cactcggtac	cattttactag	tacaaactta	gatttaattt	ttaaaaataa	4020
aatcataata	ttgttattat	ggaaaattta	gtcatagtac	ttttgtaata	tatgagatgg	4080
gttatacttg	agatataccta	aattgcttta	agatgaataa	ttgctaggta	tatcaaagat	4140
gagctaaaaa	cgatgcaagc	attccttaga	gctgctgaac	ttatgaaaaa	gaaagatgaa	4200
ctattaaagg	tttgggcaga	gcaaatacgt	gacctgtcat	atgacattga	agattccctt	4260
gatgaattta	aggtccatat	tgaaagccaa	accctatttc	gtcagttggt	gaaactcaga	4320
gaacgccacc	gaattgctat	ccgtatccac	aaccttaaat	caagagttga	agaagtgagt	4380
agcaggaaca	cacgctacag	tttagtcaag	cctatttcct	ctggcacaga	gattgacatg	4440
gattcctatg	cagaagacat	tcgtaatcag	tcagctcaca	atgtggatga	agctgagctt	4500
gttgggtttt	ctgactccaa	gaaaagggtg	cttgaaatga	tcgataccaa	tgctaattgat	4560
ggtcgggcca	aagtaatctg	tgttggtggg	atgggtggtt	taggcaagac	agctctttcg	4620
aggaagatct	ttgaaagcga	agaagacatt	aggaagaact	tcctttgcaa	tgcttggtgatt	4680
acagtgtcac	aatcatttca	caggattgag	ctactcaaag	atatgatacg	ccaacttctt	4740
ggccccagtt	ctctggatca	actcttgcaa	gaattgcaag	ggaagggtgg	ggtgcaagta	4800
catcatcttt	ctgagtacct	gatagaagag	ctcaaggaga	agaggtactt	tgttgttcta	4860
gatgatctat	ggattttaca	tgattggaat	tgataaatg	aaattgcatt	tcctaagaac	4920

aataagaagg	gcagtcgaat	agtaataacc	actcggaatg	ttgatcttgc	ggagaagtgt	4980
gccacagcct	cactggtgta	ccaccttgat	ttcttgcaga	tgaacgatgc	cataacattg	5040
ctactgagaa	aaacaaataa	aaatcatgaa	gacatggaat	caaataaaaa	tatgcaaaag	5100
atggttgaac	gaattgtaaa	taaatgtggt	cgtctaccat	tagcaatact	tacaatagga	5160
gctgtgcttg	caactaaaca	ggtgtcagaa	tgggagaaat	tctatgaaca	acttccttca	5220
gaactagaaa	taaacccaag	cctggaagct	ttgaggagaa	tggtgaccct	aggttacaac	5280
cacctacat	cccatctgaa	accatgcttt	ttgtatctaa	gtatctttcc	tgaggatttt	5340
gaaatcaaaa	ggaatcgtct	agtaggtata	tggatagcag	aagggttggt	tagaccaaag	5400
gttgggatga	cgactaagga	tgtcggagaa	agttactttt	atgagctaata	caaccgaagt	5460
atgattcaac	gatcaagagt	gggcatagca	ggaaaaatta	agacttgctg	aattcatgat	5520
atcatccgtg	atatcacagt	ttcaatctcg	agacaggaaa	attttgtatt	attaccaatg	5580
ggagatggct	ctgatttagt	tcaggaaaac	actcgccaca	tagcattcca	tgggagtatg	5640
tcctgcaaaa	caggattgga	ttggagcatt	attcgatcat	tagctatttt	tggtgacaga	5700
ccaagagtc	tagcacatgc	agtttgtcca	gatcaattga	ggatgttacg	ggtcttggat	5760
cttgaagatg	tgacattctt	aatcactcaa	aaagatttcg	accgtattgc	attgttgtgc	5820
cacttgaaat	acttgagtat	tggatatctg	tcatccatat	attcacttcc	cagatccatt	5880
ggtaaactac	agggcctaca	aactttgaac	atgccgagca	catacattgc	agcactacca	5940
agtgagatca	gtaaactcca	atgtctgcat	actcttcgtt	gtataggaca	gtttcattat	6000
gacaacttta	gtctaaacca	ccaatgaag	tgcataacta	acacaatatg	cctgcctaaa	6060
gtattcacac	ctttagttag	tcgcatgat	cgtgcaaaac	aaattgctga	attgcacatg	6120
gccaccaaaa	gttgctggtc	tgaatcattc	ggtgtgaagg	tacccaaagg	aataggttaag	6180
ttgcgagact	tgcaggttct	agagtatgta	gatatcaagc	ggaccagtag	tagagcaatc	6240
aaagagctgg	ggcagttaag	caagttgagg	aaattaggtg	tgataacaaa	aggctcgaca	6300
aaggaaaaat	gtaagatact	ttatgcagcc	attgagaagc	tctcttccct	ccaatatctc	6360
tatgtgaatg	ctgcgttatt	atcagatatt	gaaacacttg	agtgcctaga	ttctattttca	6420
tctcctcctc	ccctactaag	tacactcagg	ttgaatggaa	gtcttgaaga	gatgcctaac	6480
tggattgagc	agctcactca	cctgaagaag	ttctacttac	ggaggagcaa	actaaaggaa	6540
ggtaaaacca	tgctgatact	tggggcactg	cccaacctca	tgttcctttc	tctttatcat	6600
aattcttata	ttggggagaa	gctagtattc	aaaacgggag	cattcccaaa	tcttagaaca	6660
ctttgtattt	acgaattgga	tcagctaaga	gagatcagat	ttgaggacgg	cagctcaccc	6720
ctggttgaaa	agatagaaat	aggcaagtgc	aggttggaat	ctgggattat	tggtatcatt	6780
caccttccaa	agctcaagga	gattccaatt	acatacggaa	gtaaagtggc	tgggcttggg	6840
cagctggagg	gagaagtga	cgcacacca	aatcgccccg	tgctgctaata	ggacagtgac	6900
cgaaggtatc	acgacctggg	ggctgaagcc	gaaggatctt	ctatagaagt	gcaaacagca	6960
gatcctgttc	ctgatgccga	aggatcagtc	actgtagcag	tggaagcaac	ggatcccctt	7020
cccgagcagg	agggagagag	ctcgcagtcg	caggtgatca	cgttgacgac	gaatgatagg	7080

tcagtcactc cctagatggc agcttaatta acttgtttct aattctcttc ttgttcagta 7140
 ttagccatca ggtgagggcg atgatttcaa ctcaatttct atctctctcg ttttcttaac 7200
 ctgacagcga agagataggc acagctcaag ctggctga 7238

<210> 5

<211> 21

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi xuôi F1

<400> 5

atggcggaga cgggtgctgag c 21

<210> 6

<211> 20

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi ngược R1

<400> 6

tcagccagct tgagctgtgc 20

<210> 7

<211> 40

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi xuôi F2

<400> 7

acgaattcga gctcggtacc cttgacatcc aaaccgcacc 40

<210> 8

<211> 41

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi ngược 2

<400> 8

agtgcgaagc ttccggcgcg ctcagccagc ttgagctgtg c

41

<210> 9

<211> 20

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi xuôi F3

<400> 9

ggtatatgga tagcagaagg

20

<210> 10

<211> 20

<212> ADN

<213> trình tự nhân tạo

<220>

<223> Mỗi ngược R3

<400> 10

gcgattaagt tgggtaacgc

20