



(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



1-0023245

(51)⁷ C07K 14/415; C12N 5/14; C12N 15/29; (13) B
C12N 15/82; A01H 1/00; C12N 1/21

(21) 1-2012-03292 (22) 31/03/2011
(86) PCT/CN2011/000558 31/03/2011 (87) WO2011/127744 20/10/2011
(30) 201010146613.8 12/04/2010 CN
(45) 27/04/2020 385 (43) 25/04/2013 301A
(73) INSTITUTE OF GENETICS AND DEVELOPMENTAL BIOLOGY, CHINESE
ACADEMY OF SCIENCES (CN)

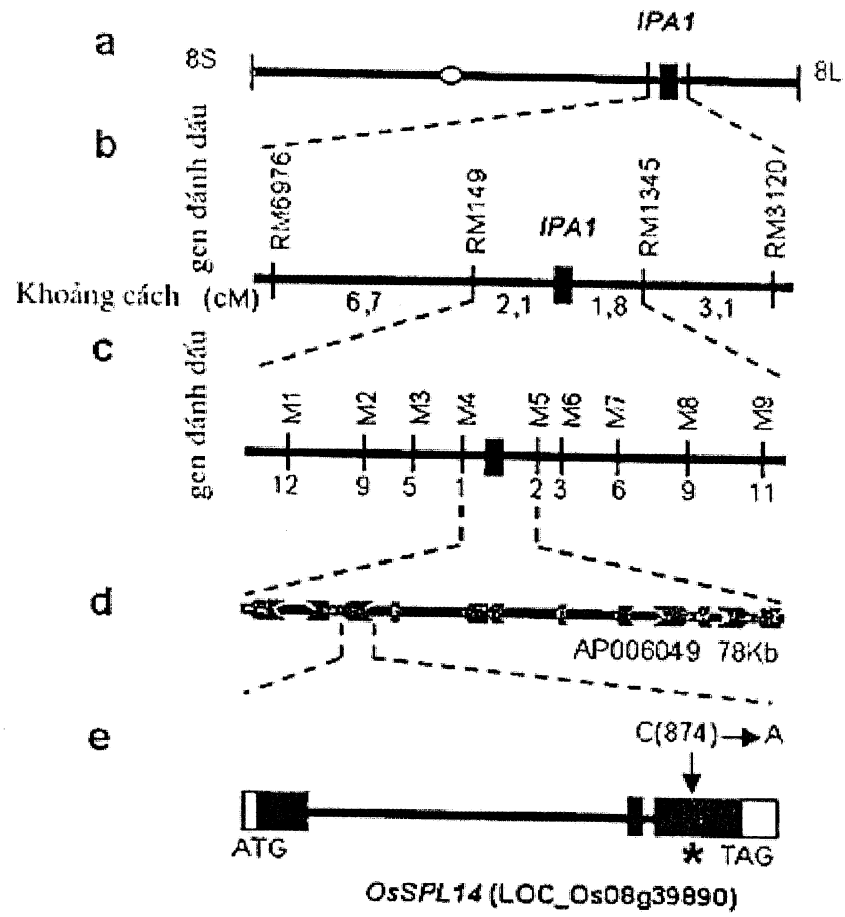
No. 2, Yard No.1 West Beichen Road, Chaoyang District Beijing 100101, China

(72) LI, Jiayang (CN); QIAN, Qian (CN); WANG, Yonghong (CN); JIAO, Yongqing
(CN); XUE, Dawei (CN); LIU, Guifu (CN); WANG, Jing (CN); DONG, Guojun (CN)

(74) Công ty cổ phần tư vấn Trung Thực (TRUNG THUC.,JSC)

(54) PHƯƠNG PHÁP LAI TẠO THỰC VẬT CHUYỂN GEN

(57) Sáng chế đề xuất gen mã hóa protein IPA1 liên quan đến thực vật và phương pháp lai tạo cây chuyển gen bằng cách sử dụng nó, trong đó protein là 1) hoặc 2) sau: 1) protein chứa trình tự axit amin là SEQ ID NO: 1 nêu trong phần danh mục trình tự; 2) protein thu được từ protein nêu trong mục 1) bằng cách thay thế và/hoặc làm khuyết đoạn và/hoặc bổ sung một hoặc vài gốc axit amin trong trình tự axit amin được xác định trong mục 1) và liên quan đến cấu trúc thực vật. Gen IPA1 có thể được sử dụng để lai tạo với sự trợ giúp của gen đánh dấu phân tử để tạo ra các giống lúa mới và tăng năng suất lúa.



Lĩnh vực kỹ thuật được đề cập

Sáng chế thuộc lĩnh vực kỹ thuật gen thực vật, cụ thể là đề cập đến protein IPA1 liên quan đến cấu trúc thực vật, gen mã hóa nó và phương pháp lai tạo cây chuyển gen bằng cách sử dụng nó.

Tình trạng kỹ thuật của sáng chế

Cấu trúc thực vật của cây lúa bao gồm các tính trạng như số lượng chồi gốc, góc chồi gốc, kiểu bông và chiều cao cây. Cấu trúc thực vật tốt là yếu tố then chốt giúp tăng năng suất lúa. Đa số các cây trồng dùng trong sản xuất hiện nay là các giống lúa lùn chứa gen nửa lùn SD1. Các giống lúa lùn có nhiều ưu điểm so với các giống lúa thân cao truyền thống, do đó “cuộc cách mạng xanh” đầu tiên đã được thiết lập và năng suất lúa đã tăng lên nhiều. Tuy nhiên, việc tăng năng suất hơn nữa bị hạn chế bởi các nhược điểm vốn có của các giống lúa lùn, bao gồm số lượng chồi gốc không sinh sản nhiều hơn, bông nhỏ hơn và chỉ số diện tích lá lớn hơn, hiện tượng che bóng lá nghiêm trọng và mức độ quang hợp trên tán lá giảm và các nhược điểm khác. Để khắc phục các nhược điểm về năng suất hạn chế của hầu hết các cây trồng hiện nay cũng như đáp ứng nhu cầu lương thực của con người, các nhà lai tạo thuộc Viện Nghiên cứu Lúa gạo Quốc tế (International Rice Research Institute) đã đưa ra khái niệm kiểu hình mới của cây lúa và các đặc điểm chính của kiểu hình mới này là số lượng chồi gốc ít hơn, không có chồi gốc không sinh sản, bông lớn hơn, số lượng hạt trên một bông nhiều hơn, thân cứng và mập và có khả năng chống đổ.

Số lượng chồi gốc của cây lúa là tính trạng nông học quan trọng trong sản xuất lúa gạo. Số lượng chồi gốc sinh sản trên một đơn vị diện tích quyết định số lượng bông, mà đến lượt mình nó là một trong ba yếu tố quan trọng để

quyết định năng suất lúa cho mỗi đơn vị diện tích. Do vậy, việc kiểm soát hợp lý quá trình đẻ chồi của cây lúa và việc giảm đến mức tối thiểu số lượng chồi gốc không sinh sản là cực kỳ quan trọng trong sản xuất lúa.

Số lượng hạt trên một bông lúa là một yếu tố quan trọng khác để xác định năng suất lúa. Đặc điểm điển hình của hầu hết các giống cho năng suất cao hiện được sử dụng trong sản xuất là sự gia tăng đáng kể số lượng hạt trên một bông. Sự gia tăng số lượng hạt trên một bông chủ yếu là do nhiều nhánh cuống cấp một và nhánh cuống cấp hai trên một bông và các hạt phát triển dày hơn. Việc gia tăng số lượng hạt trên một bông là rất quan trọng trong lai tạo các giống cho năng suất cao. Khối lượng của 1000 hạt lúa là yếu tố quan trọng thứ ba đối với năng suất lúa, là dấu hiệu phản ánh trực tiếp về mức độ tích lũy chất khô tốt hay tồi và mức độ chắc của hạt gạo và liên quan mật thiết đến kích thước của hạt lúa.

Khả năng chống đổ luôn là một khía cạnh mà các nhà lai tạo lúa cho là cực kỳ quan trọng. Nó đóng một vai trò rất quan trọng trong việc làm ổn định năng suất lúa và là yếu tố hạn chế để gia tăng tiếp theo năng suất lúa, mặc dù nó không trực tiếp làm tăng năng suất lúa. Các giống lúa lùn có khả năng chống đổ tăng do chiều cao cây giảm và vì vậy đảm bảo việc sản xuất lúa gạo ổn định và khiến cho năng suất lúa có thể tăng, so với các giống lúa thân cao truyền thống. Vì vậy, việc gia tăng khả năng chống đổ hơn nữa của cây lúa là yếu tố quyết định để gia tăng hơn nữa năng suất lúa. Liên quan đến vấn đề này, mục đích của các nhà lai tạo luôn luôn là nâng cao tính trạng của thân và để lai tạo được giống có thân mập và cứng hơn và khả năng chống đổ của chúng là cao hơn.

Các đặc tính cơ bản của kiểu hình cây lúa mới được đề xuất bởi Viện Nghiên cứu lúa gạo quốc tế là số lượng chồi gốc ít hơn, thân mập hơn và bông

lớn hơn. Các nghiên cứu mô phỏng đã gợi ý rằng năng suất của giống theo kiểu hình mới tăng khoảng 25% so với các giống hiện có vào mùa khô ở vùng nhiệt đới. Điều quan trọng là giải thích cơ sở di truyền và cơ chế phân tử của việc tạo chồi gốc, sự phát triển của thân và bông nhằm có được giống lúa có năng suất cao hơn. Hiện nay, chưa có thông báo về các gen có thể thay đổi một cách có hệ thống kiểu hình của cây lúa nói chung để tạo ra các đặc tính kiểu hình mới, mặc dù nhiều gen liên quan đến năng suất lúa đã được tách dòng.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là nhằm thu được protein liên quan đến cấu trúc thực vật và gen mã hóa nó.

Protein liên quan đến cấu trúc thực vật theo sáng chế có tên là IPA1, thu được từ cây lúa (*Oryza sativa* L.), là protein được thể hiện trong phần 1) hoặc 2) dưới đây:

1) protein chứa trình tự axit amin như SEQ ID NO: 1 nêu trong phần danh mục trình tự; và

2) protein liên quan đến cấu trúc thực vật thu được từ trình tự axit amin như SEQ ID NO: 1 nêu trong phần danh mục trình tự bằng cách thay thế và/hoặc làm khuyết đoạn và/hoặc bổ sung một hoặc nhiều gốc axit amin.

Để tạo điều kiện thuận lợi cho việc tinh chế IPA1 nêu trong phần 1), đầu tận cùng amino hoặc đầu tận cùng carboxyl của protein bao gồm trình tự axit amin như SEQ ID NO: 1 nêu trong phần danh mục trình tự có thể được bổ sung các nhãn như được liệt kê trong Bảng 1.

Bảng 1. Trình tự của đuôi

Đuôi	Số gốc	Trình tự
Poly-Arg	5 đến 6 (thường là 5) (SEQ ID NO: 8)	RRRRR (SEQ ID NO: 8)
Poly-His	2 đến 10 (thường là 6)	HHHHHH

	(SEQ ID NO: 10)	(SEQ ID NO: 11)
FLAG	8	DYKDDDDK (SEQ ID NO: 12)
Strep-tag II	8	WSHPQFEK (SEQ ID NO: 13)
c-myc	10	EQKLISEEDL (SEQ ID NO: 14)

IPA1 như nêu trong mục 2) nêu trên có thể được tổng hợp theo cách nhân tạo; hoặc thu được theo cách khác bằng cách biểu hiện gen mã hóa nó trong vi sinh vật. Gen mã hóa IPA1 như nêu trong mục 2) nêu trên có thể thu được nhờ việc làm mất hoặc bổ sung các codon cho một hoặc nhiều gốc axit amin trong trình tự ADN từ 124 cặp bazơ đến 1377 cặp bazơ ở đầu tận cùng 5' của SEQ ID NO: 2 trong phần danh mục trình tự và/hoặc nhờ một hoặc nhiều đột biến sai nghĩa của cặp bazơ và/hoặc nhờ việc bổ sung các đuôi như nêu trong Bảng 1 vào đầu tận cùng 5' và/hoặc đầu tận cùng 3' của nó.

Gen mã hóa protein liên quan đến cấu trúc thực vật như nêu trên (được đặt tên là gen IPA1) cũng thuộc phạm vi của sáng chế.

Gen mã hóa protein liên quan đến cấu trúc thực vật như nêu trên cụ thể là một trong số các gen được mô tả trong các mục từ 1) đến 5) sau:

1) trình tự mã hóa nó như được liệt kê có chiều dài từ 124 cặp bazơ đến 1377 cặp bazơ ở đầu tận cùng 5' của SEQ ID NO: 2 trong phần danh mục trình tự;

2) trình tự nucleotit của nó là SEQ ID NO: 2 trong phần danh mục trình tự;

3) trình tự nucleotit của ADN di truyền như được liệt kê có chiều dài từ 1 cặp bazơ đến 7229 cặp bazơ ở đầu tận cùng 5' của SEQ ID NO: 3 trong phần danh mục trình tự;

4) gen mà được lai với gen ở phần 1), 2) hoặc 3) trong các điều kiện

nghiêm ngặt và mã hóa protein đã nêu; và

5) gen đồng nhất với gen trong các mục 1), 2) hoặc 3) với mức độ hơn 90% và mã hóa protein này.

SEQ ID NO:2 trong phần danh mục trình tự chứa 1624 bazơ; khung đọc mở (open reading frame - ORF) của nó là từ bazơ 124 đến bazơ 1377 ở đầu tận cùng 5' và nó mã hóa protein IPA1 có trình tự axit amin là SEQ ID NO: 1 trong phần danh mục trình tự.

Các điều kiện nghiêm ngặt nêu trên có thể sử dụng 0,1xSSPE (hoặc 0,1xSSC) và dung dịch SDS 0,1%, lai ở 65°C và rửa màng trong dung dịch thử nghiệm lai hóa ADN hoặc ARN.

Các đoạn môi để khuếch đại gen IPA1 với chiều dài đầy đủ hoặc phần bất kỳ của nó cũng thuộc phạm vi của sáng chế.

Catxet biểu hiện, vật truyền tái tổ hợp, dòng tế bào chuyển gen và vi khuẩn tái tổ hợp chứa gen mã hóa protein liên quan đến kiểu hình nêu trên cũng thuộc phạm vi của sáng chế.

Vật truyền biểu hiện tái tổ hợp chứa gen IPA1 có thể được xây dựng bằng cách sử dụng các vật truyền biểu hiện thực vật hiện có. Vật truyền biểu hiện thực vật nêu trên bao gồm vật truyền *Agrobacterium* nhị thể và vật truyền dùng để bắn phá thực vật và các vật truyền tương tự, như pCAMBIA3301, pCAMBIA 1300, pB11121, pBin19, pCAMBIA2301, pCAMBIA1301-UbiN hoặc các vật truyền khác thu được từ vật truyền biểu hiện thực vật.

Gen khởi đầu tăng cường, gen khởi đầu cơ định, gen khởi đầu đặc hiệu mô hoặc gen khởi đầu cảm ứng bất kỳ, như gen khởi đầu virus khảm súp lơ (cauliflower mosaic virus - CAMV) 35S và gen khởi đầu ubiquitin (pUbi) và các gen khởi đầu tương tự, ở dạng một mình hoặc kết hợp với các gen khởi đầu thực vật khác, có thể được thêm vào nucleotit khởi đầu phiên mã khi sử dụng

gen IPA1 để xây dựng vật truyền biểu hiện tái tổ hợp. Ngoài ra, cũng có thể sử dụng một vùng tăng cường, bao gồm vùng tăng cường dịch mã hoặc vùng tăng cường phiên mã khi sử dụng gen theo sáng chế để xây dựng vật truyền biểu hiện thực vật, các vùng tăng cường này là mã khởi đầu ATG hoặc mã khởi đầu trong vùng liền kề, nhưng chúng phải giống với khung đọc của trình tự mã hóa để đảm bảo dịch mã đúng toàn bộ trình tự. Gốc của tín hiệu điều khiển dịch mã nêu trên và mã khởi đầu có số lượng lớn; chúng có thể là tự nhiên hoặc tổng hợp. Vùng khởi đầu dịch mã có thể thu được từ vùng khởi đầu phiên mã hoặc gen cấu trúc.

Để tạo điều kiện thuận lợi cho việc xác định và sàng lọc tế bào thực vật hoặc thực vật chuyển gen, vật truyền biểu hiện thực vật được sử dụng có thể được cải biến như bổ sung các enzym tạo ra sự đổi màu trong quá trình biểu hiện thực vật hoặc các gen khiến cho các hợp chất phát sáng (gen GUS, gen GFP và gen luciferaza và v.v.), gen kháng thuốc kháng sinh (kháng gentamycin, kháng kanamycin, v.v.) hoặc gen kháng các tác nhân hóa học (như gen dùng cho tác nhân chống bệnh gỉ sắt) và các gen tương tự.

Cụ thể là, vật truyền biểu hiện tái tổ hợp nêu trên có thể là vật truyền biểu hiện tái tổ hợp được tạo ra bằng cách cài xen gen mã hóa protein liên quan đến kiểu hình như nêu trên vào vị trí đa tách dòng của vật truyền biểu hiện thực vật pCAMBIA1300.

Một mục đích khác của sáng chế là đề xuất phương pháp lai tạo thực vật chuyển gen.

Phương pháp lai tạo thực vật chuyển gen theo sáng chế bao gồm bước đưa gen IPA1 mã hóa protein liên quan đến kiểu hình như nêu trên hoặc ADN hệ gen vào thực vật để tạo ra thực vật chuyển gen; thực vật chuyển gen nêu trên có số lượng chồi gốc ít hơn, thân cứng và mập hơn, số lượng bông và nhánh

cuồng nhiều hơn và số lượng hạt trên một bông nhiều hơn so với thực vật đích nêu trên.

Gen IPA1 mã hóa protein nêu trên liên quan đến kiểu hình như nêu trên được đưa vào thực vật đích nhờ vật truyền biểu hiện tái tổ hợp như nêu trên.

Vật truyền biểu hiện thực vật mang gen IPA1 mã hóa protein liên quan đến kiểu hình theo sáng chế có thể được biến nạp vào tế bào hoặc mô thực vật theo các phương pháp sinh học thông thường như Ti plasmit, Ri plasmit, vật truyền virus thực vật, biến nạp trực tiếp ADN, tiêm lượng rất nhỏ, sự tải nạp bằng điện, phương pháp qua trung gian *Agrobacterium* và v.v.. Thực vật chủ đã được biến nạp (thực vật đích) là thực vật hai lá mầm hoặc thực vật một lá mầm, tốt hơn là cây lúa, tốt hơn nữa là giống lúa Nipponbare.

Một mục đích khác của sáng chế là đề xuất phương pháp lai tạo thực vật chuyển gen. Phương pháp này bao gồm bước đưa vật truyền can thiệp vào thực vật đích để thu được thực vật chuyển gen; thực vật chuyển gen nêu trên có số lượng chồi gốc nhiều hơn nhiều, thân gầy hơn, số lượng nhánh cuống cấp một ít hơn và số lượng hạt trên một bông ít hơn so với thực vật đích nêu trên; vật truyền can thiệp này là vật truyền tái tổ hợp thu được bằng cách cài xen tuần tự trình tự nucleotit được liệt kê như SEQ ID NO: 4 nêu trong phần danh mục trình tự và trình tự nucleotit được liệt kê như SEQ ID NO: 5 nêu trong phần danh mục trình tự vào vị trí giữa *Bam*HI và *Kpn*I và vị trí giữa vị trí *Spe*I và vị trí *Sac*I trong vật truyền pTCK303. Thực vật đích là thực vật hai lá mầm hoặc thực vật một lá mầm, tốt hơn là cây lúa, tốt hơn nữa là Ri 22.

SEQ ID NO: 4 là một mảnh có chiều dài từ 1014 cặp bazơ đến 1623 cặp bazơ của SEQ ID NO: 2 và SEQ ID NO: 5 là trình tự bổ trợ ngược chiều của SEQ ID NO: 4. Đã khẳng định được rằng không có trình tự tương đồng nào khác với SEQ ID NO: 4 và 5 trong hệ gen cây lúa khi phân tích so sánh toàn bộ

bộ gen.

Gen IPA1 đa tính trạng có thể điều khiển số lượng chồi gốc và sự phát triển của thân và bông được phân lập bằng cách áp dụng phương pháp tách dòng dựa trên bản đồ gen, và chức năng của gen này được khẳng định thông qua thử nghiệm bổ trợ chức năng.

Thử nghiệm này đã chứng minh được rằng số lượng chồi gốc của cây lúa giảm, thân trở nên mập và cứng, số lượng bông và nhánh cuống tăng lên và số lượng hạt trên một bông tăng sau khi gen theo sáng chế được biểu hiện quá mức ở cây lúa; chiều cao cây giảm, số lượng chồi gốc tăng, thân trở nên gầy và số lượng bông và nhánh cuống giảm sau khi gen theo sáng chế mất chức năng hoặc hoạt tính của gen này bị giảm, điều đó gợi ý rằng gen này có thể điều khiển cấu trúc thực vật của cây lúa. Do vậy, gen IPA1 tạo ra công cụ mạnh để lai tạo một giống lúa với kiểu hình mới bằng cách áp dụng phương pháp lai tạo có sự trợ giúp của gen đánh dấu phân tử và bằng cách áp dụng kỹ thuật di truyền, nhờ đó tiếp tục tăng năng suất lúa, và gen này có ý nghĩa quan trọng về mặt lý thuyết và có tiềm năng ứng dụng lớn.

Mô tả vắn tắt các hình vẽ

Hình 1 thể hiện kiểu hình của cây lúa, bao gồm Shaoniejing (SNJ) với số lượng chồi gốc ít và giống lúa Indica thông thường TN1.

Hình 2 là sơ đồ thể hiện việc tách dòng dựa trên bản đồ gen của gen IPA1. Hình 2a và 2b là các hình vẽ thể hiện phương pháp phân tích QTL và định vị bằng cách sử dụng BC_2F_2 . Hình 2c thể hiện bản đồ định vị chính xác; số nằm bên dưới nhãn biểu thị cá thể tái tổ hợp; Hình 2d thể hiện gen dự đoán khoảng 78kb, và mũi tên thể hiện gen dự đoán; Hình 2e là sơ đồ thể hiện cấu trúc của gen *IPA1* theo sáng chế, hộp rộng màu trắng thể hiện vùng chưa dịch mã ở đầu tận cùng 5' và 3', hộp màu đen thể hiện một exon, đường nằm ngang

thể hiện một intron và dấu hoa thị màu đỏ thể hiện vị trí đích miARN156. Sự thay đổi bazơ phía trên hộp thể hiện đột biến bazơ trong vật liệu Shaoniejin. Số ở trong ngoặc đơn thể hiện vị trí đột biến bazơ.

Hình 3 thể hiện biểu đồ trình tự của cADN (SEQ ID NO: 2) và protein (SEQ ID NO: 1) của gen IPA1, nucleotit màu xanh da trời thể hiện vùng chưa mã hóa 5' và 3', trình tự protein được gạch chân là miền cấu trúc SBP, dấu hoa thị màu đỏ thể hiện vị trí đích miARN156 và chữ cái màu đỏ thể hiện đột biến nucleotit ở vật liệu Shaoniejing và sự thay đổi axit amin do đột biến.

Hình 4 thể hiện bản đồ vật truyền gIPA1 và phân tích thống kê về kiểu hình và tính trạng nông học của cây lúa chuyển gen trong thử nghiệm hỗ trợ chức năng. Hình 4a thể hiện bản đồ vật truyền gIPA1; Hình 4b thể hiện kiểu hình của cây lúa chuyển gen gIPA1; Hình 4c thể hiện mức độ biểu hiện của IPA1 được xác định bởi phương pháp RT-PCR; Hình 4d thể hiện so sánh về mặt thống kê của các tính trạng nông học liên quan đến vật liệu chuyển gen gIPA1 nhờ thử nghiệm T, dấu hoa thị đơn thể hiện tính trạng quan trọng và dấu hoa thị kép thể hiện tính trạng rất quan trọng; Nipponbare là mẫu đối chứng kiểu đại và gIPA1 là thực vật chuyển gen trên Hình 4b, 4c và 4d.

Hình 5 thể hiện so sánh kiểu hình và các tính trạng nông học của cây lúa chuyển gen thu được bằng cách can thiệp bằng ARN, Hình 5a thể hiện kiểu hình của cây lúa chuyển gen thu được bằng cách can thiệp bằng ARN; Hình 5b thể hiện mức độ biểu hiện của IPA1 trên thực vật chuyển gen xác định được bởi phương pháp RT-PCR; Hình 5c thể hiện so sánh các tính trạng nông học liên quan đến cây lúa chuyển gen thu được bằng cách can thiệp bằng ARN nhờ thử nghiệm T-test và dấu hoa thị kép thể hiện tính trạng rất quan trọng; Ri 22 thể hiện mẫu đối chứng không chuyển gen và IPA1-ARNi thể hiện thực vật chuyển gen trên các Hình 5a, 5b và 5c.

Mô tả chi tiết sáng chế

Các phương án nêu dưới đây được đưa ra nhằm minh họa rõ hơn sáng chế, tuy nhiên sáng chế không giới hạn ở các phương án này.

Các phương pháp nêu trong các ví dụ dưới đây đều là các phương pháp thông thường, trừ khi có quy định khác.

Cây lúa trong các ví dụ nêu dưới đây thu được nhờ phương pháp trồng lúa sau: (1) cấy lúa trên cánh đồng: hạt lúa được ngâm trong nước trong 2 ngày, tiếp theo được chuyển vào buồng nuôi cấy ở 37°C để thúc đẩy sự nảy mầm của hạt lúa trong 3 ngày, sau đó hạt lúa có mầm màu trắng được gieo vào một luống để tạo ra mạ và mạ được cấy trên cánh đồng khi có 4 lá.

Ví dụ thực hiện sáng chế

Ví dụ 1. Phát hiện gen

Hạt giống Shaoniejing với số lượng chồi gốc ít hơn, thân cứng và bông lớn và giống lúa Indica thông thường TN1 (*Oryza sativa* L.) được trồng theo phương pháp canh tác trên đồng ruộng như nêu trên và hình thái học của cây lúa trưởng thành được thể hiện trên Hình 1. ADN được phân lập từ lá.

Phân lập ADN hệ gen của cây lúa:

ADN hệ gen được phân lập bằng cách áp dụng phương pháp CTAB đã được làm thích ứng (Mou Z, He Y, Dai Y, et al. Deficiency in fatty acid synthase leads to premature cell death and dramatic alterations in plant morphology. *The Plant Cell*. 2000, 12, 405-418.) từ lá lúa. 100mg lá lúa được gom lại và làm lạnh trong nitor lỏng và được xay thành bột trong một cối nhỏ có đường kính 5cm, được chuyển vào ống ly tâm dung tích 1,5ml để phân lập ADN và ADN thu được được hòa tan vào 100 μ MQ H₂O.

Việc tách dòng dựa trên bản đồ gen được thực hiện theo các bước sau:

1. Lập bản đồ sơ bộ của gen IPA1

Để phân lập gen IPA1, quần thể lập bản đồ trước tiên được xây dựng bằng cách sử dụng Shaoniejing và TN1. Tiếp theo, việc phân tích QTL và lập bản đồ được thực hiện bằng cách sử dụng cá thể BC_2F_2 của quần thể phân tách này. Kết quả lập bản đồ cho thấy IPA1 được sắp xếp chủ yếu giữa hai gen đánh dấu RM149 và RM1345 trên nhiễm sắc thể 8 và khoảng cách di truyền lần lượt là 2,1cm và 1,8cm (Hình 2a và hình 2b).

2. Lập bản đồ chính xác của gen IPA1

Để thu hẹp hơn nữa vùng xác định thuộc gen đích, 5500 cá thể thực vật có kiểu hình gần với TN1 được chọn từ quần thể phân tách BC_2F_2 để tiến hành lập bản đồ chính xác. Đồng thời, nghiên cứu vị trí phân biệt trong vùng lập bản đồ nhờ trình tự của hệ gen giữa cây lúa Indica và cây lúa Japonica bằng cách sử dụng 93-11 và Nipponbare được công bố để khai thác các gen đánh dấu STS và SSR mới. Một cá thể thực vật có gen được trao đổi với gen đích được sàng lọc trước tiên bằng cách sử dụng các gen đánh dấu RM149 và RM1345 và sau đó các cá thể thực vật đã được trao đổi gen này được sàng lọc bằng cách sử dụng một gen đánh dấu phân tử mới để phát hiện ra rằng các gen đánh dấu phân tử M3 và M6 được liên kết chặt chẽ với gen đích khi tiến hành lập bản đồ chính xác. Các thể tái tổ hợp 5 và 3 lần lượt được xác định. Cuối cùng, locut IPA1 được lập bản đồ chính xác trong đoạn khoảng 78kb nằm giữa gen đánh dấu M4 và M5 (Hình 2c và hình 2d).

3. Xác định và phân tích trình tự của các gen được xem xét

Việc dự đoán gen được xem xét được thực hiện trên đoạn 78kb và tiến hành so sánh trình tự giữa các vật liệu TN1 và Shaoniejing. Kết quả chỉ ra rằng một đột biến điểm từ C thành A xảy ra ở exon thứ ba của gen OsSPL14 (LOC_0s08g39890) trong vật liệu Shaoniejing. Tuy nhiên, không có đột biến nào trong vùng tương tự của vật liệu Nipponbare và Zhonghua11 và các vật

liệu tương tự. Bằng cách phân tích sâu hơn, các tác giả sáng chế đã phát hiện ra rằng vị trí đột biến được định vị ở vị trí đích của miARN156 và vì vậy nó có thể ảnh hưởng đến sự điều hòa của miARN156 trên OsSPL14 (Hình 2e). Trên cơ sở thông tin nêu trên, gen này được xác định là gen được xem xét. Trong cơ sở dữ liệu KOME có một trình tự cADN với chiều dài đầy đủ tương ứng của gen được xem xét này là AK107191. Việc phân tích trình tự protein đã chỉ ra rằng gen được xem xét này chứa một miền SBP bảo toàn (protein gắn kết gen khởi đầu SQUAMOSA) (Hình 3).

Ví dụ 2. Thu nhận và kiểm tra thực vật chuyển gen

1. Thu nhận thực vật chuyển gen

(1) Xây dựng vật truyền biểu hiện tái tổ hợp

(a) Tách dòng gen

Một mảnh ADN chứa gen IPA1 (Việc phân tích trình tự đã chỉ ra rằng trình tự nucleotit của mảnh đã được chuẩn bị này là SEQ ID NO: 3 nêu trong phần danh mục trình tự) được khuếch đại từ ADN hệ gen của Nipponbare (Công chúng có thể có được tài liệu này từ Viện Di truyền và sinh học phát triển, Viện Khoa học Trung Quốc (Institute of Genetics and Developmental Biology, Chinese Academy of Science) và tài liệu này được nêu trong tư liệu phi sáng chế: Lin H, Wang R, Qian Q, et al. *DWARF27, an iron-containing protein required for the biosynthesis of strigolactones, regulated rice tillers bud outgrowth. Plant Cell* 2009, 21, 1512-1525). Mảnh gen chuẩn bị này được phân cắt bởi *Kpn* I và *Xba* I để thu được mảnh ADN hệ gen cuối cùng chứa IPA1 với chiều dài đầy đủ (trình tự nucleotit được liệt kê là 1 cặp bazơ đến 7229 cặp bazơ tính từ đầu tận cùng 5' của SEQ ID NO: 3 nêu trong phần danh mục trình tự).

Trong ADN hệ gen được liệt kê là từ 1 cặp bazơ đến 7229 cặp bazơ của

SEQ ID NO: 3, từ 1118 cặp bazơ đến 1566 cặp bazơ là exon thứ nhất, từ 1567 cặp bazơ đến 3996 cặp bazơ là intron thứ nhất, từ 3997 cặp bazơ đến 4130 cặp bazơ là exon thứ hai, từ 4131 cặp bazơ đến 4232 cặp bazơ là intron thứ hai và từ 4233 cặp bazơ đến 4903 cặp bazơ là exon thứ ba.

Trình tự nucleotit của cADN tương ứng với ADN hệ gen được liệt kê là từ 1 cặp bazơ đến 7229 cặp bazơ của SEQ ID NO: 3 được thể hiện là SEQ ID NO: 2 nêu trong phần danh mục trình tự. SEQ ID NO: 2 này có 1624 cặp bazơ; khung đọc mở (ORF) của nó là từ bazơ 124 đến bazơ 1377 ở đầu tận cùng 5' và nó mã hóa protein IPA1 có trình tự axit amin là SEQ ID NO: 1 nêu trong phần danh mục trình tự.

Bảng 2. Trình tự của các đoạn môi

Tên của các đoạn môi	Trình tự của các đoạn môi (theo chiều từ đầu tận cùng 5' đến đầu tận cùng 3')
gIPA11F	AGGTACCGCAATGTAGAGCCACGTAGGCAAG (SEQ ID NO: 15)
gIPA11R	AGGGCCCGCTTACCAGCTATTGGTTACACATATT (SEQ ID NO: 16)

Vị trí phân cắt enzym *Kpn* I được bổ sung vào đầu tận cùng 5' của đoạn môi gIPA11F và vị trí phân cắt enzym *Apa* I được bổ sung vào đầu tận cùng 5' của đoạn môi gIPA11R.

(b) Xây dựng vật truyền biểu hiện

Mảnh cuối cùng của ADN hệ gen chứa gen *IPA1* với chiều dài đầy đủ thu được ở bước (a) được cài xen vào vị trí nằm giữa vị trí *Kpn* I và vị trí *Xba* I của vật truyền pCAMBIA1300 (mua được từ công ty Cambia) để thu được vật truyền biểu hiện tái tổ hợp gIPA1 (Hình 4a) và đã khẳng định rằng việc xây dựng này là đúng.

(2) Thu nhận cây lúa chuyển gen

Plasmid gIPA1 được biến nạp vào *Agrobacterium tumefaciens* dòng EHA05 (Công chúng có thể có được vật liệu này từ Viện Di truyền và sinh học phát triển, Viện Khoa học Trung Quốc (Institute of Genetics and Developmental Biology, Chinese Academy of Science) và vật liệu này đã được thông báo trong tư liệu phi sáng chế: Lin H, Wang R, Qian Q, et al. DWARF27, an iron-containing protein required for the biosynthesis of strigolactones, regulated rice tillers bud outgrowth. Plant Cell 2009, 21, 1512-1525) bằng cách tạo lỗ bằng xung điện để sàng lọc các chủng *Agrobacterium tumefaciens* tái tổ hợp chứa plasmid gIPA1 tái tổ hợp.

Mô sẹo của Nipponbare được gây nhiễm chủng *Agrobacterium tumefaciens* tái tổ hợp chứa plasmid gIPA1 tái tổ hợp, được nuôi cấy ở 25°C trong bóng tối trong 3 ngày và mô sẹo và thực vật chuyển gen có tính kháng được sàng lọc trên môi trường chọn lọc chứa 50mg/L hygromycin. Cây kháng hygromycin được làm thích nghi trong bóng râm và cấy trên cánh đồng để thu được thực vật chuyển gen thế hệ T₀. Hạt giống thuộc cây thế hệ T₀ được thu hoạch và được trồng theo phương pháp trồng trên cánh đồng như nêu trên và thu được thực vật chuyển gen T1 với gen gIPA1 đã được biến nạp theo phương pháp phát hiện phân tử thông thường.

Nipponbare được biến nạp bằng vật truyền rỗng pCAMBIA1300 để thu được cây đối chứng với vật truyền rỗng theo phương pháp thu nhận thực vật chuyển gen T1 với gIPA1 đã được biến nạp.

2. Xác định kiểu hình của thực vật chuyển gen

(1) Xác định mức độ biểu hiện của gen IPA1 nhờ RT-PCR

ARN tổng số của cây lúa được phân lập từ thực vật chuyển gen và cây đối chứng Nipponbare bằng cách sử dụng TRIZOL (mua được từ Invitrogen

Company) và việc phiên mã ngược được thực hiện bằng cách sử dụng kit phiên mã ngược (mua được từ Công ty Promega) để thu được cADN. Sự biểu hiện của gen IPA1 được phát hiện nhờ PCR bằng cách sử dụng các đoạn mồi IPA1RT1F và IPA1RT1R (Trình tự của mảnh đã được khuếch đại là vùng từ 681 cặp bazơ đến 1362 cặp bazơ của đầu tận cùng 5' của SEQ ID NO: 2). Gen Ubiquitin được khuếch đại để làm chất chuẩn nội bằng cách sử dụng các đoạn mồi UbiRTF và UbiRTR và trình tự của các đoạn mồi này được liệt kê trong Bảng 3. Kết quả này đã chỉ ra rằng mức độ biểu hiện của IPA1 tăng ở thực vật chuyển gen (Hình 4c).

Bảng 3. Trình tự của các đoạn mồi

Tên của các đoạn mồi	Trình tự của các đoạn mồi (theo chiều từ đầu tận cùng 5' đến đầu tận cùng 3')
IPA1RT1F	CGGTCGACTAGCTGCATCTGTTGG (SEQ ID NO: 17)
IPA1RT1R	CATCGTGTTGCTGGTTTGGTCGAAG (SEQ ID NO: 18)
UbiRTF	CCCTCCACCTCGTCCTCAG (SEQ ID NO: 19)
UbiRTR	AGATAACAACGGAAGCATAAAAGTC (SEQ ID NO: 20)

(2) Đặc điểm kiểu hình của thực vật chuyển gen

Thực hiện thống kê kiểu hình trên các thực vật chuyển gen T_1 với gIPA1 đã được biến nạp, cây lúa đối chứng Nipponbare và cây lúa đối chứng chứa vật truyền rỗng, và 10 cá thể cây được lấy ra từ mỗi vật liệu nghiên cứu. Kết quả đã được thể hiện trên Hình 4b và Hình 4d (kiểu hình của cây lúa đối chứng chứa vật truyền rỗng là giống với cây lúa đối chứng Nipponbare trên Hình 4b và 4d, không được thể hiện trên phần hình vẽ), trong đó số lượng chồi gốc

giảm (trung bình từ 11,9 xuống 6,7), số lượng nhánh cuống cấp một tăng lên đáng kể (trung bình từ 10,8 lên 15,5), số lượng nhánh cuống cấp hai cũng tăng lên đáng kể (trung bình từ 21,9 lên 25,6), thân trở nên mập hơn (đường kính của nhánh thứ ba tăng trung bình từ 0,35cm lên 0,47cm) và số lượng hạt trên một bông tăng đáng kể (trung bình từ 117,1 lên 135,7) khi so sánh thực vật chuyển gen T1 được biến nạp gen gIPA1 với cây đã được biến nạp bằng vật truyền rỗng và Nipponbare.

Ví dụ 3. Thu nhận và kiểm tra thực vật chuyển gen

1. Thu nhận thực vật chuyển gen

(1) Thu nhận mảnh can thiệp

Việc khuếch đại nhờ PCR được thực hiện bằng cách sử dụng cặp đoạn mồi ARNi1F/ARNi1R và cặp đoạn mồi ARNi2F/ARNi2R trong Bảng 4 và sử dụng dòng cADN với chiều dài đầy đủ (số AK107191) trên cơ sở dữ liệu KOME làm khuôn mẫu (có thể mua được từ Genome Resource Center, National Institute of Agrobiological Science, Japan), và tiến hành xác định trình tự đối với sản phẩm thu được. Trình tự nucleotit của mảnh gen thu được bằng cách khuếch đại nhờ sử dụng hai cặp đoạn mồi được lần lượt thể hiện trong SEQ ID NO: 4 và SEQ ID NO: 5 trong phần danh mục trình tự.

Bảng 4. Trình tự của các đoạn mồi

Tên của các đoạn mồi	Trình tự của các đoạn mồi (theo chiều từ đầu tận cùng 5' đến đầu tận cùng 3')
ARN can thiệp 1F	AGGATCCCCAGCCATGGGATACTACTACC (SEQ ID NO: 21)
ARN can thiệp 1R	AGGTACCCAGCATTAACACTGATACTTAAA (SEQ ID NO: 22)
ARN can thiệp 2F	AACTAGTCAGCATTAACACTGATACTTAAA (SEQ ID NO: 23)

ARN can thiệp 2R	AGAGCTCCCAGCCATGGGATACTACTACC (SEQ ID NO: 24)
------------------	--

Đầu tận cùng 5' của đoạn mỗi xuôi chiều và ngược chiều của ARN can thiệp 1 lần lượt được thêm vào các mối nối *Bam*HI và *Kpn*I và đầu tận cùng 5' của đoạn mỗi xuôi chiều và ngược chiều của ARN can thiệp 2 lần lượt được thêm vào các mối nối *Spe*I và *Sac*I.

SEQ ID NO: 4 là một mảnh có chiều dài từ 1014 cặp bazơ đến 1623 cặp bazơ của SEQ ID NO:2 và SEQ ID NO: 5 là trình tự bổ trợ ngược chiều của SEQ ID NO: 4. Đã khẳng định được rằng không có trình tự nào khác tương đồng với SEQ ID NO: 4 và 5 trong hệ gen cây lúa khi phân tích bằng cách so sánh trình tự toàn bộ hệ gen.

(2) Xây dựng vật truyền can thiệp

Sản phẩm đã được khuếch đại bằng cách sử dụng cặp đoạn mỗi ARNi1F/ARNi1R được phân cắt ở các vị trí *Bam*HI và *Kpn*I và *Bam*HI và *Kpn*I của vật truyền pTCK303 mang gen khởi đầu Ubiquitin (Công chúng có thể có được vật liệu này từ Viện Di truyền và Sinh học Phát triển, Viện Khoa học Trung Quốc (Institute of Genetics and Developmental Biology, Chinese Academy of Science) và vật liệu này được báo cáo trong tài liệu phi sáng chế Wand Z, Chen C, Xu Y, et al. Một vật truyền thiết thực để làm giảm một cách hữu hiệu mức độ biểu hiện gen ở cây lúa (*Oryza sativa* L.). *Plant Mol. Biol. Rep.* 2004, 22, 409-417) được cài xen để thu được vật truyền 1; sản phẩm thu được nhờ khuếch đại bằng cách sử dụng cặp đoạn mỗi ARNi2F/ARNi2R được phân cắt bởi *Spe*I và *Sac*I, và các mảnh thu được được cài xen vào vật truyền nhằm thu được vị trí *Spe*I và *Sac*I của vật truyền 1 và sau đó thu được vật truyền biểu hiện tái tổ hợp (nghĩa là vật truyền can thiệp IPA1-ARNi). Các mảnh được cài xen để tạo ra cấu trúc cặp tóc sau khi biểu hiện.

(3) Thu nhận thực vật chuyển gen

Vật truyền can thiệp IPA1-ARNi được biến nạp vào *Agrobacterium tumefaciens* dòng EHA05 bằng phương pháp bắn phá bằng điện để sàng lọc chủng *Agrobacterium tumefaciens* tái tổ hợp chứa vật truyền can thiệp IPA1-ARNi.

Các hạt trưởng thành của vật liệu Ri22 chứa gen đột biến điểm được bóc vỏ, khử trùng và gieo vào môi trường cảm ứng sẹo. Mô sẹo được trồng từ mảnh này sau ba tuần nuôi cấy. Mô sẹo trong thời kỳ đầu có sự phát triển mạnh mẽ, màu vàng nhạt và rời ra được chọn để sử dụng làm thụ thể của quá trình biến nạp.

Mô sẹo của cây lúa Ri22 được gây nhiễm bằng chủng *Agrobacterium tumefaciens* tái tổ hợp chứa vật truyền can thiệp IPA1-ARNi và tiếp theo được nuôi cấy ở 25°C trong bóng tối trong 3 ngày và mô sẹo và thực vật chuyển gen kháng lại chủng này được sàng lọc trên môi trường chọn lọc chứa 50mg/L hygromyxin. Các cây lúa kháng hygromyxin được làm thích nghi trong bóng râm và tiếp theo được cấy trên cánh đồng để thu được thực vật chuyển gen thế hệ T₀. Hạt chuyển gen của thế hệ T₀ được thu hoạch và tiếp theo được trồng để thu được thực vật chuyển gen thế hệ T₁ với gIPA1 đã được biến nạp.

Ri22 được biến nạp bằng vật truyền rỗng pTCK303 để thu được cây lúa đối chứng chứa vật truyền rỗng theo phương pháp thu nhận thực vật chuyển gen thế hệ T₁ được biến nạp bằng IPA1-ARN can thiệp.

2. Phát hiện thực vật chuyển gen

(1) Xác định mức độ biểu hiện của gen IPA1 bằng phương pháp RT-PCR

ARN tổng số của cây lúa được phân lập từ cây lúa chuyển gen và cây lúa đối chứng Ri22 bằng cách sử dụng TRIZOL (có thể mua được từ Invitrogen Company) và thực hiện phiên mã ngược bằng cách sử dụng kit phiên mã ngược

(có thể mua được từ Promega Company) để thu được cADN. Mức độ biểu hiện của gen IPA1 được xác định theo phương pháp PCR bằng cách sử dụng các đoạn mồi IPA1RT1F và IPA1RT1R. Gen Ubiquitin được khuếch đại để dùng làm chất chuẩn nội bằng cách sử dụng UbiRTF và UbiRTR và trình tự của các đoạn mồi này được liệt kê trong Bảng 3. Kết quả này chỉ ra rằng mức độ biểu hiện của IPA1 giảm trong thực vật chuyển gen (Hình 5b).

(2) Phát hiện kiểu hình của thực vật chuyển gen

Thực hiện việc thống kê kiểu hình trên thực vật chuyển gen thế hệ T_1 với cây IPA1-ARNi đã được biến nạp, cây lúa đối chứng Ri22 và cây lúa đối chứng chứa vật truyền rỗng và 10 cá thể cây được lấy ra từ mỗi vật liệu nghiên cứu này. Kết quả được thể hiện trên Hình 5a và Hình 5c (kiểu hình của cây đối chứng chứa vật truyền rỗng - không được thể hiện trên Hình vẽ - là giống với cây đối chứng Nipponbare trên Hình 5a và Hình 5c), trong đó chiều cao cây giảm (trung bình từ 115,7cm xuống 91,2cm), số lượng chồi gốc tăng đột ngột (trung bình từ 3,7 lên 23,3), số lượng nhánh cuống cấp một giảm đáng kể (trung bình từ 15,2 xuống 6,2), số lượng nhánh cuống cấp hai cũng giảm đáng kể (trung bình từ 57,5 xuống 9,7), số lượng hạt trên một bông giảm đáng kể (trung bình từ 259,6 xuống 54,6) và thân trở nên gầy hơn (đường kính của nhánh thứ cấp giảm trung bình từ 0,68cm xuống 0,29cm) khi so sánh thực vật chuyển gen T_1 chứa IPA1-ARNi đã được biến nạp bằng cây đối chứng chứa vật truyền rỗng và cây lúa đối chứng Ri22.

YÊU CẦU BẢO HỘ

1. Phương pháp lai tạo thực vật chuyển gen bao gồm bước đưa vật truyền can thiệp vào thực vật đích để thu được thực vật chuyển gen có các kiểu hình thay đổi so với thực vật đích,

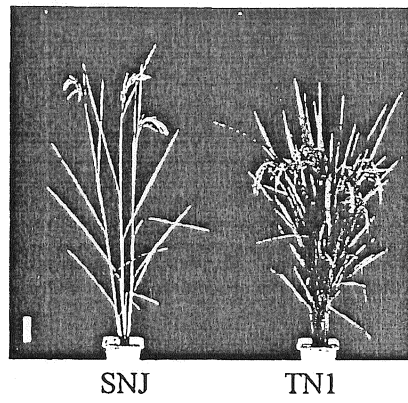
trong đó vật truyền can thiệp được xây dựng bằng cách lần lượt cài xen axit nucleic thứ nhất chứa trình tự nucleotit của SEQ ID NO: 4 và axit nucleic thứ hai chứa trình tự nucleotit của SEQ ID NO: 5 vào trong vật truyền nhờ đó cho phép hình thành cấu trúc cặp tóc khi biểu hiện; và

trong đó các kiểu hình thay đổi bao gồm chiều cao cây giảm, số lượng chồi gốc nhiều hơn, thân gầy hơn, số lượng nhánh cuống cấp một ít hơn và số lượng hạt trên mỗi bông ít hơn.

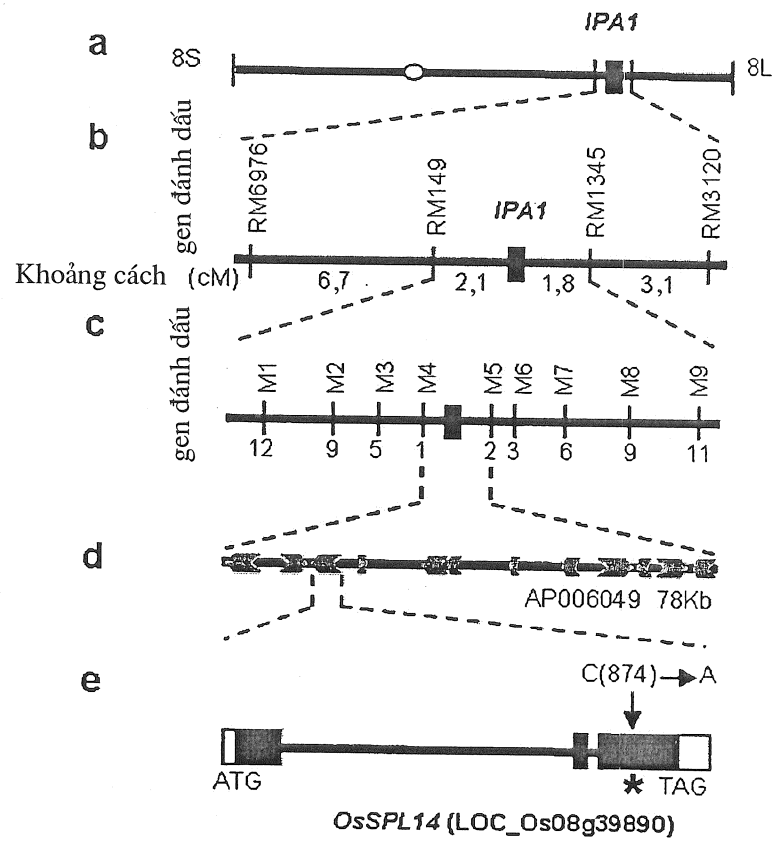
2. Phương pháp theo điểm 1, trong đó vật truyền là pTCK303, trong đó axit nucleic thứ nhất được cài xen giữa vị trí *Bam*HI và *Kpn*I, và trong đó axit nucleic thứ hai được cài xen giữa vị trí *Spe*I và *Sac*I.

3. Phương pháp theo điểm 1, trong đó thực vật là thực vật một lá mầm.

4. Phương pháp theo điểm 3, trong đó thực vật là cây lúa.

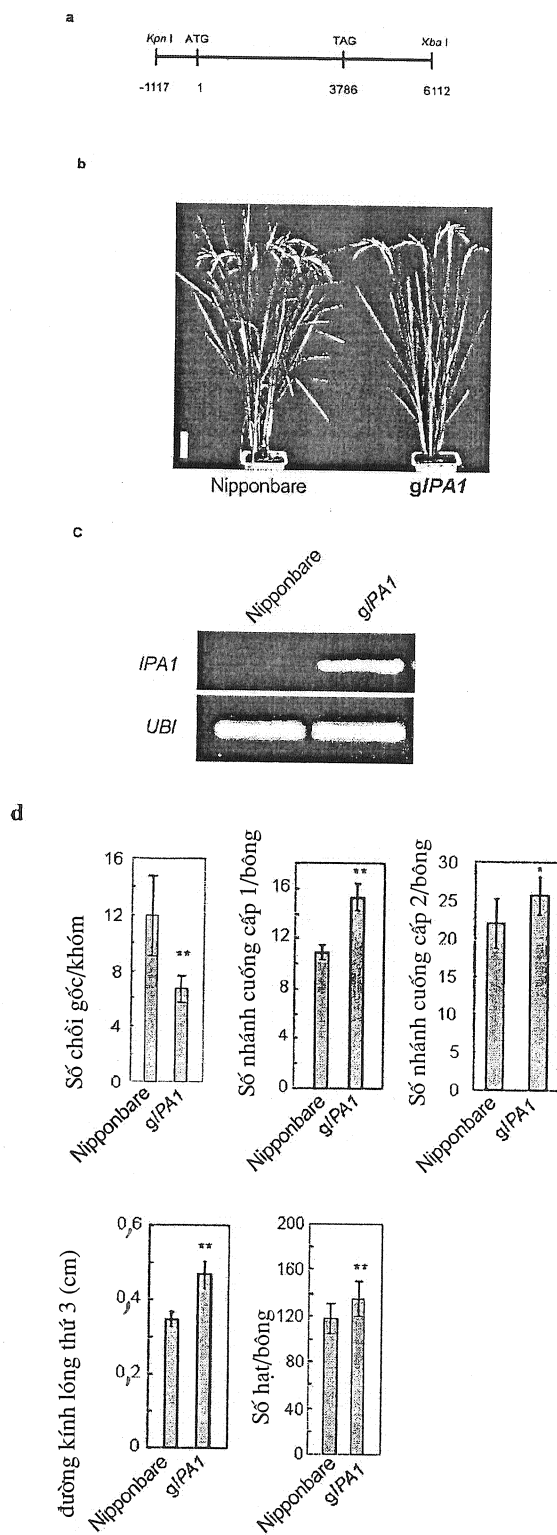


Hình 1

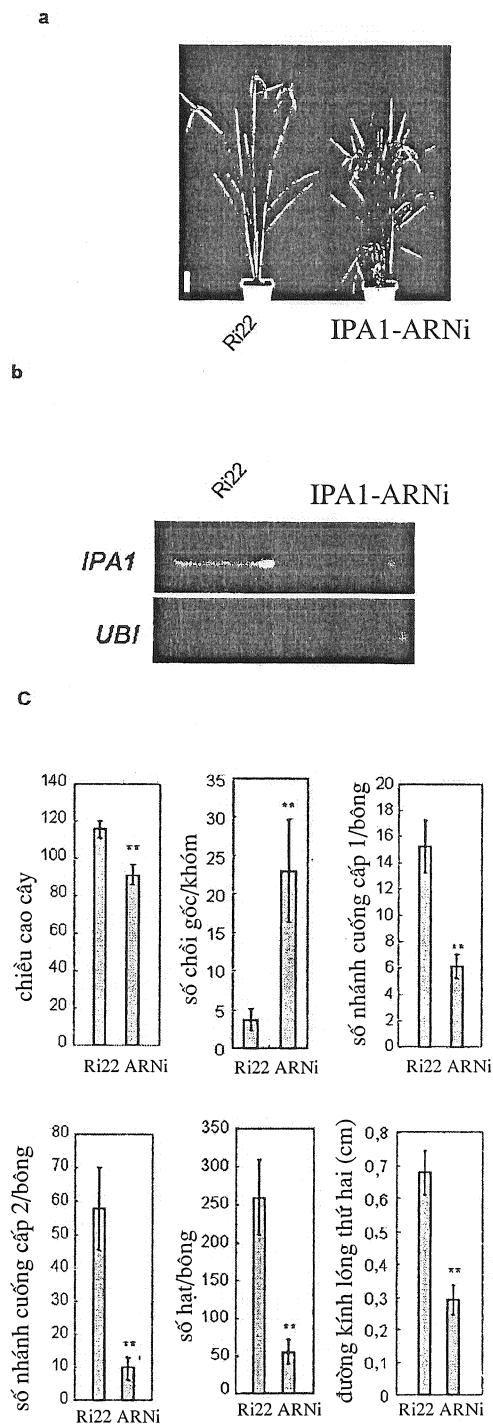


Hình 2

Hình 3



Hình 4



Hình 5

DANH MỤC TRÌNH TỰ

<110> INSTITUTE OF GENETICS AND DEVELOPMENTAL BIOLOGY, CHINESE ACADEMY OF SCIENCES

<120> Phương pháp lai tạo thực vật chuyển gen

<130> CGGNARL102214

<160> SỐ LƯỢNG 24

<210> SEQ ID NO 1

<211> CHIỀU DÀI: 417

<212> LOẠI: PRT

<213> SINH VẬT: oryza sativa

<400> TRÌNH TỰ: 1

Met Glu Met Ala Ser Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly Gly
1 5 10 15

Val Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Asp Glu His Arg Gln
20 25 30

Leu His Gly Leu Lys Phe Gly Lys Lys Ile Tyr Phe Glu Asp Ala Ala
35 40 45

Ala Ala Ala Gly Gly Gly Gly Thr Gly Ser Gly Ser Gly Ser Ala Ser
50 55 60

Ala Ala Pro Pro Ser Ser Ser Ser Lys Ala Ala Gly Gly Gly Arg Gly
65 70 75 80

Gly Gly Gly Lys Asn Lys Gly Lys Gly Val Ala Ala Ala Ala Pro Pro
85 90 95

Pro Pro Pro Pro Pro Pro Arg Cys Gln Val Glu Gly Cys Gly Ala Asp
100 105 110

Leu Ser Gly Ile Lys Asn Tyr Tyr Cys Arg His Lys Val Cys Phe Met
115 120 125

His Ser Lys Ala Pro Arg Val Val Val Ala Gly Leu Glu Gln Arg Phe
130 135 140

Cys Gln Gln Cys Ser Arg Phe His Leu Leu Pro Glu Phe Asp Gln Gly
145 150 155 160

Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg Arg Arg
165 170 175

Arg Pro Gln Thr Pro Leu Ala Ser Arg Tyr Gly Arg Leu Ala Ala Ser
180 185 190

Val Gly Glu His Arg Arg Phe Arg Ser Phe Thr Leu Asp Phe Ser Tyr
195 200 205

Pro Arg Val Pro Ser Ser Val Arg Asn Ala Trp Pro Ala Ile Gln Pro
210 215 220

Gly Asp Arg Ile Ser Gly Gly Ile Gln Trp His Arg Asn Val Ala Pro

225		230		235		240
His Gly His Ser	Ser Ala Val Ala Gly Tyr	Gly Ala Asn Thr Tyr Ser				
	245	250	255			
Gly Gln Gly Ser	Ser Ser Ser Gly Pro Pro Val Phe Ala Gly Pro Asn					
	260	265	270			
Leu Pro Pro Gly Gly Cys Leu Ala Gly Val Gly Ala Ala Thr Asp Ser						
	275	280	285			
Ser Cys Ala Leu Ser Leu Leu Ser Thr Gln Pro Trp Asp Thr Thr Thr						
	290	295	300			
His Ser Ala Ala Ala Ser His Asn Gln Ala Ala Ala Met Ser Thr Thr						
305	310	315	320			
Thr Ser Phe Asp Gly Asn Pro Val Ala Pro Ser Ala Met Ala Gly Ser						
	325	330	335			
Tyr Met Ala Pro Ser Pro Trp Thr Gly Ser Arg Gly His Glu Gly Gly						
	340	345	350			
Gly Arg Ser Val Ala His Gln Leu Pro His Glu Val Ser Leu Asp Glu						
	355	360	365			
Val His Pro Gly Pro Ser His His Ala His Phe Ser Gly Glu Leu Glu						
	370	375	380			
Leu Ala Leu Gln Gly Asn Gly Pro Ala Pro Ala Pro Arg Ile Asp Pro						
385	390	395	400			
Gly Ser Gly Ser Thr Phe Asp Gln Thr Ser Asn Thr Met Asp Trp Ser						
	405	410	415			
Leu						

<210> SEQ ID NO 2
 <211> CHIỀU DÀI: 1624
 <212> LOẠI: ADN
 <213> SINH VẬT: oryza sativa
 <220> DẤU HIỆU:
 <221> TÊN/CHÌA KHÓA: CDS
 <222> VỊ TRÍ: (124)..(1374)
 <400> TRÌNH TỰ: 2

ttccgtctctt	ttctctctctc	ttctctctctcc	ccctctctctg	gaggagagag	aggagaagag	60
gaggggggggc cgcgccaaga gccacgcgcg ctacagtctc cttcccaccc gcgaccgcga						120
gca atg gag atg gcc agt gga gga ggc gcc gcc gcc gcc gcc gcc ggc ggc						168
Met Glu Met Ala Ser Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly						
1	5	10	15			
gga gta ggc ggc agc ggc ggc ggt ggt ggt gga ggg gac gag cac cgc						216
Gly Val Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Gly Asp Glu His Arg						

20					25					30						
cag	ctg	cac	ggt	ctc	aag	ttc	ggc	aag	aag	atc	tac	ttc	gag	gac	gcc	264
Gln	Leu	His	Gly	Leu	Lys	Phe	Gly	Lys	Lys	Ile	Tyr	Phe	Glu	Asp	Ala	
			35					40					45			
gcc	gcg	gca	gca	ggc	ggc	ggc	ggc	act	ggc	agt	ggc	agt	ggc	agc	gcg	312
Ala	Ala	Ala	Ala	Gly	Gly	Gly	Gly	Thr	Gly	Ser	Gly	Ser	Gly	Ser	Ala	
			50				55					60				
agc	gcc	gcg	ccg	ccg	tcc	tcg	tct	tcc	aag	gcg	gcg	ggt	ggt	gga	cgc	360
Ser	Ala	Ala	Pro	Pro	Ser	Ser	Ser	Ser	Lys	Ala	Ala	Gly	Gly	Gly	Arg	
			65			70					75					
ggc	gga	ggg	ggc	aag	aac	aag	ggg	aag	ggc	gtg	gcc	gcg	gcg	gcg	cca	408
Gly	Gly	Gly	Gly	Lys	Asn	Lys	Gly	Lys	Gly	Val	Ala	Ala	Ala	Ala	Pro	
					85					90					95	
ccg	ccg	ccg	ccg	ccg	ccg	ccg	ccg	tgc	cag	gtg	gag	ggg	tgc	ggc	gcg	456
Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Arg	Cys	Gln	Val	Glu	Gly	Cys	Gly	Ala
					100					105					110	
gat	ctg	agc	ggg	atc	aag	aac	tac	tac	tgc	cgc	cac	aag	gtg	tgc	ttc	504
Asp	Leu	Ser	Gly	Ile	Lys	Asn	Tyr	Tyr	Cys	Arg	His	Lys	Val	Cys	Phe	
			115					120					125			
atg	cat	tcc	aag	gct	ccc	cgc	gtc	gtc	gtc	gcc	ggc	ctc	gag	cag	cgc	552
Met	His	Ser	Lys	Ala	Pro	Arg	Val	Val	Val	Ala	Gly	Leu	Glu	Gln	Arg	
			130				135					140				
ttc	tgc	cag	cag	tgc	agc	agg	ttc	cac	ctg	ctg	cct	gaa	ttt	gac	caa	600
Phe	Cys	Gln	Gln	Cys	Ser	Arg	Phe	His	Leu	Leu	Pro	Glu	Phe	Asp	Gln	
			145			150					155					
gga	aaa	cgc	agc	tgc	cgc	aga	cgc	ctt	gca	ggt	cat	aat	gag	cgc	cgg	648
Gly	Lys	Arg	Ser	Cys	Arg	Arg	Arg	Leu	Ala	Gly	His	Asn	Glu	Arg	Arg	
					165					170					175	
agg	agg	ccg	caa	acc	cct	ttg	gca	tca	cgc	tac	ggt	cga	cta	gct	gca	696
Arg	Arg	Pro	Gln	Thr	Pro	Leu	Ala	Ser	Arg	Tyr	Gly	Arg	Leu	Ala	Ala	
				180					185					190		
tct	gtt	ggt	gag	cat	cgc	agg	ttc	aga	agc	ttt	acg	ttg	gat	ttc	tcc	744
Ser	Val	Gly	Glu	His	Arg	Arg	Phe	Arg	Ser	Phe	Thr	Leu	Asp	Phe	Ser	
			195					200					205			
tac	cca	agg	gtt	cca	agc	agc	gta	agg	aat	gca	tgg	cca	gca	att	caa	792
Tyr	Pro	Arg	Val	Pro	Ser	Ser	Val	Arg	Asn	Ala	Trp	Pro	Ala	Ile	Gln	
			210				215					220				
cca	ggc	gat	cgg	atc	tcc	ggt	ggt	atc	cag	tgg	cac	agg	aac	gta	gct	840
Pro	Gly	Asp	Arg	Ile	Ser	Gly	Gly	Ile	Gln	Trp	His	Arg	Asn	Val	Ala	
			225			230					235					
cct	cat	ggt	cac	tct	agt	gca	gtg	gcg	gga	tat	ggt	gcc	aac	aca	tac	888
Pro	His	Gly	His	Ser	Ser	Ala	Val	Ala	Gly	Tyr	Gly	Ala	Asn	Thr	Tyr	
					245					250					255	

agc ggc caa ggt agc tct tct tca ggg cca ccg gtg ttc gct ggc cca	936
Ser Gly Gln Gly Ser Ser Ser Ser Gly Pro Pro Val Phe Ala Gly Pro	
260 265 270	
aat ctc cct cca ggt gga tgt ctc gca ggg gtc ggt gcc gcc acc gac	984
Asn Leu Pro Pro Gly Gly Cys Leu Ala Gly Val Gly Ala Ala Thr Asp	
275 280 285	
tcg agc tgt gct ctc tct ctt ctg tca acc cag cca tgg gat act act	1032
Ser Ser Cys Ala Leu Ser Leu Leu Ser Thr Gln Pro Trp Asp Thr Thr	
290 295 300	
acc cac agt gcc gct gcc agc cac aac cag gct gca gcc atg tcc act	1080
Thr His Ser Ala Ala Ala Ser His Asn Gln Ala Ala Ala Met Ser Thr	
305 310 315	
acc acc agc ttt gat ggc aat cct gtg gca ccc tcc gcc atg gcg ggt	1128
Thr Thr Ser Phe Asp Gly Asn Pro Val Ala Pro Ser Ala Met Ala Gly	
320 325 330 335	
agc tac atg gca cca agc ccc tgg aca ggt tct cgg ggc cat gag ggt	1176
Ser Tyr Met Ala Pro Ser Pro Trp Thr Gly Ser Arg Gly His Glu Gly	
340 345 350	
ggt ggt cgg agc gtg gcg cac cag cta cca cat gaa gtc tca ctt gat	1224
Gly Gly Arg Ser Val Ala His Gln Leu Pro His Glu Val Ser Leu Asp	
355 360 365	
gag gtg cac cct ggt cct agc cat cat gcc cac ttc tcc ggt gag ctt	1272
Glu Val His Pro Gly Pro Ser His His Ala His Phe Ser Gly Glu Leu	
370 375 380	
gag ctt gct ctg cag ggg aac agt cca gcc cca gca cca cgc atc gat	1320
Glu Leu Ala Leu Gln Gly Asn Gly Pro Ala Pro Ala Pro Arg Ile Asp	
385 390 395	
cct ggg tcc ggc agc acc ttc gac caa acc agc aac acg atg gat tgg	1368
Pro Gly Ser Gly Ser Thr Phe Asp Gln Thr Ser Asn Thr Met Asp Trp	
400 405 410 415	
tct ctg tagaggctgt tccagctgcc atcgatctgt cgtcccgcga ggcgagtcac	1424
ser Leu	
ggaactgaag aacctcatgc tgctgccct tattttgtgt tcaaattttc ctttccagta	1484
tggaaggaa attctaagg gactggcgat taatctccct gtgatgaata ataatgcgcg	1544
cccttgaaact caattaattg ctgtgccgca tccatctatg taactctcca tgaattttta	1604
agtatcagtg ttaatgctgt	1624

<210> SEQ ID NO 3

<211> CHIỀU DÀI: 7335

<212> LOẠI: ADN

<213> SINH VẬT: oryza sativa

<400> TRÌNH TỰ: 3

gcaatgtaga gccacgtagg caagtcgctt gcgtggagga gagaggggag tggggaccgt 60
 tcccaaccca gcttcgtgtg accaagtttg gccacacggg ccaaacgaac ctgagcaact 120
 tttgtcagaa agaaaagaac ctccgcgaga aacaagaaag cgagagaggg agagaaagga 180
 ggctcgtcgg agtaggggag ctccgggtat ggggctcggc ggaggctcgc tggagtaggg 240
 gccgccactg cgtggggctc gccggagtag gggcgtcggc ggagcctccc gagatccgcc 300
 gctcagcggc gccgccgtct tccgggcaga gctctcgaag ctcgccctcc tcgcgcgccg 360
 gtggcggttg cggggcccg cgtgtggctac gcagctccgg tgctgcgcct ccaccgtcga 420
 cgacagcgcc gcttggcgct gccgccgtct tccgccgcgc cgctggacgc cgccagatct 480
 gctgctcgtc gcccggtggg ccgctccacc cggttggagg aggagaggcg gcgccgcgct 540
 tgggctgccc caccgccgag ctctgccgcg ccgttcgccg gtgctgccga gctccgccgc 600
 gcctgccgga gcacgctgcc atggccgccc tggagaagac acgagagaat taggtggagg 660
 gtgggggaag ggtgagattt tttatattat ctatgggtcc cattataaat tttctaaacc 720
 acacttatac tgtgggtgca gtgtcattta gagttcccaa accacctatg ttgcagctgt 780
 ggtataacaa tttgctagga cgcattgcta ctgcccttgt accctgctat aagaagataa 840
 ccaatgacat ctccactcga ttttctcggc gcgcgtgtga ggggtgtgagg ataattttta 900
 ttttaagtgg tttttaaggg cggagagaga gagagagaga gggcaccgca ctacttctac 960
 ttgtgtgtgt gtcgctcgtt gggcttcgcc acctttccgt ctctttctct tctcttctct 1020
 ctccccctct cctggaggag agagaggaga agaggagggg gggccgcgcc aagagccacg 1080
 cgcgctacag tctccttccc acccgcgacc gcgagcaatg gagatggcca gtggaggagg 1140
 cgccgccgcc gccgccggcg gcggagtagg cggcagcggc ggcgggtggtg gtggagggga 1200
 cgagcaccgc cagctgcacg gtctcaagtt cggcaagaag atctacttcg aggacgccgc 1260
 cgcggcagca ggcggcgcg gcactggcag tggcagtggc agcgcgagcg ccgcgccgcc 1320
 gtccctcgtct tccaaggcgg cgggtggttg acgcggcgga gggggcaaga acaaggggaa 1380
 gggcggtggc gcggcgcgcc caccgccgcc gccgccgccg ccgcggtgcc aggtggaggg 1440
 gtgcggcgcg gatctgagcg ggatcaagaa ctactactgc cgccacaagg tgtgcttcat 1500
 gcattccaag gctccccgcg tcgtcgtcgc cggcctcgag cagcgcttct gccagcagt 1560
 cagcaggtca ctctctcact cacctcgcca ttgctgatgt caccactgct tttgctttgc 1620

tttgcttgct ctccctcctc ttccacctat ctctcttggt tatttgcttc ttgttcttgt 1680
 ttagtgctag tacatgtggt gttattgttg tgccgttttg tcttttggtt tattgtgttg 1740
 ttgttactac tcgttttact atagggtttt aagggttatg agcacggcca ccacattaga 1800
 tgcactgtca agtgggtgtg gtgggacctt tcctgctaaa acaagctgat ttcaactctc 1860
 tgaaacttcc tgcatttcat ctatttttat ctttgattgt gttgggagta ctacactagt 1920
 agtgtaata ttttgactgg tgcttatgag atttttaagt tggtaggttg atgaggaaaa 1980
 tactccttta tatggttgag tgatgtgact tgccgtgtctg cctgcctgcc tgccgctttg 2040
 cataagattc ctctgtgtta gtaagagcca ctgtttattt gtactggtgc ttactctact 2100
 tagttaatta gccattagct ataaaattcc gttgatgttg caagcttagc aatggccacg 2160
 gtaagaatgg gagagagaag ttggctaaag ctgttgcttt gtagtttgta ctatatatgt 2220
 gtctttgtgt tgcaagatat gcaactccta ctatgctgtg acttgagctc aaggttttca 2280
 gttatctata gatccttact actactgagc atactaccac ttctgtatgg tagcatatgg 2340
 tagcatagtc caagttccaa cgccctcgcca gttgttcata atctatacta ccacttctgt 2400
 gcatttgta cttttattta atagtttgtc tcattagctg acaagcatat gcctgttttg 2460
 atatctgccc ctcttgtaat agtctatgga tagcttggac tgtttgatgc ttttaattttt 2520
 tactagcaac acttagggcc cctttgaaat ggaggattag caaaggaatt ttggaggatt 2580
 cattttccta aggatttttt cctatagagc cctttgattc atagaaagag gataggaaaa 2640
 cttccgtagg attgcattcc tatgatcaat tccataggaa aataagcaag aggttagacc 2700
 tcttgtgaaa ctttcctttg ttgagtgtat cttgtggtat aatcaaaggg ctcttctctc 2760
 catttcatgt gttttcaatt cctgtaggat tggaaaaaca tacaacttca attcctacgt 2820
 ttttcctatt cctatgtttt tcctatcctg cgtttcaaag gggcccttaa ggatgaaggg 2880
 aagtaagaga aacatactag agaatatgta gtagtatttc tacattccat atttgtagca 2940
 ctagcccaca aatatctttg ccttgtaactt acttcatacc agttcccccc ttttcagagc 3000
 aaaccaacaa tttctgttgc cttatatatc tagtgtcttc gtactaatat atctgttcca 3060
 aatgtacct gtccaaattc atagctagaa atagctttat ttaggacgga agtaataact 3120
 gttgttagag acttggttca gacttttggt tatgttgagg ctactatcat ttcctttacg 3180
 ggccaaatta ctacaaatga gaattcataa aaatgtcaag attttatgat tgttgtagct 3240
 ttatttagga cggaggtagt aattgttggt agagacttgg ttcagacttt tggttacgtt 3300
 gaagctacta tcatttcctt tatggtcaaa ttactaacia tgagtattca taaaaatgtc 3360

aagattttat aattgagctg tgccagtgc aagtgtgtca ctatctgatg ccataatgca 3420
tcattataaa agccagatgg accattagct tttatgtgta ggacacctgc cgtccaatta 3480
gatggataac catctagtgt ttgtgtactg ttattttaag cccgacatct cacaactcca 3540
tgaatgatta cagtcttctt ttcacatggt gtccttttgt tgtgttagga atagcatttt 3600
ttatttatgg gtgtaattat gaaaggcact aggagagttg ctgctttatc ttgatgggat 3660
ttgtagtaat accatcttta ggatgacaag aaatcttggt ctgagttagc atgggctgcc 3720
ttttgacctg agctacgggt tgctatgttt ggcttgcac atgcagatct attaggataa 3780
taagcatata aaagttgctt gcattgtgca ttgcttggtt taccttgatt catgtaggag 3840
taatttgctc gccatgcctc gttttgcttt ctgagtcaac agccaaattt agatgatgta 3900
ccttctgttg cttcaaaaac tcagtcactg cacagcagca gtggatagga ttcagaatca 3960
atctatccat gattctctgt tcacataata tgacagggtc cacctgctgc ctgaatttga 4020
ccaaggaaaa cgcagctgcc gcagacgcct tgcagggtcat aatgagcgcc ggaggaggcc 4080
gcaaaccctt ttggcatcac gctacgggtc actagctgca tctgttggtg gtatcatcag 4140
aggctcttgt tttctttgca tcttgtgtgt ttgttggtta ctactggttg cattcgctga 4200
tgtgttggtt gttgcgattc ttgatccaga agagcatcgc aggttcagaa gctttacgtt 4260
ggatttctcc tacccaaggg ttccaagcag cgtaaggaat gcatggccag caattcaacc 4320
aggcgatcgg atctccggtg gtatccagtg gcacaggaac gtagctcctc atggtcactc 4380
tagtgcaagt gcgggatatg gtgccaacac atacagcggc caaggtagct cttcttcagg 4440
gccaccggtg ttcgctggcc caaatctccc tccagggtga tgtctcgag gggtcggtgc 4500
cgccaccgac tcgagctgtg ctctctctct tctgtcaacc cagccatggg atactactac 4560
ccacagtgcc gctgccagcc acaaccaggc tgcagccatg tccactacca ccagctttga 4620
tggcaatcct gtggcaccct ccgcatggc gggtagctac atggcaccaa gccctggac 4680
aggttctcgg ggccatgagg gtggtggtcg gagcgtggcg caccagctac cacatgaagt 4740
ctcacttgat gaggtgcacc ctggtcctag ccatcatgcc cacttctccg gtgagcttga 4800
gcttgctctg caggggaacg gtccagcccc agcaccacgc atcgatcctg ggtccggcag 4860
caccttcgac caaaccagca acacgatgga ttggtctctg tagaggctgt tccagctgcc 4920
atcgatctgt cgtcccgcaa ggcgagtcac ggaactgaag aacctcatgc tgctgcct 4980
tattttgtgt tcaaattttc ctttcagta tggaaaggaa attctaaggt gactggcgat 5040

taatctccct gtgatgaata ataatgcgcg cccttgaact caattaattg ctgtgccgca 5100
 tccatctatg taactctcca tgaattttta agtatcagtg ttaatgctgt attgtcgagg 5160
 acttctgctc gatatgttat ttctcttatg ttgttcatca tgaatctttt tctgcttatt 5220
 attctggtgc cgggttgtcc ttaccacaga agattcagtt tcggttggcg agagtaaaca 5280
 ccttccctgg ttgtgacaaa agctccaacc ttttcacttc tcggcctgta tttgatcttc 5340
 cccttctgac gctgttatac tactttttaag cctgtatggt tccagccttc cagggtgaagg 5400
 gccatactga agagaaaaca tgctttcagg gtttgatgca ttgtgtactt tacaagtgtg 5460
 ctttaagattt tgtacaattt atatatgtac ctgctctgct gctgagtatt gtaggaaaga 5520
 atcagttcga agggcgtgtg ttcatgtaaa gtgagaccac atgcacagcg tggatttgca 5580
 gcatgctctc tgcaccagtg gtgttctggt gatgcctttg atgggctggc tgaggtgaga 5640
 ggaggatgat ccatgttggc agcttcttca ctctgaaaaa taaaagagaa gaaatgttca 5700
 gatttgcaga caagtggaga gcagtgatat attctacaat aaaacattac caccttgctt 5760
 ttctgtgatg atagatactc catggaattt tgcacaaagc atctcttggt ttccagccac 5820
 tgtttgctgg gttgttgctt caatttcgct ccaattgatt ggtcaccttt gggtgtgact 5880
 tgagagcact gagcactgaa acttttgctg tcagcaggca atgcacctca tccatgtcac 5940
 gacagagga gagagccac ataaatggcc aaagaggaca catacagtgg cactgatgca 6000
 gtcattgcaa cataattgac atcatgctaa acagtgggtg aaccatatgt tagctagcct 6060
 gtgatcagca aacagtgatt atggatctta atgtcacatg caagatttga cacagttgta 6120
 aaaccatcat tgcattgaag atagatccca gcaacaggtg tatgatgtat tgctagaatg 6180
 aatcaaaaat atcagtgcc aactaaacac agtactacca acttgaacag ttatcaccgt 6240
 gattggaaaa cagaaatgta taattgcttt ggcgccatct gcttatcatt atcatatgtc 6300
 gcagatcact tgttccattg acacgactct ttttcaactgt gaggagaggc accttgattt 6360
 ggacttttca agagctgtag caagggtcc tttgaagcct tctacatgga ggagcagagc 6420
 ataccatagc cagaactgta aactcttctc gaagctttcc agttccaccc ttgtagcttt 6480
 aagctgccgc aagagaatta tcatttctaa cattgagatg tgatactgaa atgtgaaagg 6540
 tgattcgcag tataggtccc aaaatatcgt ttacagcaac ttgcaaacc tgcagtatac 6600
 agttaattca tcaaaatatt agaccattag tactacagtc taaaaatacc ccttaactga 6660
 acatgtatga taaggacaag attctgaagc tccagtgcac caggaatcca acgcagtatg 6720
 caaatcatta ctgaacaaga ttctgcact tacagaatca tcacctgttg taacaaggac 6780

cattctttgt tgccccagac acagcgaatt aatgggtcatc ttcatttggc ccaggacatt 6840
 catttgccaa tgcttctgct gattcatact gaaaagggga caatgcgtcc aatttttaaaa 6900
 gcatggaaga tgctataaaa gatcacccta tttaaaatgc agagaaaacc aaagatccaa 6960
 catgatatgg taatcacaga ttcccaacag taatgccgtc cagtaggcag taggggcatg 7020
 cacataaaca ctagtactat gtagagctga agcttatttc cagaatgaag ctgaccttgc 7080
 aaccgcaata aaggcaatag tagtggtcat cgcgagctt agccaatata ttttgcaa 7140
 cctgcaagaa taaacacagg tcaatctcgt cctttcagca aaatttgcag tcttgcataa 7200
 gatttctcag ataaaaaagg aagtctagac aagaatatga aggaaagaat gtgcaa 7260
 aagtatatcc ataattcaaa gttcgagcta tttccattgt cacaacataa tatgtgtaac 7320
 caatagctgg taagc 7335

<210> SEQ ID NO 4
 <211> CHIỀU DÀI: 610
 <212> LOẠI: ADN
 <213> SINH VẬT: *oryza sativa*

<400> TRÌNH TỰ: 4

ccagccatgg gatactacta cccacagtgc cgctgccagc cacaaccagg ctgcagccat 60
 gtccactacc accagctttg atggcaatcc tgtggcacc tccgccatgg cgggtagcta 120
 catggcacca agccccctgga caggttctcg gggccatgag ggtgggtggc ggagcgtggc 180
 gcaccagcta ccacatgaag tctcacttga tgaggtgcac cctggtccta gccatcatgc 240
 ccacttctcc ggtgagcttg agcttgctct gcaggggaac ggtccagccc cagcaccacg 300
 catcgatcct gggtcgggca gcacctctga ccaaaccagc aacacgatgg attggctctc 360
 gtagaggctg ttccagctgc catcgatctg tcgtcccgca aggcgagtca tggaaactgaa 420
 gaacctcatg ctgcctgccc ttattttgtg ttcaaatttt cctttccagt atggaaagga 480
 aattctaagg tgactggcga ttaatctccc tgtgatgaat aataatgcgc gcccttgaac 540
 tcaattaatt gctgtgccgc atccatctat gtaactctcc atgaattttt aagtatcagt 600
 gttaatgctg 610

<210> SEQ ID NO 5
 <211> CHIỀU DÀI: 610
 <212> LOẠI: ADN
 <213> SINH VẬT: *oryza sativa*

<400> TRÌNH TỰ: 5

cagcattaac actgatactt aaaaattcat ggagagttac atagatggat gcggcacagc 60
 aattaattga gttcaagggc gcgcattatt attcatcaca gggagattaa tcgccagtca 120
 ccttagaatt tcctttccat actggaaagg aaaatttgaa cacaaaataa gggcaggcag 180
 catgaggttc ttcagttcca tgactcgcct tgccgggacga cagatcgatg gcagctggaa 240
 cagcctctac agagaccaat ccatcgtgtt gctggtttgg tcgaagggtg tcgccggaccc 300
 aggatcgatg cgtgggtgctg gggctggacc gttcccctgc agagcaagct caagctcacc 360
 ggagaagtgg gcatgatggc taggaccagg gtgcacctca tcaagtgaga cttcatgtgg 420
 tagctggtgc gccacgctcc gaccaccacc ctcatggccc cgagaacctg tccaggggct 480

tggtgccatg tagctacccg ccatggcgga ggggtgccaca ggattgccat caaagctggt 540
 ggtagtggac atggctgcag cctggttgtg gctggcagcg gcactgtggg tagtagtatc 600
 ccatggctgg 610

<210> SEQ ID NO 6
 <211> CHIỀU DÀI: 1624
 <212> LOẠI: ADN
 <213> SINH VẬT: oryza sativa
 <220> DẤU HIỆU:
 <221> TÊN/CHÌA KHÓA: CDS
 <222> VỊ TRÍ: (124)..(1374)

<400> TRÌNH TỰ: 6

ttccgtctct ttcctctctc ttctctctcc ccctctcctg gaggagagag aggagaagag 60
 gagggggggc cgcgccaaga gccacgcgcg ctacagtctc cttcccaccc gcgaccgcga 120
 gcaatggaga tggccagtgg aggaggcgcc gccgcccgcg ccggcgggcg agtaggcggc 180
 agcggcgggc gtggtggtgg aggggacgag caccgccagc tgcacggtct caagttcggc 240
 aagaagatct acttcgagga cgcgcgcgcg gcagcaggcg gcggcgggcac tggcagtggc 300
 agtggcagcg cgagcgccgc gccgccgtcc tcgtcttcca aggcggcggg tggtaggacg 360
 ggcggagggg gcaagaacaa ggggaaggcg gtggccgcgg cggcgccacc gccgcgcgcg 420
 ccgcccgcgc ggtgccagggt ggaggggtgc ggcgcggtatc tgagcgggat caagaactac 480
 tactgccgcc acaagggtgtg cttcatgcat tccaaggctc cccgcgtcgt cgtcgccggc 540
 ctcgagcagc gcttctgccca gcagtgcagc aggttccacc tgctgcctga atttgaccaa 600
 ggaaaacgca gctgccgcag acgccttgca ggtcataatg agcgccggag gaggccgcaa 660
 accccttttg catcacgcta cggtcgacta gctgcatctg ttggtgagca tcgcagggtc 720
 agaagcttta cgttggtatt ctcctaccca agggttccaa gcagcgtaag gaatgcatgg 780
 ccagcaattc aaccaggcga tcggatctcc ggtggtatcc agtggcacag gaacgtagct 840
 cctcatggtc actctagtgc agtggcggga tatggtgcca acacatacag cggccaagggt 900
 agctcttctt cagggccacc ggtgttcgct ggcccaaate tccctccagg tggatgtctc 960
 gcaggggtcg gtgcccgcac cgactcgagc tgtgctatct ctcttctgtc aaccagcca 1020
 tgggatacta ctaccacag tgccgctgcc agccacaacc aggtgcagc catgtccact 1080
 accaccagct ttgatggcaa tcctgtggca ccctccgcca tggcggttag ctacatggca 1140
 ccaagccctt ggacagggtc tcggggccat gagggtggtg gtcggagcgt ggcgcaccag 1200
 ctaccacatg aagtctcact tgatgaggtg caccctggtc ctagccatca tgcccacttc 1260
 tccggtgagc ttgagcttgc tctgcagggg aacggtccag cccagcacc acgcatcgat 1320
 cctgggtccg gcagcacctt cgaccaaacc agcaacacga tggattggtc tctgtagagg 1380
 ctgttccagc tgccatcgat ctgtcgtccc gcaaggcgag tcatggaact gaagaacctc 1440
 atgctgctg cccttatttt gtgttcaaat tttcctttcc agtatggaaa ggaaattcta 1500
 aggtgactgg cgattaatct ccctgtgatg aataataatg cgcgcccttg aactcaatta 1560
 attgctgtgc cgcattccatc tatgtaactc tccatgaatt tttaagtatc agtggttaatg 1620
 ctgt 1624

<210> SEQ ID NO 7
 <211> CHIỀU DÀI: 417
 <212> LOẠI: PRT
 <213> SINH VẬT: oryza sativa

<400> TRÌNH TỰ: 7

Met Glu Met Ala Ser Gly Gly Gly Ala Ala Ala Ala Gly Gly Gly
 1 5 10 15

Val Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Asp Glu His Arg Gln
 20 25 30

Leu His Gly Leu Lys Phe Gly Lys Lys Ile Tyr Phe Glu Asp Ala Ala
 35 40 45
 Ala Ala Ala Gly Gly Gly Gly Thr Gly Ser Gly Ser Gly Ser Ala Ser
 50 55 60
 Ala Ala Pro Pro Ser Ser Ser Ser Lys Ala Ala Gly Gly Gly Arg Gly
 65 70 75 80
 Gly Gly Gly Lys Asn Lys Gly Lys Gly Val Ala Ala Ala Ala Pro Pro
 85 90 95
 Pro Pro Pro Pro Pro Pro Arg Cys Gln Val Glu Gly Cys Gly Ala Asp
 100 105 110
 Leu Ser Gly Ile Lys Asn Tyr Tyr Cys Arg His Lys Val Cys Phe Met
 115 120 125
 His Ser Lys Ala Pro Arg Val Val Val Ala Gly Leu Glu Gln Arg Phe
 130 135 140
 Cys Gln Gln Cys Ser Arg Phe His Leu Leu Pro Glu Phe Asp Gln Gly
 145 150 155 160
 Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg Arg Arg
 165 170 175
 Arg Pro Gln Thr Pro Leu Ala Ser Arg Tyr Gly Arg Leu Ala Ala Ser
 180 185 190
 Val Gly Glu His Arg Arg Phe Arg Ser Phe Thr Leu Asp Phe Ser Tyr
 195 200 205
 Pro Arg Val Pro Ser Ser Val Arg Asn Ala Trp Pro Ala Ile Gln Pro
 210 215 220
 Gly Asp Arg Ile Ser Gly Gly Ile Gln Trp His Arg Asn Val Ala Pro
 225 230 235 240
 His Gly His Ser Ser Ala Val Ala Gly Tyr Gly Ala Asn Thr Tyr Ser
 245 250 255
 Gly Gln Gly Ser Ser Ser Ser Gly Pro Pro Val Phe Ala Gly Pro Asn
 260 265 270
 Leu Pro Pro Gly Gly Cys Leu Ala Gly Val Gly Ala Ala Thr Asp Ser
 275 280 285
 Ser Cys Ala Leu Ser Leu Leu Ser Thr Gln Pro Trp Asp Thr Thr Thr
 290 295 300
 His Ser Ala Ala Ala Ser His Asn Gln Ala Ala Ala Met Ser Thr Thr
 305 310 315 320
 Thr Ser Phe Asp Gly Asn Pro Val Ala Pro Ser Ala Met Ala Gly Ser
 325 330 335

Tyr Met Ala Pro Ser Pro Trp Thr Gly Ser Arg Gly His Glu Gly Gly
340 345 350

Gly Arg Ser Val Ala His Gln Leu Pro His Glu Val Ser Leu Asp Glu
355 360 365

Val His Pro Gly Pro Ser His His Ala His Phe Ser Gly Glu Leu Glu
370 375 380

Leu Ala Leu Gln Gly Asn Gly Pro Ala Pro Ala Pro Arg Ile Asp Pro
385 390 395 400

Gly Ser Gly Ser Thr Phe Asp Gln Thr Ser Asn Thr Met Asp Trp Ser
405 410 415

Leu

<210> SEQ ID NO 8

<211> CHIỀU DÀI: 6

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: Peptit tổng hợp

<220> DẤU HIỆU:

<221> TÊN/CHÌA KHÓA: dấu hiệu_khác

<222> VỊ TRÍ: (1)..(6)

<223> THÔNG TIN KHÁC: Trình tự này có thể bao gồm 5 đến 6 gốc

<400> TRÌNH TỰ: 8

Arg Arg Arg Arg Arg Arg
1 5

<210> SEQ ID NO 9

<211> CHIỀU DÀI: 5

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<400> TRÌNH TỰ: 9

Arg Arg Arg Arg Arg
1 5

<210> SEQ ID NO 10

<211> CHIỀU DÀI: 10

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<220> DẤU HIỆU:

<221> TÊN/CHÌA KHÓA: dấu hiệu_khác

<222> VỊ TRÍ: (1)..(10)

<223> THÔNG TIN KHÁC: Trình tự này có thể bao gồm 2 đến 10 gốc

<400> TRÌNH TỰ: 10

His His His His His His His His His His
1 5 10

<210> SEQ ID NO 11

<211> CHIỀU DÀI: 6

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<400> TRÌNH TỰ: 11

His His His His His His
1 5

<210> SEQ ID NO 12

<211> CHIỀU DÀI: 8

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<400> TRÌNH TỰ: 12

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> SEQ ID NO 13

<211> CHIỀU DÀI: 8

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<400> TRÌNH TỰ: 13

Trp Ser His Pro Gln Phe Glu Lys
1 5

<210> SEQ ID NO 14

<211> CHIỀU DÀI: 10

<212> LOẠI: PRT

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: peptit tổng hợp

<400> TRÌNH TỰ: 14

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10

<210> SEQ ID NO 15
<211> CHIỀU DÀI: 31
<212> LOẠI: ADN
<213> SINH VẬT: Trình tự nhân tạo
<220> DẤU HIỆU:
<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 15

aggtaccgca atgtagagcc acgtaggcaa g 31

<210> SEQ ID NO 16
<211> CHIỀU DÀI: 34
<212> LOẠI: ADN
<213> SINH VẬT: Trình tự nhân tạo
<220> DẤU HIỆU:
<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 16

agggcccgct taccagctat tggttacaca tatt 34

<210> SEQ ID NO 17
<211> CHIỀU DÀI: 24
<212> LOẠI: ADN
<213> SINH VẬT: Trình tự nhân tạo
<220> DẤU HIỆU:
<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 17

cggtcgacta gctgcatctg ttgg 24

<210> SEQ ID NO 18
<211> CHIỀU DÀI: 25
<212> LOẠI: ADN
<213> SINH VẬT: Trình tự nhân tạo
<220> DẤU HIỆU:
<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 18

catcgtgttg ctggtttggt cgaag 25

<210> SEQ ID NO 19
<211> CHIỀU DÀI: 19
<212> LOẠI: ADN
<213> SINH VẬT: Trình tự nhân tạo
<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 19

ccctccacct cgtcctcag

19

<210> SEQ ID NO 20

<211> CHIỀU DÀI: 25

<212> LOẠI: ADN

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 20

agataacaac ggaagcataa aagtc

25

<210> SEQ ID NO 21

<211> CHIỀU DÀI: 29

<212> LOẠI: ADN

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 21

aggatcccca gccatgggat actactacc

29

<210> SEQ ID NO 22

<211> CHIỀU DÀI: 30

<212> LOẠI: ADN

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 22

aggtacccag cattaacact gataactaaa

30

<210> SEQ ID NO 23

<211> CHIỀU DÀI: 30

<212> LOẠI: ADN

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 23

aactagtcag cattaacact gataactaaa

30

<210> SEQ ID NO 24

<211> CHIỀU DÀI: 29

<212> LOẠI: ADN

<213> SINH VẬT: Trình tự nhân tạo

<220> DẤU HIỆU:

<223> THÔNG TIN KHÁC: Mô tả Trình tự nhân tạo: đoạn mỗi tổng hợp

<400> TRÌNH TỰ: 24

agagctccca gccatgggat actactacc

29