



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

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

(11)



1-0052327

(51)⁷ C12N 1/34; C12Q 1/68; C12N 5/04

(13) B

(21) 1-2018-02542

(22) 17/11/2016

(86) PCT/US2016/062548 17/11/2016

(87) WO 2017/087682 26/05/2017

(30) 62/256,902 18/11/2015 US; 62/300,507 26/02/2016 US

(45) 27/10/2025 451

(43) 26/11/2018 368A

(73) SYNGENTA CROP PROTECTION AG (CH)

Rosentalstrasse 67, 4058 Basel, Switzerland

(72) KELLIHER, Timothy Joseph (US); DELZER, Brent (US); CHINTAMANANI, Satya (US); SKIBBE, David Stewart (US); CHEN, Zhongying (US); STARR, Dakota (US); WENDEBORN, Sebastian (DE); LEDSON, Timothy Mark (US); FOWLER, Jeffrey, David (US).

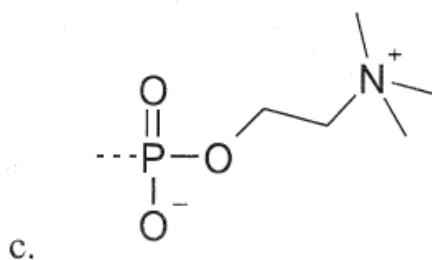
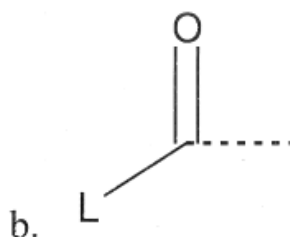
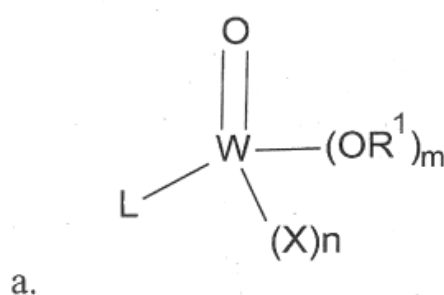
(74) Công ty TNHH Ban Ca (BANCA)

(54) PHƯƠNG PHÁP KÍCH THÍCH PHÔI ĐƠN BỘI TRONG PHÉP LAI GIỮA CÁC CÂY

(21) 1-2018-02542

(57) Sáng chế đề cập đến phương pháp kích thích các phôi đơn bội khi lai giữa hai cây, trong đó phương pháp bao gồm: (a) biểu hiện phospholipaza giống patatin đột biến ở cây; hoặc (b) dùng cho cây phân tử ARN can thiệp kích thước nhỏ bao gồm từ 18 đến 21 nucleotit của gen mã hóa phospholipaza giống patatin; hoặc (c) biến nạp cây với phospholipaza giống patatin đột biến; hoặc (d) gây đột biến trình tự phospholipaza giống patatin của cây sử dụng chỉnh sửa gen; trong đó cây được dùng làm cây bố mẹ trong khi lai giữa hai cây, sao cho việc lai tạo ra ít nhất một phôi đơn bội.

HÌNH 31.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc phát hiện chẩn đoán kích thích tạo đơn bội ("HI") hoặc sự vắng mặt và/hoặc có mặt của nó ở cây là hoặc không phải là dòng kích thích tạo đơn bội. Cụ thể hơn là, sáng chế đề cập đến axit nucleic mà có thể sử dụng để kích thích tạo HI ở cây và/hoặc hoạt tính sinh học mà có thể được biến đổi để tạo ra hoặc ngăn ngừa HI hoặc là ở cây mà có thể biểu hiện khác đi HI hoặc ở cây mà có thể không biểu hiện khác đi HI. Còn cụ thể hơn là, sáng chế đề cập đến phân tử axit nucleic mà mã hóa phân tử hoạt tính sinh học cũng như phương pháp sử dụng chúng để điều chỉnh HI ở cây.

Sáng chế đề xuất hàng loạt đột biến do con người tạo ra độc lập, thấy ở ít nhất một gen phospholipaza giống patatin A2 α ("PLA") của ngô; cây ngô có các đột biến này trong ít nhất một trong các gen PLA của chúng; và phương pháp tạo ra và nhận diện đột biến tương tự và/hoặc bổ sung trong gen PLA bằng cách sàng lọc cây ngô hỗn hợp và/hoặc cá thể. Cây ngô theo sáng chế kích thích tạo tính đơn bội do đột biến không chuyển gen ở ít nhất một trong các gen PLA của chúng. Sáng chế còn đề xuất phương pháp kích thích tạo mới kích thích tạo đơn bội bằng cách áp dụng hóa học lên mô sinh sản của cây. Sáng chế còn đề xuất phương pháp làm tăng tỷ lệ tạo hạt và tỷ lệ kích thích tạo đơn bội ("HIR", ở đây gọi là số lượng hạt đơn bội sống sót trên tổng số hạt sau khi bông được thụ phấn với phấn hoa của dòng kích thích tạo đơn bội), bằng cách áp dụng hóa học lên mô sinh sản của cây trong khi thụ phấn.

Tình trạng kỹ thuật của sáng chế

Sự thụ phấn là quá trình phức tạp. Hạt phấn của cây hạt kín chứa tế bào sinh dưỡng lớn và hai giao tử đực (tinh tử). Sau khi đáp xuống đầu nhụy, hạt nảy ra ống phấn mà thể hiện sự phát triển đầu nhanh chóng do nó hướng xuống đường truyền cái, được dẫn bởi các chất hấp dẫn hóa học tiết ra bởi hai trợ cầu ở đầu lỗ noãn của túi phôi. Trong quá trình truyền xuống ống phấn, tinh tử được nối với nhau và nhân sinh dưỡng bằng tế bào chất có sợi gọi là đơn vị mầm đực. Ngay sau khi tiếp xúc với một trong hai trợ cầu, ống phấn vỡ ra và hai tinh tử được đẩy ngang qua tế bào chất trợ cầu chết để kích hoạt lập với trứng và tế bào trung tâm của túi phôi, hoàn thành việc thụ phấn kép. Thậm chí

sau tiếp xúc ban đầu, sự kiện thụ phấn hồng có thể cấp cứu bằng ống phần thứ hai mà thụ phấn túi phôi thông qua tương tác với trợ cầu không rụng.

Người nhân giống lai các dòng bố mẹ nội dòng, một vật lai đóng vai trò là đực và một vật lai là cái, để tạo ra hạt lai. Quy trình phát triển dòng lai bố mẹ nội dòng mà cơ bản đồng nhất thường cần lai vật lai được lựa chọn và tự thụ phấn (tự lai) cho nhiều thế hệ để trở nên gần đồng nhất. Quy trình này tốn thời gian và đắt. Để rút ngắn thời gian để phát triển lai cùng dòng đồng nhất ở ngô, lúa, lúa mỳ, lúa mạch, và cây trồng khác, người nhân giống có thể chọn để sử dụng dòng kích thích tạo đơn bội để kích thích tạo sản xuất hạt đơn bội trên bố mẹ lai. Nhiễm sắc thể của cây đơn bội sau đó được nhân đôi, ví dụ bằng tác nhân nhân đôi nhiễm sắc thể như colchixin, để tạo ra dòng lai cùng dòng đồng nhất đơn bội kép.

Kích thích tạo đơn bội ("HI") là nhóm hiện tượng cây đặc trưng bởi việc mất nhiễm sắc thể kích thích tạo trong quá trình phát triển phôi. WO2012/030893, được kết hợp vào đây bằng cách viện dẫn, mô tả vùng của nhiễm sắc thể 1 của ngô mà có thể chịu trách nhiệm kích thích tạo đơn bội. Các gen đánh dấu được xác định trong vùng tăng kích thích tạo đơn bội được mô tả là ở giữa 48,249,509–51,199,249, mà liên quan đến gen đánh dấu công bố umc1169 có vị trí vật lý của (60,213,661). Vùng này có vẻ không thẳng hàng với vùng kích thích tạo đơn bội trong Stock 6. Dong và cộng sự, (2013) Theor. Appl. Genet. 126: 1713–1720, mô tả QTL ở trong bin 1.04 mà giải thích đến 66% sự biến thiên kiểu gen cho HIR.

Kích thích tạo đơn bội được quan sát thấy ở nhiều loài cây, như lúa miến, lúa mạch, lúa mỳ, và các cây thân cỏ khác. Ở ngô, HI xuất hiện là kết quả của việc sắp xếp lại, các đột biến trong, và/hoặc tái tổ hợp, chèn, hoặc loại bỏ trong vùng nhiễm sắc thể 1 (với loại trừ đáng chú ý của kích thích tạo đơn bội loại *ig*, mà là kết quả của đột biến trong gen *INDETERMINATE GAMETOPHYTE1* trên nhiễm sắc thể 3). Dòng HI nổi tiếng đã được nghiên cứu và xác định thô. Tuy nhiên, bằng chứng thử nghiệm chứng minh tác nhân di truyền gây ra HI ở ngô không có mặt. Cũng không có gen đánh dấu được liệt kê ở đây liên quan đến đặc điểm được xác định trước đó.

Ở hạt hoặc phôi đơn bội của ngô được tạo ra một cách đặc biệt bằng cách tạo lai giữa cây đực dòng kích thích tạo đơn bội (tức là, "phần hoa của dòng kích thích tạo đơn bội") và hầu như bất kỳ bông mà người ta chọn – bông có thể là của bất kỳ cùng dòng,

lai, hoặc vật liệu di truyền khác. Các thể đơn bội được tạo ra phần hoa của dòng kích thích tạo đơn bội ADN không được truyền đầy đủ và/hoặc duy trì thông qua sự phân chia tế bào đầu tiên của phôi. Kiểu hình thu được không thâm qua đủ, với một số noãn chứa các phôi đơn bội, và noãn khác chứa phôi lưỡng bội, phôi không bội, phôi khảm, hoặc phôi thui. Hạt đơn bội có phôi mà chỉ chứa ADN của mẹ cộng nội nhũ tam bội thường. Sau khi kích thích tạo đơn bội, các phôi hoặc hạt đơn bội thường được tách ra từ anh em lưỡng bội và không bội sử dụng sàng lọc gen đánh dấu kiểu hình hoặc di truyền và phát triển hoặc được nuôi cấy trong cây đơn bội. Các cây này sau đó được chuyển đổi hoặc là tự nhiên hoặc qua tác động hóa học (tức là, colchixin) vào cây đơn bội kép (DH) mà sau đó tạo ra hạt cùng dòng.

Các dòng HI chứa vị trí tính trạng định lượng ("QTL") trên Nhiễm sắc thể 1 chịu trách nhiệm đối với ít nhất 66% biến đổi trong kích thích tạo đơn bội. QTL gây ra kích thích tạo đơn bội ở các tỷ lệ khác nhau khi nó thâm nhập vào các nền khác nhau. Tất cả các dòng kích thích tạo đơn bội dùng trong ngành hạt là dẫn xuất của dòng HI tìm thấy, gọi là Stock6, và tất cả đều có dòng kích thích tạo đơn bội nhiễm sắc thể 1 QTL đột biến. Ở đây, tác giả đề mở đột biến chính trong QTL mà, khi được bổ sung, hỗ trợ sinh sản bình thường. Trong khi nguồn gốc của đột biến này là không rõ ràng, nó ở trong tất cả các dòng kích thích tạo, bao gồm Stock6.

Việc nhân giống cây được tạo điều kiện bằng cách sử dụng các cây đơn bội kép (DH). Việc sản xuất cây DH cho phép người nhân giống cây thu được dòng lai cùng dòng mà không lai cùng dòng đa thế hệ, do đó làm giảm thời gian cần để tạo ra cây đồng nhất. Cây DH tạo ra công cụ vô giá cho người nhân giống cây, đặc biệt là tạo ra dòng lai cùng dòng, bản đồ QTL, chuyển đổi bào tương, đưa vào gen đặc điểm, và sàng lọc F2 đối với cải thiện đặc điểm lưu lượng cao. Lượng lớn thời gian được tiết kiệm do các dòng đồng nhất cơ bản tạo ra trên một thế hệ, phù hợp nhu cầu đối với việc lai cùng dòng đa thế hệ thông thường. Cụ thể là, do cây DH hoàn toàn đồng nhất, chúng rất dễ để nghiên cứu di truyền định tính. Việc sản xuất hạt đơn bội là cần thiết cho quá trình lai đơn bội kép. Hạt đơn bội được tạo ra trên vật liệu di truyền của mẹ khi được thụ phấn với phần hoa từ dòng kích thích tạo phát triển do thể cái, như Stock 6.

Cây dòng kích thích tạo đơn bội ngô tạo ra phần hoa mà, khi lai lên vật liệu di truyền dòng không kích thích tạo, tạo ra sự phát triển do thể cái của các hạt đơn bội.

Không may là, quá trình này thường thu được tần suất hạt đơn bội thấp. Tần suất kích thích tạo đơn bội không đủ là yếu tố hạn chế trong chương trình lai đơn bội kép ở ngô.

Bản chất kỹ thuật của sáng chế

HIR cao cho phép tần suất hạt đơn bội cao hơn được tạo thành trên cây bố mẹ mong muốn. Cây bố mẹ có thể được sàng lọc trước bằng gen đánh dấu di truyền kết hợp với các đặc điểm mong muốn hoặc các đặc điểm được quan sát về kiểu hình để làm giàu khả năng di truyền của cây bố mẹ. Khi cây bố mẹ mong muốn được thụ phấn bởi dòng kích thích tạo đơn bội mà có HIR càng cao, thì thu được khả năng đơn bội kép mong muốn càng cao với kiểu gen và kiểu hình mong muốn.

Mặc dù quá trình đơn bội kép tạo ra sản xuất nhanh hơn vật lai cùng dòng đồng nhất, thể tích của vật lai cùng dòng đơn bội kép mà được tạo ra có thể bị giới hạn. Các dòng kích thích tạo đã biết, bao gồm nhưng không giới hạn ở: Stock 6, MHI (Dòng kích thích tạo đơn bội Moldovian), đột biến thể giao tử vô định ("*ig*"), KEMS, RWK, ZEM, ZMS, và KMS. Tất cả có HIR tương đối thấp. Stock 6, ví dụ, chỉ kích thích tạo 1–3% hạt đơn bội. Như vậy, việc kích thích tạo đơn bội là bước giới hạn tỷ lệ trong quy trình sản xuất dòng đơn bội kép.

Tác giả đã sáng chế ra cách để kích thích tạo sản xuất đơn bội và/hoặc làm tăng tỷ lệ kích thích tạo đơn bội ở cây bằng cách xử lý cây với hợp chất lipit, chất ức chế phospholipaza, và chất ức chế desaturaza axit béo. Một bộ trong các phương pháp này bao gồm áp dụng chất hóa học cụ thể lên mô sinh sản khi lai với phấn hoa kiểu đại (không kích thích tạo đơn bội). Lần thứ nhất, tác giả phát động kích thích tạo đơn bội *mới* bằng cách hóa học. Điều này được thực hiện bằng cách dùng dịch đặc của chất ức chế phospholipaza methyl alpha linolenyl flophosphonat (MALFP) lên hoa trong quá trình thụ phấn, mà dẫn đến tỷ lệ kích thích tạo đơn bội cao: lên đến 9% HIR. Một cách riêng biệt, tác giả phát động kích thích tạo đơn bội *mới* bằng cách dùng dịch đặc của arachidonyl flophosphonat (MAFP) lên hoa trong quá trình thụ phấn. Một cách riêng biệt, tác giả phát động kích thích tạo đơn bội *mới* bằng cách dùng dịch đặc của 1,2-distearoyl-sn-glyxero-3-phosphatidyl cholin (còn gọi là distearyl-phosphatidyl cholin; "DSPC") lên hoa trong quá trình thụ phấn. Tác giả cũng phát động kích thích tạo đơn bội *mới* bằng cách dùng hợp chất có

sẵn, bao gồm dầu ngô và dầu hạt lanh, cũng như axit linoleic tổng hợp hóa học, etyl este của axit oleic (OAE), metyl este của axit arachidonic, (AAME) và chất ức chế phospholipaza manoolit. Tác giả cũng làm tăng tỷ lệ kích thích tạo đơn bội khi lai các cây với dòng kích thích tạo đơn bội (ví dụ, RWK, Stock 6, hoặc ZMS) bằng cách cho dùng dịch đặc của chất ức chế phospholipaza metyl alpha linolenyl flophosphonat (MALFP). Ở các nồng độ bằng 2% MALFP+hỗn hợp chất hoạt động bề mặt 91 được nhũ hóa với đệm DML được đệm, tác giả đã nhân đôi tỷ lệ kích thích tạo đơn bội trong quá trình thụ phấn sử dụng RWK làm dòng thụ phấn. Tỷ lệ kích thích tạo thông thường trong RWK là khoảng 10 – 18%. Với MALFP được áp dụng tỷ lệ kích thích tạo tăng đến 20–35%. Tác giả cũng đã đẩy tỷ lệ kích thích tạo đơn bội bằng cách áp dụng các dịch đặc của axit linoleic (LLE), etyl este của axit linoleic (LLAE), và chất ức chế phospholipaza gọi là metyl arachidonyl flophosphonat (MAFP) lên hoa trong quá trình thụ phấn. Hơn nữa, làm như vậy theo cách này còn làm giảm tỷ lệ thui phôi và làm tăng tỷ lệ tạo thành hạt cho MALFP, LLAE, MAFP, và LLE. Cùng với các tác dụng này dẫn đến làm tăng tổng số các phôi đơn bội tái tạo trên bông sau quá trình thụ phấn.

Tác giả cũng tách dòng và xác định đặc điểm của gen chịu trách nhiệm kích thích tạo đơn bội ở ngô. Gen là PLA2 và nó có biểu hiện đặc hiệu phần hoa. Protein PLA2 được định vị lên tế bào chất tinh tử, có thể mạng lưới nội chất hoặc thể golgi. Việc xác định gen dẫn đến sáng chế ra một vài kỹ thuật mới để cải thiện quá trình kích thích tạo đơn bội, được xác định là đóng vai trò sản xuất các phôi đơn bội, hạt, hạt, hoặc cây bằng cách giao phấn bất kỳ bông với phần hoa của dòng kích thích tạo đơn bội. Việc xác định gen còn dẫn đến sáng chế ra các phương pháp mới để tạo ra đơn bội. Bộ khác bao gồm các phương pháp để tạo ra dòng kích thích tạo đơn bội mới bằng cách thay đổi trình tự của gen nguyên nhân, hoặc là thông qua đột biến nhắm đích, TILLING, hoặc CRISPR/Cas9. Biểu hiện của protein PLA2 có thể là điều chỉnh xuống sử dụng ARNi hoặc bằng cách sử dụng đột biến nhắm đích trong gen khởi động, 3' UTR, 5' UTR, hoặc vị trí nối.

Dựa trên việc xác định đột biến này, tác giả sáng chế ra các cách để biến đổi và cải thiện quá trình kích thích tạo đơn bội và, lần thứ nhất, tác giả bộc lộ phương pháp để tạo ra các thể đơn bội mới thông qua xử lý trong quá trình thụ phấn. Tác giả thể hiện các phương pháp làm tăng tỷ lệ kích thích tạo đơn bội ("HIR," tức là, phần trăm các phôi đơn bội tìm thấy trên bông được kích thích tạo đơn bội được cho) và cả phương pháp để

tăng tỷ lệ sống của hạt trong khi kích thích tạo đơn bội. Tác giả cũng bàn về các phương pháp để làm tăng dòng kích thích tạo đơn bội mới sử dụng biến đổi di truyền ("GM") hoặc chiến lược đột biến nhắm đích.

Quá trình kích thích tạo đơn bội có thể được cải thiện thông qua nhiều phương pháp. Thứ nhất, người ta có thể cố gắng cải thiện HIR trung bình. HIR là giới hạn tỷ lệ trong sản xuất cây đơn bội kép phạm vi rộng ("DH") vì HIR tương đối thấp. Xem Hình 1. Với loại trừ ít, HIR thường ít hơn 25% và thường thấy nhất là trong khoảng 10-20%, tức là 75-90% hạt là lưỡng bội. Với tần suất khá thấp, phôi không bội cũng được tạo ra trong khi kích thích tạo đơn bội. Các ví dụ của các hợp chất mà thể hiện làm tăng số các thể đơn bội tạo thành và/hoặc tỷ lệ tạo thành đơn bội trong lai dòng kích thích tạo đơn bội mà không làm tăng tỷ lệ thui phôi bao gồm methyl alpha linolenyl flophosphonat (MALFP), methyl arachidonyl flophosphonat (MAFP), axit linoleic (LLA), và etyl este của axit linoleic (LLAEE). Thứ hai, người ta có thể làm giảm tỷ lệ thui hạt và thụ phấn hỏng trong khi kích thích tạo đơn bội, dẫn đến hình thành nhiều nhân mỗi bông. Thui hạt xảy ra sau khi xảy ra thụ phấn thành công nhưng khi phôi chức năng hoặc nội nhũ không phát triển, và do hạt nhỏ, không màu phát triển thiếu phôi trưởng thành bên trong. Tỷ lệ thui hạt cao trong quá trình kích thích tạo đơn bội—trong khoảng 10–50% noãn được thụ phấn trên bông. Điều này hạn chế số hạt đơn bội người ta có thể tái tạo mỗi bông. Thụ phấn hỏng gây ra khi noãn không được thụ phấn bởi hạt phấn, và đặc trưng bởi sự vắng mặt bất kỳ sự phát triển hạt sau khi thụ phấn. Tỷ lệ này có xu hướng cao trong khi kích thích tạo đơn bội—trong khoảng 10–70% hạt trên một số bông kích thích tạo đơn bội thụ phấn hỏng, phụ thuộc vào sự tương ứng về độ chín của hoa đực và cái, kiểu lai được thực hiện, và di truyền đực và cái. Các ví dụ của các hợp chất mà thể hiện làm tăng tỷ lệ hình thành hạt sống được bằng cách làm giảm hỏng thụ phấn và thui phôi bao gồm bốn hợp chất đề cập ở trên: MALFP, MAFP, LLA, và LLAEE. Bốn phân tử này, khi được áp dụng cho phần hoa, núm tua hoặc phần khác của hoa trong khi lai kích thích tạo đơn bội, làm tăng số lượng các thể đơn bội tạo thành bằng cách tăng cả tỷ lệ kích thích tạo đơn bội và số hạt.

Cuối cùng, người ta có thể cải thiện quá trình kích thích tạo đơn bội bằng cách phủ nhận nhu cầu dòng kích thích tạo đơn bội đực, bằng cách cho phép kích thích tạo đơn bội *mới* để tạo ra trên bông hoặc trong noãn khi nó tự lai hoặc lai bằng phấn hoa từ dòng không kích thích tạo đơn bội. Các phương pháp kích thích tạo đơn bội mới này,

bao gồm việc áp dụng phospholipit như DSPC, axit béo như axit linolenic (LNA), hỗn hợp thông thường của triglycerit như dầu ngô và dầu hạt lanh, hoặc chất ức chế phospholipaza như MALFP, có thể được áp dụng trong bất kỳ lai xa hoặc tự thụ phấn để tạo ra đơn bội ở ngô. Ở đây tác giả mô tả các ví dụ mà rơi vào một hoặc nhiều trong số các loại này, tạo thành sự cải thiện hoặc sáng chế mới trong quy trình kích thích tạo đơn bội.

Sáng chế nhằm đến phương pháp để kích các phôi đơn bội trong việc lai giữa hai cây bố mẹ. Điều này được thực hiện bằng cách thay đổi biểu hiện của phospholipaza ở một trong số các cây bố mẹ. Việc thay đổi này có thể được thực hiện bằng một vài cách: hoặc là bằng cách khiến một trong số các cây bố mẹ biểu hiện phospholipaza đột biến; hoặc bằng cách dùng ARN can thiệp kích thước nhỏ cho một hoặc cả hai cây bố mẹ, mà khiến triệt phospholipaza; hoặc bằng cách biến nạp một trong số các cây bố mẹ với phospholipaza đột biến; hoặc bằng cách chỉnh sửa một trong số các phospholipaza của cây bố mẹ, ví dụ bằng cách đột biến định hướng điểm như các công nghệ trên cơ sở CRISPR hoặc TALEN. Khi biểu hiện của phospholipaza ở một trong các cây bố mẹ được thay đổi bởi một trong các kỹ thuật này, sau đó khi cây bố mẹ này được dùng trong việc lai, ít nhất một phôi đơn bội được tạo ra.

Theo một phương án của phương pháp, phospholipaza là phospholipaza giống patatin. Theo phương án khác, phospholipaza giống patatin là gen đồng đẳng của pPLAII α , mà được mã hóa bởi trình tự nucleotit bao gồm SEQ ID NO: 1 hoặc trình tự ít nhất 70% tương đồng với SEQ ID NO: 1. Trình tự nucleotit mã hóa phospholipaza giống patatin có thể được đột biến, và theo một phương án trình tự nucleotit có đột biến dịch khung mà tạo ra bộ ba mã kết thúc nhân tạo. Trình tự đột biến dịch khung bao gồm SEQ ID NO: 3 hoặc trình tự ít nhất 70% tương đồng với SEQ ID NO: 3.

Các đột biến khác là có thể và nằm trong phạm vi sáng chế. Sử dụng đột biến định hướng điểm có thể được dùng để tạo ra nhiều đột biến của phospholipaza. CRISPR/Cas9, TALENs, ngón tay kẽm, và meganucleaza là các phương pháp thực hiện đột biến định hướng điểm theo các phương án của sáng chế.

Sáng chế hữu dụng trong nhiều kiểu lai giữa các cây. Theo một phương án, các cây bố mẹ dùng trong việc lai là cây một lá mầm, như ngô, lúa, lúa mạch, và lúa mỳ. Các cây bố mẹ có thể là thuộc cùng loài cây một lá mầm, hoặc chúng có thể là các loài

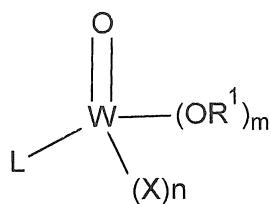
khác nhau. Theo phương án khác, các cây bố mẹ dùng trong việc lai là cây hai lá mầm, như đậu tương, hoa hướng dương, cà chua, ớt, củ cải đường, hoặc cải Bruxen. Theo phương án được ưu tiên, các cây bố mẹ là cây ngô hoặc cây lúa. Thuộc phạm vi sáng chế là phôi đơn bội được tạo ra bằng phương pháp, hạt đơn bội bao gồm phôi đơn bội, và cây đơn bội phát triển từ hạt đơn bội. Cũng thuộc phạm vi sáng chế là đơn bội kép được tạo ra bằng cách phôi lộ phôi đơn bội ra tác nhân nhân đôi nhiễm sắc thể, như colchixin hoặc trifluralin.

Sáng chế nhằm đến cADN bao gồm SEQ ID NO: 3, hoặc trình tự tương đồng gen với SEQ ID NO: 3, hoặc trình tự 70% tương đồng với SEQ ID NO: 3. Theo các phương án được ưu tiên, trình tự tương đồng gen với SEQ ID NO: 3 bao gồm các phospholipaza giống patatin từ ngô, lúa, lúa mỳ, đậu tương, và hoa hướng dương. Cụ thể là, trình tự tương đồng gen với SEQ ID NO: 3 bao gồm gen của lúa Os03g27610. Theo phương án được ưu tiên hơn, trình tự tương đồng gen với SEQ ID NO: 3 bao gồm các SEQ ID NO: 23 và 73–81.

Sáng chế nhằm đến cây chứa đột biến không chuyển gen do con người trong gen phospholipaza giống patatin của nó. Theo phương án được ưu tiên, gen phospholipaza giống patatin là pPLAII α . Theo phương án khác, đột biến tạo ra bộ ba mã kết thúc sớm được mã hóa trong gen. Theo phương án được ưu tiên hơn, cây là bất kỳ cây một lá mầm hoặc bất kỳ cây hai lá mầm, nhưng được đặc biệt ưu tiên là ngô hoặc lúa.

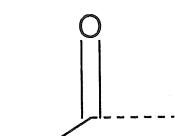
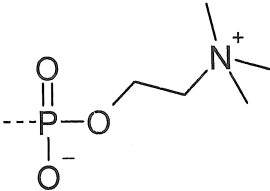
Sáng chế còn nhằm đến phương pháp kích thích tạo sản xuất các phôi và hạt đơn bội bằng cách xử lý mô sinh sản của cây với hợp chất bao gồm lipid hoặc chất ức chế phospholipaza. Theo một phương án, việc xử lý xảy ra trước, trong, hoặc ngay sau khi thụ phấn. Cây được xử lý có thể là bất kỳ cây một lá mầm hoặc cây hai lá mầm, nhưng theo các phương án được ưu tiên cây là ngô, lúa, lúa mỳ, đậu tương, hoa hướng dương, và củ cải đường. Theo phương án khác, lipid hoặc chất ức chế phospholipaza được chọn từ nhóm nêu trong Bảng 7. Theo các phương án được ưu tiên, việc xử lý hợp chất bao gồm methyl alpha-linolenoyl flophosphonat ("MALFP"), etyl este của axit linoleic ("LLAEE"), axit linoleic ("LLA"), dầu ngô, distearyl-phosphatidyl cholin ("DSPC"), hoặc methyl arachidonyl flophosphonat ("MAFP").

Theo phương án khác, lipid hoặc chất ức chế phospholipaza phù hợp với công



thức (I) sau:

. Trong phạm vi sáng chế, W có công thức (I) là cacbon ("C"), phospho ("P"), hoặc lưu huỳnh ("S"); m có thể là 0 hoặc 1; n có thể là 0 hoặc 1; X được chọn từ nhóm bao gồm OH, CN, O(C₁–C₄ alkyl), halogen, C₁–C₄ alkyl, và C₁–C₄ alkyl được thế bởi một, hai, hoặc ba halogen hoặc carbonyl; R¹ được chọn từ nhóm bao gồm H, C₁–C₆ alkyl, và C₁–C₆ alkyl được thế bởi một hoặc nhiều nhóm hydroxyl trong đó tùy ý một hoặc nhiều nhóm hydroxyl này được este hóa với gốc độc

lập được chọn từ nhóm bao gồm:  và ; hoặc R¹ được

liên kết với W khi W là S; mỗi L độc lập là mạch cacbon C₂–C₃₀, mạch cacbon này tùy ý bao gồm một hoặc nhiều nhóm độc lập được chọn từ alkenyl, alkynyl, phenyl, và heteroaryl, và mạch cacbon này tùy ý bị cắt bởi 1–6 nguyên tử oxy. Khi W là C và m là 1, sau đó n là 0; tuy nhiên, khi W là C và m là 0, sau đó n là 1.

Theo phương án được ưu tiên, X là F, Cl, CF₃, CCl₃, CF₂H, CCl₂H, CF₂CF₃, CCl₂CCl₃, CF₂Cl, CF₂CH₃, C(O)CH₃ hoặc CN. Theo phương án được ưu tiên hơn, halogen là F hoặc Cl.

Quan trọng là cần lưu ý rằng trong R¹, C₁–C₆ alkyl bao gồm các nhóm alkyl mạch thẳng, mạch nhánh, và mạch vòng. Theo phương án được ưu tiên, R¹ là C₁–C₆ alkyl được thế bởi một đến sáu nhóm hydroxyl.

Theo một phương án, mỗi L có công thức (I) độc lập là mạch cacbon C₂–C₃₀, bao gồm mạch nhánh, mà có thể được bão hòa, không bão hòa, hoặc không bão hòa đa. Theo phương án được ưu tiên, mạch cacbon của L bao gồm một đến bốn nhóm độc lập được chọn từ alkenyl, alkynyl, phenyl, và heteroaryl. Không bão hòa là dưới dạng liên kết đôi hoặc liên kết ba. Alkenyl hoặc alkynyl có thể trong mạch cacbon, hoặc đầu so với mạch cacbon. Các vòng phenyl và/hoặc heteroaryl có thể nối vào mạch cacbon ở vị trí ortho, meta, hoặc para, hoặc có thể ở đầu so với mạch cacbon. Các vòng aryl có thể tùy ý được

thể. Theo phương án được ưu tiên, mạch cacbon bị cắt bởi một đến sáu nguyên tử oxy. Như được sử dụng ở đây, "bị cắt bởi" nghĩa là mạch cacbon bao gồm ít nhất hai cacbon trong trình tự, sau nguyên tử oxy. Ví dụ, $-\text{CH}_2-\text{CH}_2-\text{O}-\text{CH}_2-\text{CH}_2-\text{CH}_3$ là mạch cacbon bị cắt bởi nguyên tử oxy. Theo phương án được ưu tiên, mạch cacbon bị cắt bởi một đến hai nguyên tử oxy.

Các ví dụ của mạch cacbon thích hợp theo các yêu cầu của L bao gồm: $(\text{CH}_2)_8-(\text{CH})_2-\text{CH}_2-(\text{CH})_2-\text{CH}_2-(\text{CH})_2-\text{CH}_2-\text{CH}_3$; $(\text{CH}_2)_3-(\text{CH})_2-\text{CH}_2-(\text{CH})_2-\text{CH}_2-(\text{CH})_2-\text{CH}_2-(\text{CH})_2-(\text{CH}_2)_4-\text{CH}_3$; $(\text{CH}_2)_7-(\text{CH})_2-(\text{CH}_2)_7-\text{CH}_3$; $(\text{CH}_2)_8-(\text{CH})_2-\text{CH}_2-\text{phenyl}-\text{CH}_2-(\text{CH})_2-\text{CH}_2-\text{CH}_3$; $(\text{CH}_2)_8-(\text{CH})_2-(\text{CH}_2)_2-\text{O}-\text{CH}_2-(\text{CH})_2-\text{CH}_2-\text{CH}_3$; và $(\text{CH}_2)_8-(\text{CH})_2-\text{CH}_2-\text{phenyl}-\text{O}-(\text{CH}_2)_3-\text{CH}_3$.

Việc xử lý các hợp chất này được thực hiện bằng cách áp dụng hợp chất bằng bất kỳ trong số các kỹ thuật sau: nhúng, tiêm, áp dụng tiếp xúc trên cơ sở phun, ống phun, áp dụng tiếp xúc trên cơ sở hút pipet, và áp dụng tiếp xúc trên cơ sở chải, và bất kỳ áp dụng tiếp xúc khác. Các phương án được ưu tiên sử dụng phun hoặc ống phun.

Sáng chế còn nhắm đến phương pháp làm tăng bộ hạt và làm giảm thui phôi ở cây trong khi kích thích tạo đơn bội, bao gồm xử lý mô sinh sản của cây, như tơ, núm tua, phấn hoa, bông, hạt, hoặc các mô tạo hoa khác, với nồng độ thích hợp của hợp chất trước, trong, hoặc sau khi thụ phấn. Theo một phương án, hợp chất được chọn từ nhóm bao gồm các thành phần của Bảng 7. Theo phương án khác, hợp chất là metyl alpha-linolenyl flophosphonat (MALFP). Theo phương án khác, hợp chất là axit linoleic (LLA), etyl este của axit linoleic (LLAEE), axit linolenic (LNA), distearoyl-phosphatidylcholin (DSPC), hoặc metyl arachidonyl flophosphonat (MAFP).

Sáng chế còn nhắm đến phương pháp làm tăng tỷ lệ kích thích tạo đơn bội ở cây, bao gồm việc áp dụng chế phẩm lipid cho các mô của cây ngay trước, trong, hoặc ngay sau khi thụ phấn. Theo một phương án, cây là cây một lá mầm hoặc cây hai lá mầm; hoặc cây là cây ngô hoặc cây lúa. Theo phương án khác, lipid đóng vai trò là chất ức chế phospholipaza và/hoặc chất ức chế desaturaza axit béo. Theo phương án khác, lipid là axit béo (ví dụ, LLA) hoặc este của axit béo (ví dụ, LLAEE) có chiều dài mạch và độ bão hòa cụ thể (mười tám cacbon, và hai liên kết đôi), mà là nhóm chiều dài axit béo mà không có phần hoa của dòng kích thích tạo đơn bội. Bằng cách minh họa và không giới hạn, lipid là, ví dụ, chất ức chế phospholipaza metyl alpha linolenyl flophosphonat

(MALFP), được hòa tan trong dung dịch DMSO đậm ở các nồng độ của MALFP nằm trong khoảng từ 0,0001 mg/mL đến 1g/mL, hoặc được hòa tan trong chế phẩm chất hoạt động bề mặt và sau đó được nhũ hóa trong dung dịch dimethylactamit (DML) đậm ở các nồng độ của MALF nằm trong khoảng từ 0,0001 mg/mL và 1 g/mL. Bằng cách minh họa thêm và không giới hạn, chế phẩm lipid được áp dụng bằng cách nhúng, tiêm, phun, phun sương, ống phun, rót, chải, hoặc bất kỳ phương pháp áp dụng khác lên mô sinh sản của cây. Theo một phương án, chế phẩm lipid kết hợp với phấn hoa trong hỗn hợp, hỗn hợp này sau đó được áp dụng cho các mô của cây. Theo phương án khác, hỗn hợp được áp dụng cho mô sinh sản của cây, ví dụ, phấn hoa hoặc tơ của cây ngô.

Sáng chế nhằm đến phương pháp kích mới để kích thích tạo đơn bội ở cây, bao gồm việc dùng hợp chất lipid cho ít nhất mô sinh sản của cây trong khi thụ phấn, trước khi thụ phấn, hoặc sau khi thụ phấn. Theo một phương án, cây được chọn từ nhóm bao gồm cây một lá mầm và cây hai lá mầm. Theo phương án khác, cây được chọn từ nhóm bao gồm lúa, ngô, lúa mỳ, lúa miến, cà chua, củ cải đường, cây kê, lúa mạch, đậu tương, hoa hướng dương, bông, yến mạch, thuốc lá, rau, cây ăn quả, và bất kỳ cây trồng khác.

Theo một phương án ví dụ, sáng chế bao gồm cây ngô hoặc cây lúa có khả năng kích thích tạo tính đơn bội do đột biến do con người ở gen phospholipaza giống patatin AII α ("PLA"), cũng như hạt, phấn hoa, các bộ phận của cây và thể hệ con của cây đó.

Theo phương án ví dụ khác nữa, sáng chế bao gồm cây ngô hoặc cây lúa có khả năng kích thích tạo các thể đơn bội được tạo ra bởi các bước thu được vật liệu cây từ cây ngô hoặc cây lúa bố mẹ, kích thích tạo ít nhất một đột biến ở ít nhất một bản sao của gen PLA của vật liệu cây bằng cách xử lý vật liệu cây với tác nhân gây đột biến để tạo ra vật liệu cây bị đột biến, nuôi cấy vật liệu cây bị đột biến để tạo ra cây lúa hoặc cây ngô thể hệ con, phân tích cây lúa hoặc cây ngô thể hệ con để phát hiện ít nhất một đột biến trong ít nhất một bản sao của gen PLA, lựa chọn cây lúa hoặc cây ngô thể hệ con mà có khả năng để tạo ra đơn bội so với cây lúa hoặc cây ngô bố mẹ; và lặp lại chu kỳ nuôi cấy cây lúa hoặc cây ngô thể hệ con để tạo ra cây thể hệ con bổ sung có khả năng để kích thích tạo đơn bội.

MÔ TẢ VẮN TẮT CÁC TRÌNH TỰ TRONG DANH MỤC TRÌNH TỰ

SEQ ID NO: 1 là trình tự cDNA cả đột biến unphospholipaza thấy ở GRMZM2G471240-NIL. Alen đột biến unphospholipaza ở đây gọi là *MATRILINEAL*.

SEQ ID NO: 2 là trình tự axit amin được mã hóa bởi SEQ ID NO: 1.

SEQ ID NO: 3 là trình tự nucleotit cADN của đột biến phospholipaza thấy ở GRMZM2G471240-mtl, bao gồm chèn 4 cặp bazơ. Alen đột biến phospholipaza ở đây gọi là *matrilineal*.

SEQ ID NO: 4 là trình tự axit amin được mã hóa bởi SEQ ID NO: 3.

SEQ ID NO: 5 là đoạn môi GRMZM2G471240_nil.F1.

SEQ ID NO: 6 là đoạn môi GRMZM2G471240_nil.R1.

SEQ ID NO: 7 là đoạn môi GRMZM2G471240_rwk.F1.

SEQ ID NO: 8 là đoạn môi GRMZM2G471240_rwk.R1.

SEQ ID NO: 9 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 39A ID T1 cá thể 22808-3954 alen 1.

SEQ ID NO: 10 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 23A T1 cá thể ID 22808-3924 alen 1.

SEQ ID NO: 11 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 81A T1 cá thể ID 22808-3932, Sự kiện 81A cá thể ID 22808-3317, và Sự kiện 81A cá thể ID 22808-3303.

SEQ ID NO: 12 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 39A ID 22808-3954 alen 2.

SEQ ID NO: 13 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 23A ID 22808-3924 alen 2.

SEQ ID NO: 14 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 38A T1 cá thể ID 22808-4108 alen 1.

SEQ ID NO: 15 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 18A T1 cá thể ID 22807-4016.

SEQ ID NO: 16 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 27A T1 cá thể ID 22807-4073 alen 1.

SEQ ID NO: 17 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 27A T1 cá thể ID 22807-4081 alen 1.

SEQ ID NO: 18 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 76A T1 cá thể ID 22873-3999.

SEQ ID NO: 19 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 32A T1 cá thể ID 22873-3991.

SEQ ID NO: 20 là trình tự nucleotit cho ARN dẫn CRISPR.

SEQ ID NO: 21 là trình tự nucleotit bộ gen cho Os03g27610, gen tương đồng PLA2 của lúa.

SEQ ID NO: 22 là trình tự cDNA cho SEQ ID NO: 21.

SEQ ID NO: 23 là trình tự axit amin được mã hóa bởi SEQ ID NO: 22.

SEQ ID NO: 24 là trình tự nucleotit của GRMZM2G471240-B73 không đột biến.

SEQ ID NO: 25 là trình tự nucleotit của GRMZM2G471240-RWK không đột biến.

SEQ ID NO: 26 là trình tự nucleotit của GRMZM2G471240-ST6 không đột biến.

SEQ ID NO: 27 là trình tự axit amin được mã hóa bởi SEQ ID NO: 24.

SEQ ID NO: 28 là trình tự axit amin được mã hóa bởi SEQ ID NO: 25.

SEQ ID NO: 29 là trình tự axit amin được mã hóa bởi SEQ ID NO: 26.

SEQ ID NO: 30 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22466, bao gồm *MATRILINEAL* kiểu dại.

SEQ ID NO: 31 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22467, bao gồm *PHOSPHOGLYXERAT MUTASE* kiểu dại.

SEQ ID NO: 32 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22503, bao gồm trình tự mã hóa cấu trúc vòng thân nhắm đích exon 2 của *MATRILINEAL*.

SEQ ID NO: 33 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22513, bao gồm trình tự mã hóa cấu trúc vòng thân nhắm đích exon 4 của *MATRILINEAL*.

SEQ ID NO: 34 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22807, bao gồm các trình tự mã hóa CRISPR/Cas9 chỉnh sửa cơ cấu nhắm đích *MATRILINEAL* trong NP2222.

SEQ ID NO: 35 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22808, bao gồm các trình tự mã hóa CRISPR/Cas9 chỉnh sửa cơ cấu nhắm đích *MATRILINEAL* trong NP2222.

SEQ ID NO: 36 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 22873, bao gồm các trình tự mã hóa CRISPR/Cas9 chỉnh sửa cơ cấu nhắm đích *MATRILINEAL* trong NP2222.

SEQ ID NO: 37 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 23123, bao gồm các trình tự mã hóa TALEN chỉnh sửa cơ cấu nhắm đích *MATRILINEAL* trong NP2222.

SEQ ID NO: 38 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 23501, gARN của lúa nhắm đích exon 4 với dẫn kép.

SEQ ID NO: 39 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 23501, gARN của lúa nhắm đích exon 4 dẫn đơn.

SEQ ID NO: 40 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 23501, gARN của lúa nhắm đích exon 1 với dẫn kép.

SEQ ID NO: 41 là trình tự nucleotit cho cat-xet biểu hiện của cấu trúc 23501, gARN của lúa nhắm đích exon 1 với dẫn đơn.

SEQ ID NO: 42 là trình tự nucleotit cho đột biến MTL do TALEN trong Sự kiện 38A ID 22808-4108 alen 2.

SEQ ID NO: 43 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 27A ID 22807-4073 alen 2.

SEQ ID NO: 44 là trình tự nucleotit cho đột biến MTL do CRISPR trong Sự kiện 27A ID 22807-4081 alen 2.

SEQ ID NO: 45 là trình tự nucleotit cho TILLING dòng 1139.

SEQ ID NO: 46 là trình tự nucleotit cho TILLING dòng 3594.

SEQ ID NO: 47 là trình tự nucleotit cho TILLING dòng 0505.

SEQ ID NO: 48 là trình tự nucleotit cho TILLING dòng 2658.

SEQ ID NO: 49 là trình tự nucleotit cho TILLING dòng 1983.

SEQ ID NO: 50 là trình tự nucleotit cho TILLING dòng 2732.

SEQ ID NO: 51 là trình tự nucleotit cho TILLING dòng 2414.

SEQ ID NO: 52 là trình tự axit amin được mã hóa bởi SEQ ID NO: 45.

SEQ ID NO: 53 là trình tự axit amin được mã hóa bởi SEQ ID NO: 46.

SEQ ID NO: 54 là trình tự axit amin được mã hóa bởi SEQ ID NO: 47.

SEQ ID NO: 55 là trình tự axit amin được mã hóa bởi SEQ ID NO: 48.

SEQ ID NO: 56 là trình tự axit amin được mã hóa bởi SEQ ID NO: 49.

SEQ ID NO: 57 là trình tự axit amin được mã hóa bởi SEQ ID NO: 50.

SEQ ID NO: 58 là trình tự axit amin được mã hóa bởi SEQ ID NO: 51.

SEQ ID NO: 59 là trình tự axit amin được mã hóa bởi SEQ ID NO: 9.

SEQ ID NO: 60 là trình tự axit amin được mã hóa bởi SEQ ID NO: 10.

SEQ ID NO: 61 là trình tự axit amin được mã hóa bởi SEQ ID NO: 11.

SEQ ID NO: 62 là trình tự axit amin được mã hóa bởi SEQ ID NO: 12.

SEQ ID NO: 63 là trình tự axit amin được mã hóa bởi SEQ ID NO: 13.

SEQ ID NO: 64 là trình tự axit amin được mã hóa bởi SEQ ID NO: 14.

SEQ ID NO: 65 là trình tự axit amin được mã hóa bởi SEQ ID NO: 15.

SEQ ID NO: 66 là trình tự axit amin được mã hóa bởi SEQ ID NO: 16.

SEQ ID NO: 67 là trình tự axit amin được mã hóa bởi SEQ ID NO: 17.

SEQ ID NO: 68 là trình tự axit amin được mã hóa bởi SEQ ID NO: 18.

SEQ ID NO: 69 là trình tự axit amin được mã hóa bởi SEQ ID NO: 19.

SEQ ID NO: 70 là trình tự axit amin được mã hóa bởi SEQ ID NO: 42.

SEQ ID NO: 71 là trình tự axit amin được mã hóa bởi SEQ ID NO: 43.

SEQ ID NO: 72 là trình tự axit amin được mã hóa bởi SEQ ID NO: 44.

SEQ ID NO: 73 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Sorghum bicolor*.

SEQ ID NO: 74 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Setaria italica*.

SEQ ID NO: 75 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Hordeum vulgare*.

SEQ ID NO: 76 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Brachypodium distachyon*.

SEQ ID NO: 77 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Oryza sativa* v. *indica*.

SEQ ID NO: 78 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Triticum aestivum*.

SEQ ID NO: 79 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Musa acuminata*.

SEQ ID NO: 80 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Elaeis guineensis*.

SEQ ID NO: 81 là trình tự axit amin cho gen đồng đẳng *MTL* thấy ở *Arabidopsis thaliana*.

Mô tả vắn tắt các hình vẽ

Hình 1 là sơ đồ tạo đồ được dùng để tạo bản đồ đặc điểm kích thích tạo đơn bội ở RWK.

Hình 2 thể hiện việc tạo bản đồ tỉ mỉ thu hẹp QTL chính đến khoảng cách rất nhỏ trong bin 1.04, trong khoảng từ 67,85 Mb đến 68,42 Mb. Vùng này có bảy gen được chú giải. Tác giả tạo trình tự và lắp các gen trong khoảng cách này ở một vài dòng. Hai gen với đột biến đột ngột nhất ở dòng kích thích tạo đơn bội được thể hiện ở dưới cùng bên phải (GRMZM2G471240 và GRMZM2G062320).

Hình 3 thể hiện sự khác nhau về biểu hiện của GRMZM2G471240 ở dòng kích thích tạo đơn bội và phần hoa không kích thích tạo và bao phần sau khi nở hoa (mô bào tử với hạt phần đã loại ra). Gen này được biểu hiện đặc hiệu trong thể giao tử đực.

Hình 4a thể hiện qRT-PCR đặc hiệu nổi tạo ra cho GRMZM2G471240. Ba mẫu lặp sinh học của bao phần giai đoạn R1 được thử thành bản sao ba kỹ thuật, và Ct trung bình và dung sai chuẩn được tính toán cho mỗi phản ứng. Số lượng tương đối của mỗi loại bản mã được so với đối chứng nội sinh sử dụng hồi quy log₂ của delta Ct. Hai bộ đoạn mỗi được dùng để đánh giá độ nhiễu tương đối của mỗi trong hai biến thể nổi được chú giải so với bộ đoạn mỗi mà không biết trước so với biến thể nổi. Biến thể bản mã ngắn hơn có độ nhiễu ít hơn tương đối so với bản mã dài ở cả kiểu gen NP2222 (kiểu đại) và NP2222-HI (dòng kích thích tạo đơn bội). Biểu hiện của các bản sao đột biến của gen ở NP2222-HI là cao hơn đáng kể đối với tất cả ba cặp đoạn mỗi được thử.

Hình 4b thể hiện năm mẫu lặp sinh học của phần hoa sạch từ cây NP2222 và *MTL^{TAL-FS}* (T1 mà đồng nhất đối với alen giống *mtl* được chỉnh sửa) được thử thành ba mẫu lặp kỹ thuật trên đoạn mỗi chung, và Ct trung bình và dung sai chuẩn được tính toán cho mỗi phản ứng. Số lượng tương đối của mỗi loại bản mã được so sánh với đối chứng nội sinh sử dụng hồi quy log₂ của delta Ct. Phần hoa *MTL^{TAL-FS}* có độ dồi dào bản mã thấp hơn so với phần hoa NP2222 (kiểu đại).

Hình 5a thể hiện sắp thẳng hàng axit amin của trình tự protein dự báo B73 của biến thể nổi dài của gen GRMZM2G471240 ở B73 và RWK-NIL, với trình tự dự báo của alen *mtl* thấy ở RWK và Stock 6 (S6). Axit amin mà khác là màu đỏ; axit amin mà khớp được thể hiện chữ màu xám bình thường, và bộ ba mã hóa kết thúc được thể hiện bằng dấu chấm. Hai đột biến điểm tạo ra sự thể axit amin, histidin (H) thành tyrosin (Y), và lysin (K) thành arginin (N). Các thay đổi này không bảo tồn; có thể là một hoặc cả hai làm biến đổi kiểu hình kích thích tạo đơn bội – đề xuất rằng chuỗi alen có thể không được bao trùm bằng cách khảo sát thêm các biến thể.

Hình 5b thể hiện MTL kiểu đại và đột biến (cắt cụt) MTL được mã hóa bởi alen *mtl* có hoạt tính phospholipaza trong ống nghiệm. Hoạt tính phospholipaza PLA2 được đo bằng thử nghiệm liposom huỳnh quang trên protein tinh chế, tái tổ hợp được tạo ra sử dụng cADN *MTL* và *mtl*. Thanh lỗi chỉ ra lỗi chuẩn dựa trên trung bình của bốn mẫu lặp.

Hình 6 thể hiện *mtl* chịu trách nhiệm đối với kiểu hình nhiễu tính trạng kết hợp với kích thích tạo đơn bội. **6A:** Tỷ lệ nảy mầm ống phần là tương tự ở các dòng kích thích tạo và dòng không kích thích tạo (n=200). **6B:** Sự dài ra của ống phần ban đầu

cũng tương tự (n=25). **6C**: RWK nhưng không phải RWK-NIL được đưa qua biến dạng tách (SD) dựa trên việc truyền đặc điểm thấp (25%) ở thể hệ con nảy mầm (n=300). **6D**: MTL/0 dòng bổ sung cũng thể hiện SD đối với mtl ở thể hệ con nảy mầm (n=400). **6E**: Đồ thị Venn thể hiện các kết quả nhận diện ARN-seq của hai cặp dòng kích thích tạo đơn bội-gần isogenic (trái, RWK với RWK-NIL; phải, NP2222-HI với NP2222; chữ đỏ, điều chỉnh lên; chữ xanh lá cây, điều chỉnh xuống). Chỉ 60 gen được thấy là thay đổi đáng kể trong cùng hướng.

Hình 7 thể hiện sắp thẳng hàng axit amin của gen MTL ở ngô với ortholog MTL hiện có rộng rãi trong tám cây thân cỏ, hai cây một lá mầm không phải thân cỏ, và *Arabidopsis* (thale cress). Sự sắp thẳng hàng này bao gồm ngô (*Zea mays*), lúa miến (*Sorghum bicolor*, 92% tương đồng trình tự với MTL), kê đuôi chồn (*Setaria italica*, 85% tương đồng), lúa mạch (*Hordeum vulgare*, 78% tương đồng), *Brachypodium distachyon* (78% tương đồng), lúa giống Indica và Japonica (*Oryza sativa* v. *indica* và *japonica*, Os3g27610, 78 và 79% tương đồng, tương ứng), lúa mì làm bánh (*Triticum aestivum*, 55% tương đồng), chuối (*Musa acuminata*, 57% tương đồng), cọ dầu (*Elaeis guineensis*, 56% tương đồng), và *Arabidopsis thaliana* (52% tương đồng).

Hình 8. Tương quan biểu hiện của phospholipaza của lúa (thích hợp với Singh, A., và cộng sự, *Rice phospholipase A superfamily: organization, phylogenetic and expression analysis during abiotic stresses and development*, PLOS ONE 7: e30947 (2012)). Tương đồng nhất với MTL là gen của lúa *OspPLAII* (Os3g27610).

Hình 9. Đồ thị thể hiện cách để chỉnh sửa Os3g27610 để tạo dòng kích thích tạo đơn bội. Có thể nhắm đích phần bất kỳ của gen (được thể hiện ở đây – nhắm đích exon thứ nhất và thứ tư) và kỳ vọng để tạo ra đột biến dịch khung mà có thể dẫn đến bất hoạt và mất chức năng của gen, và mà sẽ dẫn đến kích thích tạo đơn bội.

Hình 10 thể hiện cấu trúc nguyên tử của methyl alpha-linolenoyl flophosphonat (MALFP).

Hình 11 thể hiện cấu trúc nguyên tử của methyl arachidonyl flophosphonat (MAFP).

Hình 12 thể hiện cấu trúc nguyên tử của palmityl triflometylketon (PACOCF3).

Hình 13 thể hiện cấu trúc nguyên tử của arachidonyl triflometylketon (AACOCF₃).

Hình 14 thể hiện cấu trúc nguyên tử của manoalit.

Hình 15 thể hiện cấu trúc nguyên tử của etyl este của axit linoleic (LLAEE).

Hình 16 thể hiện cấu trúc nguyên tử của etyl este của axit linolenic (LNAEE).

Hình 17 thể hiện cấu trúc nguyên tử của metyl este của axit arachidonic (AAME).

Hình 18 thể hiện cấu trúc nguyên tử của axit oleic metyl este (OAME).

Hình 19 thể hiện cấu trúc nguyên tử của etyl este của axit oleic (OAEE).

Hình 20 thể hiện cấu trúc nguyên tử của etyl este của axit palmitic (PAEE).

Hình 21 thể hiện cấu trúc nguyên tử của etyl este của axit palmitoleic (PLAEE).

Hình 22 thể hiện cấu trúc nguyên tử của axit alpha-linolenic (αLNA).

Hình 23 thể hiện cấu trúc nguyên tử của axit gamma-linolenic (γLNA).

Hình 24 thể hiện cấu trúc nguyên tử của axit oleic.

Hình 25 thể hiện cấu trúc nguyên tử của axit linoleic.

Hình 26 thể hiện cấu trúc nguyên tử của axit arachidonic.

Hình 27 thể hiện cấu trúc nguyên tử của axit stearic.

Hình 28 thể hiện cấu trúc nguyên tử của axit linoleic liên hợp 9(Z)-11(E).

Hình 29 thể hiện cấu trúc nguyên tử của distearoyl phosphatidylcholin (DSPC).

Hình 30 thể hiện cấu trúc nguyên tử của 2-oleoyl-1-palmitoyl-sn-glyxero-3-phospho-etanolamin.

Hình 31 thể hiện cấu trúc nguyên tử chung cho các phân tử có thể hoạt động theo sáng chế yêu cầu bảo hộ.

Mô tả chi tiết sáng chế

ĐỊNH NGHĨA

Trong khi các thuật ngữ sau đây được tin là hiểu đúng bởi một trong số các chuyên gia trong lĩnh vực, các định nghĩa sau đây được nêu để tạo điều kiện giải thích đối tượng của sáng chế.

Tất cả các thuật ngữ kỹ thuật và khoa học được sử dụng ở đây, trừ khi được định nghĩa ở dưới đây theo cách khác, được dự tính có nghĩa giống như thường được hiểu bởi các chuyên gia trong lĩnh vực. Việc dẫn đến các kỹ thuật dùng ở đây là nhằm đề cập đến các kỹ thuật như được hiểu thông thường trong lĩnh vực, bao gồm các biến thể của các kỹ thuật này và/hoặc thay thế của các kỹ thuật tương đương mà có thể rõ ràng đối với một trong số các chuyên gia trong lĩnh vực. Trong khi các thuật ngữ sau đây được tin là hiểu đúng bởi một trong số các chuyên gia trong lĩnh vực, các định nghĩa sau đây được nêu để tạo điều kiện giải thích đối tượng của sáng chế

Theo điều ước luật sáng chế có từ lâu, thuật ngữ mạo từ "a", "an", và "the" đề cập đến "một hoặc nhiều" khi dùng trong sáng chế, bao gồm yêu cầu bảo hộ. Ví dụ, câu "tế bào" đề cập đến một hoặc nhiều tế bào, và theo một số phương án có thể đề cập đến mô và/hoặc cơ quan. Tương tự, câu "ít nhất một", khi được dùng ở đây để đề cập đến toàn bộ, đề cập đến, ví dụ, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, hoặc hơn của toàn bộ đó, bao gồm nhưng không giới hạn ở tất cả toàn bộ giá trị số giữa 1 và 100 cũng như toàn bộ số lớn hơn 100.

Trừ khi có quy định khác, tất cả các số biểu hiện số lượng của thành phần, điều kiện phản ứng, và v.v. dùng trong bản mô tả và yêu cầu bảo hộ được hiểu là được thay đổi trong tất cả các trường hợp bởi thuật ngữ "khoảng." Thuật ngữ "khoảng," như được sử dụng ở đây khi đề cập đến giá trị đo được như lượng của khối lượng, trọng lượng, thời gian, thể tích, nồng độ hoặc phần trăm nghĩa là bao gồm các biến thể theo một số phương án $\pm 20\%$, theo một số phương án $\pm 10\%$, theo một số phương án $\pm 5\%$, theo một số phương án $\pm 1\%$, theo một số phương án $\pm 0,5\%$, và theo một số phương án $\pm 0,1\%$ từ lượng cụ thể, như vậy các biến thể là thích hợp để thực hiện các phương pháp được bộc lộ và/hoặc sử dụng các chế phẩm, axit nucleic, polypeptit, v.v. được bộc lộ. Theo đó, trừ khi có quy định khác, các tham số bằng số được nêu trong bản mô tả này và yêu cầu bảo hộ kèm theo là xấp xỉ mà có thể thay đổi phụ thuộc vào đặc điểm mong muốn cần đạt được bởi đối tượng của sáng chế.

Như được sử dụng ở đây, thuật ngữ "alen" đề cập đến dạng trình tự biến thể hoặc biến đổi ở vị trí di truyền. Trong lưỡng bội, alen đơn được thừa hưởng bởi thể hệ con cá thể tách biệt với mỗi bố mẹ ở mỗi vị trí. Hai alen của vị trí được cho có mặt trong sinh vật lưỡng bội chiếm vị trí tương ứng trên cặp nhiễm sắc thể tương đồng, mặc dù một

trong số các chuyên gia trong lĩnh vực hiểu rằng các alen trong bất kỳ cá thể cụ thể không cần đại diện tất cả các alen mà có trong các loài.

Như được sử dụng ở đây, thuật ngữ "và/hoặc" khi dùng trong bối cảnh của danh mục các thực thể, đề cập đến các thực thể có mặt đơn lẻ hoặc trong kết hợp. Do đó, ví dụ, câu "A, B, C, và/hoặc D" bao gồm A, B, C, và D riêng lẻ, mà còn bao gồm bất kỳ và tất cả các kết hợp và mức dưới kết hợp A, B, C, và D (ví dụ, AB, AC, AD, BC, BD, CD, ABC, ABD, và BCD). Theo một số phương án, một trong nhiều yếu tố mà "và/hoặc" đề cập đến có thể còn có mặt riêng lẻ trong các xuất hiện đơn hoặc đa trong (các) kết hợp và/hoặc (các) mức dưới kết hợp.

Như được sử dụng ở đây, câu "kết hợp với" đề cập đến mối quan hệ nhận diện được và/hoặc phân tích được giữa hai thực thể. Ví dụ, câu "kết hợp với HI" đề cập đến đặc điểm, vị trí, gen, alen, gen đánh dấu, kiểu hình, v.v., hoặc biểu hiện của chúng, sự có mặt hoặc không có mặt của chúng có thể ảnh hưởng đến phạm vi và/hoặc mức độ mà cây hoặc thể hệ con của nó biểu hiện HI. Như vậy, gen đánh dấu "kết hợp với" đặc điểm khi được liên kết với nó và khi sự có mặt của gen đánh dấu là chỉ báo của việc có hay không và/hoặc đến mức độ nào đặc điểm mong muốn hoặc dạng đặc điểm sẽ xảy ra ở cây/vật liệu di truyền bao gồm gen đánh dấu. Tương tự, gen đánh dấu "kết hợp với" alen khi được liên kết với nó và khi sự có mặt của gen đánh dấu là chỉ báo của việc có hay không alen có mặt ở cây/vật liệu di truyền bao gồm gen đánh dấu. Ví dụ, "gen đánh dấu kết hợp với HI" đề cập đến gen đánh dấu mà sự có mặt hoặc không có mặt của nó có thể dùng để dự báo việc có hay không và/hoặc đến mức độ nào cây sẽ thể hiện kích thích tạo đơn bội.

Thuật ngữ "bao gồm," mà đồng nghĩa với "gồm," "chứa," và "khác biệt ở chỗ," là tính toán bộ hoặc mở cuối và không loại trừ các chi tiết và/hoặc các bước của phương pháp bổ sung, không kể ra. "Bao gồm" là thuật ngữ trong lĩnh vực mà có nghĩa là các chi tiết và/hoặc các bước có tên có mặt, nhưng các chi tiết và/hoặc các bước khác có thể được thêm và vẫn thuộc phạm vi của đối tượng liên quan.

Như được sử dụng ở đây, câu "bao gồm" không gồm bất kỳ chi tiết, bước, hoặc thành phần không được liệt kê cụ thể. Khi câu "gồm có" xuất hiện trong mệnh đề của thân yêu cầu bảo hộ, hơn là ngay sau lời nói đầu, nó giới hạn chỉ chi tiết được nêu trong mệnh đề này; các chi tiết khác không được loại trừ khỏi yêu cầu bảo hộ về mặt tổng thể.

Như được sử dụng ở đây, câu "cơ bản gồm có" giới hạn phạm vi phần mô tả liên quan hoặc yêu cầu bảo hộ cho các vật liệu và/hoặc các bước cụ thể, cộng với các loại mà không ảnh hưởng về mặt vật chất đến (các) đặc điểm cơ bản và mới của đối tượng được bộc lộ và/hoặc được yêu cầu bảo hộ.

Liên quan đến thuật ngữ "bao gồm," "cơ bản gồm có," và "gồm có," khi một trong ba thuật ngữ này được sử dụng ở đây, đối tượng được bộc lộ theo sáng chế và được yêu cầu bảo hộ có thể bao gồm theo một số phương án việc sử dụng của một trong số hai thuật ngữ còn lại. Ví dụ, nếu đối tượng theo một số phương án đề cập đến axit nucleic mà mã hóa polypeptit bao gồm trình tự axit amin mà ít nhất 95% tương đồng với SEQ ID NO: 2 hoặc 3. Cần hiểu rằng đối tượng được bộc lộ do đó còn bao gồm axit nucleic mà mã hóa polypeptit theo một số phương án cơ bản gồm có trình tự axit amin mà ít nhất 95% tương đồng với SEQ ID NO: 2 hoặc 3 cũng như axit nucleic mà mã hóa polypeptit mà theo một số phương án gồm có trình tự axit amin mà ít nhất 95% tương đồng với SEQ ID NO: 2 hoặc 3. Tương tự, cũng cần hiểu rằng theo một số phương án phương pháp cho đối tượng được bộc lộ bao gồm các bước mà được bộc lộ ở đây, theo một số phương án phương pháp cho đối tượng của sáng chế cơ bản gồm các bước mà được bộc lộ, và theo một số phương án phương pháp cho đối tượng của sáng chế gồm có các bước mà được bộc lộ ở đây.

Như được sử dụng ở đây, thuật ngữ "kích thích tạo đơn bội mới" đề cập đến việc khởi đầu kích thích tạo đơn bội bằng cách đưa vào tác nhân kích thích tạo đơn bội tự phát. Việc đưa vào này có thể được thực hiện bằng các phương pháp phun tiếp xúc, thụ phấn bằng tay, đột biến, hoặc chuyển gen. Thuật ngữ "kích thích tạo đơn bội mới," "HI mới," và "kích thích tạo đơn bội mới" được sử dụng hoán đổi nhau trong toàn bộ bản mô tả.

Như được sử dụng ở đây, thuật ngữ "gen" đề cập đến đơn vị di truyền bao gồm trình tự của ADN mà chiếm vị trí cụ thể trên nhiễm sắc thể và chứa lệnh di truyền đối với đặc điểm cụ thể hoặc đặc điểm ở sinh vật.

"Biểu đồ di truyền" là mô tả mối quan hệ liên kết di truyền trong số các vị trí trên một hoặc nhiều nhiễm sắc thể trong các loài được cho, thường được mô tả dưới dạng biểu đồ hoặc dạng bảng.

Như được sử dụng ở đây, cây được đề cập đến là "thể đơn bội" có bộ đơn (bộ gen) nhiễm sắc thể và số giảm nhiễm sắc thể (n) ở cây đơn bội bằng với của giao tử.

Như được sử dụng ở đây, cây đề cập đến "thể đơn bội kép" được phát triển bằng cách nhân đôi bộ đơn bội của nhiễm sắc thể. Cây hoặc hạt mà thu được từ cây đơn bội kép mà tự lai bất kỳ số thế hệ có thể vẫn được xác định là cây đơn bội kép. Cây đơn bội kép được coi là cây đồng hợp tử. Cây được coi là thể đơn bội kép nếu nó có khả năng sinh sản, ngay cả khi toàn bộ phần sinh dưỡng của cây không gồm các tế bào có bộ nhiễm sắc thể kép; tức là, cây sẽ được coi là thể đơn bội kép nếu nó chứa giao tử sống được, ngay cả khi nó là thể khảm.

Như được sử dụng ở đây, thuật ngữ "đột biến do con người" đề cập đến bất kỳ đột biến mà xảy ra do hoặc là tác động trực tiếp hoặc gián tiếp của con người. Thuật ngữ này bao gồm, nhưng không giới hạn ở, các đột biến thu được bằng bất kỳ phương pháp gây đột biến nhắm đích.

Như được sử dụng ở đây, thuật ngữ "mẫu dò gen đánh dấu" và "mẫu dò" đề cập đến trình tự nucleotit hoặc phân tử axit nucleic mà có thể được sử dụng để phát hiện sự có mặt hoặc vắng mặt của trình tự trong trình tự lớn hơn, ví dụ, mẫu dò axit nucleic mà bổ sung cho tất cả hoặc một phần gen đánh dấu hoặc vị trí đánh dấu, thông qua lai axit nucleic. Mẫu dò gen đánh dấu bao gồm khoảng 8, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 hoặc hơn nucleotit liên tiếp có thể được dùng để lai axit nucleic.

Như được sử dụng ở đây, thuật ngữ "gen đánh dấu phân tử" có thể được sử dụng để đề cập đến gen đánh dấu di truyền, như được xác định ở trên, hoặc sản phẩm mã hóa của chúng (ví dụ, protein) dùng làm điểm tham khảo khi xác định sự có mặt /không có mặt của vị trí liên quan đến HI. Gen đánh dấu phân tử có thể được dẫn xuất từ trình tự nucleotit di truyền hoặc từ trình tự nucleotit được biểu hiện (ví dụ, từ ARN, cADN, v.v.). Thuật ngữ còn đề cập đến trình tự nucleotit bổ sung cho hoặc kẹp sườn trình tự gen đánh dấu, như trình tự nucleotit dùng làm mẫu dò và/hoặc đoạn môi có khả năng khuếch đại trình tự gen đánh dấu. Trình tự nucleotit là "bổ sung" khi chúng lai đặc hiệu trong dung dịch (ví dụ, theo quy luật bắt cặp bazơ Watson-Crick). Thuật ngữ này còn đề cập đến gen đánh dấu di truyền mà chỉ ra đặc điểm bằng cách không có mặt trình tự nucleotit bổ sung cho hoặc kẹp sườn trình tự gen đánh dấu, như trình tự nucleotit dùng làm mẫu dò và/hoặc đoạn môi có khả năng khuếch đại trình tự gen đánh dấu.

Như được sử dụng ở đây, thuật ngữ "trình tự nucleotit," "polynucleotit," "trình tự axit nucleic," "phân tử axit nucleic," và "đoạn axit nucleic" đề cập đến polyme của ARN

hoặc ADN mà là mạch đơn hoặc mạch đôi, tùy ý chứa bazơ nucleotit tổng hợp, không tự nhiên, và/hoặc được thay đổi. "Nucleotit" là đơn vị monome mà các polyme ADN hoặc ARN được cấu trúc từ đó và gồm có bazơ purin hoặc pyrimidin, pentoza, và nhóm axit phosphoric. Các nucleotit (thường được tìm thấy ở dạng 5'-monophosphat của chúng) được đề cập đến bằng ký hiệu một ký tự của chúng như sau: "A" cho adenylat hoặc deoxyadenylat (lần lượt đối với ARN hoặc ADN), "C" cho xytylalat hoặc deoxyxytylalat, "G" cho guanylat hoặc deoxyguanylat, "U" cho uridylyat, "T" cho deoxytymylalat, "R" cho purin (A hoặc G), "Y" cho pyrimidin (C hoặc T), "K" cho G hoặc T, "H" cho A hoặc C hoặc T, "I" cho inosin, và "N" cho nucleotit bất kỳ.

Như được sử dụng ở đây, thuật ngữ "sự tương đồng trình tự nucleotit" chỉ sự có mặt của các nucleotit giống nhau ở các vị trí tương ứng của hai polynucleotit. Polynucleotit có các trình tự "tương đồng" nếu trình tự của các nucleotit trong hai polynucleotit này là giống nhau khi được sắp thẳng hàng để tương ứng tối đa (ví dụ, trong cửa sổ so sánh). Việc so sánh trình tự giữa hai hoặc nhiều polynucleotit thường được thực hiện bằng cách so sánh các phần của hai trình tự này trên cửa sổ so sánh để nhận dạng và so sánh các vùng lân cận có sự tương tự về trình tự. Cửa sổ so sánh thường là khoảng từ 20 đến 200 nucleotit liên tiếp. "Tỷ lệ phần trăm tương đồng trình tự" đối với các polynucleotit, chẳng hạn như độ tương đồng trình tự bằng khoảng 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 hoặc 100 phần trăm, có thể được xác định bằng cách so sánh hai trình tự được sắp thẳng hàng tối ưu trên cửa sổ so sánh, trong đó một phần của trình tự polynucleotit trong cửa sổ so sánh có thể bao gồm sự thêm hoặc sự làm khuyết (tức là, khoảng trống) so với trình tự tham chiếu đối với sự sắp thẳng hàng tối ưu của hai trình tự này. Theo một số phương án, phần trăm được tính toán bởi: (a) xác định số lượng vị trí ở đó bazơ axit nucleic giống nhau xuất hiện ở cả hai trình tự; (b) chia số lượng của các vị trí bắt cặp cho tổng số lượng vị trí trong cửa sổ so sánh; và (c) nhân kết quả này với 100. Việc sắp thẳng hàng tối ưu các trình tự để so sánh có thể được tiến hành bằng cách thực hiện trên máy tính các thuật toán đã biết, hoặc bằng cách kiểm tra bằng mắt. Các thuật toán so sánh trình tự đã có và sắp thẳng hàng đa trình tự là, tương ứng, Basic Local Alignment Search Tool (BLAST) và chương trình ClustalW/ClustalW2/Clustal Omega có trên mạng Internet (ví dụ, trang web của EMBL-EBI). Các chương trình thích hợp khác bao gồm, nhưng không chỉ giới hạn ở, GAP, BestFit, Plot Similarity, và FASTA, mà là một phần của gói Accelrys GCG có sẵn từ

Accelrys, Inc. San Diego, California, Mỹ. *Cũng xem* Smith & Waterman, 1981; Needleman & Wunsch, 1970; Pearson & Lipman, 1988; Ausubel và cộng sự, 1988; và Sambrook & Russell, 2001.

Một ví dụ về thuật toán mà thích hợp để xác định phần trăm tương đồng trình tự và tương tự trình tự là thuật toán BLAST, mà được mô tả trong Altschul và cộng sự, 1990. Theo một số phương án, phần trăm tương đồng trình tự đề cập đến tương đồng trình tự trên chiều dài đầy đủ của một trong số gADN, cADN, hoặc trình tự protein định trước trong ORF lớn nhất của SEQ ID No: 1 được so sánh. Theo một số phương án, việc tính toán để xác định phần trăm tương đồng trình tự axit nucleic không bao gồm trng tính toán bất kỳ vị trí nucleotit trong đó mỗi trong số các axit nucleic được so sánh bao gồm "N" (tức là, khi bất kỳ nucleotit có thể có mặt ở vị trí đó).

Thuật ngữ "khung đọc mở" (ORF) đề cập đến trình tự axit nucleic mà mã hóa polypeptit. Theo một số phương án, ORF bao gồm bộ ba mã khởi đầu dịch mã, bộ ba mã kết thúc dịch mã (tức là, kết thúc), và trình tự axit nucleic giữa chúng mà mã hóa axit amin có trong polypeptit. Thuật ngữ "bộ ba mã khởi đầu" và "bộ ba mã kết thúc" đề cập đến đơn vị ba nucleotit liền nhau (tức là, bộ ba mã) trong trình tự mã hóa mà cụ thể hóa khởi đầu và kết thúc mạch, tương ứng, của tổng hợp protein (dịch mã mARN).

Phospholipaza giống patatin A2 α còn có thể được gọi là PLA, pPLA, pPLAIIA pPLAII α , PLA2alpha, hoặc PLA2, hoặc biến thể tương tự khác. Phospholipaza giống patatin AII α còn được gọi là *MATRILINEAL*. Các thuật ngữ này được sử dụng hoán đổi nhau trong bản mô tả. Gen *MATRILINEAL* bao gồm bốn cặp bazơ đột biến dịch khung ở đây gọi là *matrilineal*.

Như được sử dụng ở đây, thuật ngữ "kiểu hình," "đặc điểm kiểu hình" hoặc "đặc điểm" đề cập đến một hoặc nhiều đặc điểm của cây hoặc tế bào cây. Kiểu hình có thể quan sát được bằng mắt thường, hoặc bằng các phương tiện đánh giá bất kỳ khác được biết trong lĩnh vực này, ví dụ, kính hiển vi, phân tích sinh hóa, hoặc thử nghiệm cơ điện. Trong một số trường hợp, kiểu hình được kiểm soát trực tiếp bởi gen đơn hoặc vị trí di truyền (tức là, tương ứng với "đặc điểm gen đơn"). Trong trường hợp sử dụng kích thích tạo đơn bội của gen đánh dấu màu sắc, như R Navajo, và các gen đánh dấu khác bao gồm gen chuyển nhìn thấy bằng mắt bởi sự có mặt hoặc sự không có mặt của màu sắc trong bằng chứng hạt nếu hạt là hạt đơn bội được kích thích tạo. Việc sử dụng R Navajo

là gen đánh dấu màu sắc và sử dụng gen chuyển là đã biết trong lĩnh vực là phương tiện để phát hiện việc kích thích tạo hạt đơn bội trên cây cái. Trong các trường hợp khác, kiểu hình là kết quả của sự tương tác trong số một vài gen, mà theo một số phương án còn tạo ra từ tương tác của cây và/hoặc tế bào cây với môi trường của nó.

Như được sử dụng ở đây, thuật ngữ "cây" có thể đề cập đến toàn bộ cây, phần bất kỳ của nó, hoặc tế bào hoặc nuôi cấy mô dẫn xuất từ cây. Do đó, thuật ngữ "cây" có thể đề cập đến bất kỳ trong số: toàn bộ cây, các thành phần của cây hoặc cơ quan (ví dụ, lá, thân, rễ, v.v.), mô cây, hạt và/hoặc tế bào cây.

Tế bào cây là tế bào của cây, lấy từ cây, hoặc được dẫn xuất qua dịch nuôi cấy từ tế bào lấy từ cây. Do đó, thuật ngữ "tế bào cây" bao gồm không giới hạn tế bào trong hạt, dịch nuôi cấy huyền phù, phôi, vùng mô phân sinh, mô thể sần, lá, chồi, thể giao tử, thể bào tử, phấn hoa, và vi cầu. Câu "bộ phận của cây" đề cập đến bộ phận của cây, bao gồm tế bào đơn và các mô tế bào như tế bào cây mà nguyên vẹn ở cây, cụm tế bào, và nuôi cấy mô mà cây có thể tái sinh từ đó. Các ví dụ của các bộ phận của cây bao gồm, nhưng không giới hạn ở, tế bào đơn và các mô từ phấn hoa, noãn, lá, phôi, rễ, đầu rễ, bao phấn, hoa, quả, thân, chồi, và hạt; cũng như mầm, thân rễ, thể nguyên sinh, thể sần, và tương tự.

Như được sử dụng ở đây, thuật ngữ "đoạn môi" đề cập đến oligonucleotit mà có khả năng ủ thành đích axit nucleic (theo một số phương án, ủ đặc hiệu thành đích axit nucleic) cho phép ADN polymeraza và/hoặc transcriptaza ngược gắn vào đó, nhờ đó dùng làm điểm khởi đầu phân tích ADN khi đặc dưới các điều kiện trong đó việc tổng hợp sản phẩm mở rộng đoạn môi được kích thích tạo (ví dụ, với sự có mặt của nucleotit và tác nhân polyme hóa như ADN polymeraza và ở nhiệt độ và pH thích hợp). Theo một số phương án, một hoặc nhiều các đoạn môi được dùng để khuếch đại axit nucleic của cây (ví dụ, sử dụng phản ứng chuỗi polymeraza; PCR).

Như được sử dụng ở đây, thuật ngữ "mẫu dò" đề cập đến axit nucleic (ví dụ, axit nucleic mạch đơn hoặc mạch của axit nucleic mạch đôi hoặc thứ tự cao hơn, hoặc dưới trình tự của nó) mà có thể tạo thành phức liên kết với trình tự bổ sung trong trình tự axit nucleic đích. Thông thường, mẫu dò có chiều dài đủ để tạo thành phân tử phức ổn định và đặc hiệu trình tự với bổ sung của nó, và như vậy có thể được sử dụng theo một số phương án để phát hiện trình tự mong muốn có mặt trong nhiều axit nucleic.

Như được sử dụng ở đây, các thuật ngữ "thế hệ con" và "thực vật thế hệ con" chỉ thực vật được tạo ra từ việc sinh sản vô tính hoặc hữu tính từ một hoặc nhiều cây bố mẹ. Trong kích thích tạo đơn bội hạt trên cây cái là đơn bội, do đó không phải là thế hệ con của dòng đơn bội kích thích tạo. Thế hệ con của hạt đơn bội không phải chỉ thế hệ con mong muốn. Còn có hạt HI và cây sau đó và thế hệ con của hạt của cây kích thích tạo đơn bội. Cả hạt đơn bội và hạt HI có thể là thế hệ con. Cây thế hệ con có thể thu được bằng cách tách dòng hoặc tự lai cây bố mẹ đơn, hoặc bằng cách lai hai hoặc nhiều cây bố mẹ. Ví dụ, cây thế hệ con có thể thu được bằng cách tách dòng hoặc tự tạo cây bố mẹ hoặc bằng cách lai hai hoặc nhiều cây bố mẹ và bao gồm tự lai cũng như F_1 hoặc F_2 hoặc các thế hệ khác nữa. F_1 là thế hệ con thế hệ thứ nhất được tạo ra từ bố mẹ ít nhất một trong số đó được dùng lần đầu tiên là vật cho đặc điểm, trong khi thế hệ con của thế hệ thứ hai (F_2) hoặc thế hệ tiếp theo (F_3 , F_4 , và tương tự) là các mẫu được tạo ra từ việc tự lai, lai chéo, lai ngược, và/hoặc lai khác của F_1 s, F_2 s, và tương tự. F_1 do đó có thể là (và theo một số phương án là) cây lai có từ việc lai giữa hai bố mẹ sinh sản thực (*tức là*, bố mẹ mà là sinh sản thực là mỗi đồng nhất cho đặc điểm mong muốn hoặc alen của chúng), trong khi đó F_2 có thể là (và theo một số phương án là) thế hệ con tạo ra từ tự thụ phấn của cây lai F_1 .

Như được sử dụng ở đây, câu "tái tổ hợp" đề cập đến sự trao đổi của các đoạn ADN giữa hai phân tử ADN hoặc thanh nhiễm sắc của nhiễm sắc thể bất cặp ("bất chéo") qua trong vùng của trình tự nucleotit tương tự hoặc tương đồng. "Sự kiện tái tổ hợp" ở đây được hiểu là đề cập theo một số phương án đến bất chéo giảm phân.

Như được dùng trong bản mô tả này, thuật ngữ "trình tự tham chiếu" dùng để chỉ trình tự nucleotit xác định được sử dụng làm cơ sở để so sánh trình tự nucleotit. Theo một số phương án, bất kỳ trong số các SEQ ID NO: 1–4, 22–23, hoặc 73–81 có thể dùng làm trình tự tham chiếu để so sánh với các trình tự khác có được từ cây.

Như được sử dụng ở đây, thuật ngữ "tái sinh," và các biến thể ngữ pháp của nó, đề cập đến sự sinh sản của cây từ nuôi cấy mô.

Như được sử dụng ở đây, câu "các điều kiện lai nghiêm ngặt" đề cập đến các điều kiện mà ở đó polynucleotit lai với dưới trình tự đích của nó, thường trong hỗn hợp phức của axit nucleic, nhưng chủ yếu không có các trình tự khác. Các điều kiện nghiêm ngặt là phụ thuộc trình tự và có thể khác nhau dưới các hoàn cảnh khác nhau.

Trình tự càng dài thường lai đặc hiệu ở nhiệt độ càng cao. Hướng dẫn bao quát để lai axit nucleic thấy ở Sambrook & Russell, 2001. Nhìn chung, các điều kiện nghiêm ngặt được chọn là khoảng 5-10°C thấp hơn điểm nóng chảy nhiệt (T_m) cho trình tự cụ thể ở cường độ ion pH xác định. T_m là nhiệt độ (dưới cường độ ion, pH, và nồng độ axit nucleic xác định) tại đó 50% mẫu dò bổ sung cho đích lai với trình tự đích ở cân bằng (do trình tự đích có mặt dư, ở T_m , 50% mẫu dò chiếm ở cân bằng). Các điều kiện nghiêm ngặt ví dụ là các điều kiện trong đó nồng độ muối ít hơn khoảng nồng độ ion natri 1,0 M, thường natri ion khoảng 0,01 đến 1,0 M (hoặc muối khác) ở độ pH 7,0 đến 8,3 và nhiệt độ ít nhất khoảng 30°C cho mẫu dò ngắn (ví dụ, 10 đến 50 nucleotit) và ít nhất khoảng 60°C cho mẫu dò dài (ví dụ, lớn hơn 50 nucleotit).

Các điều kiện nghiêm ngặt cũng có thể đạt được bằng cách thêm các tác nhân làm mất ổn định như formamit. Các điều kiện lai nghiêm ngặt ví dụ bổ sung bao gồm 50% formamit, 5x SSC, và 1 % SDS ủ ở 42°C; hoặc SSC, 1 % SDS, ủ ở 65°C; với một hoặc nhiều lần rửa trong 0,2x SSC và 0,1% SDS ở 65°C. Để PCR, nhiệt độ bằng khoảng 36°C thường cho khuếch đại nghiêm ngặt thấp, mặc dù nhiệt độ tối có thể thay đổi trong khoảng từ khoảng 32°C đến 48°C (hoặc cao hơn) phụ thuộc vào chiều dài đoạn mồi. Hướng dẫn thêm để xác định các thông số lai được đề xuất trong nhiều tài liệu (xem ví dụ, Ausubel và cộng sự, 1999).

Như được sử dụng ở đây, thuật ngữ "đặc điểm" đề cập đến kiểu hình mong muốn, gen mà góp phần vào kiểu hình mong muốn, cũng như trình tự axit nucleic liên quan đến gen mà góp phần vào kiểu hình mong muốn. Ví dụ, "đặc điểm HI" đề cập đến kiểu hình kích thích tạo đơn bội cũng như gen mà góp phần vào việc kích thích tạo đơn bội và trình tự axit nucleic (ví dụ, sản phẩm gen liên quan đến HI) mà liên quan đến sự có mặt hoặc không có mặt của kích thích tạo đơn bội kiểu hình.

Như được sử dụng ở đây, thuật ngữ "gen chuyển" đề cập đến phân tử axit nucleic được đưa vào sinh vật hoặc một hoặc nhiều tổ tiên của nó bằng một số dạng kỹ thuật chuyển nhân tạo. Do đó kỹ thuật chuyển nhân tạo tạo ra "sinh vật chuyển gen" hoặc "tế bào chuyển gen." Cần hiểu rằng kỹ thuật chuyển nhân tạo có thể xảy ra ở sinh vật tổ tiên (hoặc tế bào trong đó và/hoặc có thể phát triển trong sinh vật tổ tiên) và bất kỳ cá thể thế hệ con khác mà có phân tử axit nucleic được chuyển nhân tạo hoặc đoạn của chúng cũng được coi là chuyển gen ngay cả khi một hoặc nhiều sinh sản tự nhiên và/hoặc được hỗ trợ tạo ra trong phân tử axit nucleic chuyển gen có mặt trong cá thể thế hệ con.

Như được sử dụng ở đây, thuật ngữ "đột biến nhắm đích" hoặc "chiến lược đột biến" đề cập đến bất kỳ phương pháp đột biến mà tạo ra đột biến có chủ ý của gen được chọn. Đột biến nhắm đích bao gồm các phương pháp CRISPR, TILLING, TALEN, và các phương pháp khác chưa được phát hiện nhưng có thể được dùng để thu được cùng kết quả.

Như được sử dụng ở đây, tỷ lệ kích thích tạo đơn bội ("HIR") nghĩa là số lượng hạt đơn bội sống sót trên tổng số hạt sau khi bông được thụ phấn với phấn hoa của dòng kích thích tạo đơn bội.

Vấn đề cụ thể gây khó khăn cho kích thích tạo đơn bội: tỷ lệ thụ phôi tăng và tỷ lệ thụ phấn hỏng tăng (tỷ lệ bộ hạt giảm). Vì những lý do này, vẫn có nhu cầu để xác định thành công nguyên nhân của HI, và sử dụng hiểu biết đó để xác định phương pháp tạo cây đơn bội ổn định hoặc tăng trong khi đồng thời làm giảm thụ phấn hỏng và thụ phôi.

Đặc biệt là có thể gây đột biến gen khởi động để cải thiện tiềm năng hữu ích của các yếu tố để biểu hiện gen chuyển ở cây. Đột biến của các yếu tố này có thể được thực hiện ngẫu nhiên và trình tự gen khởi động được đột biến được sàng lọc hoạt tính trong quy trình thử nghiệm do lỗi. Tùy ý, trình tự cụ thể mà tạo ra gen khởi động với đặc điểm biểu hiện mong muốn, hoặc gen khởi động với hoạt tính tăng cường biểu hiện, có thể được xác định và các trình tự này hoặc tương tự được đưa vào gen khởi động thông qua đột biến. Còn bao gồm là có thể gây đột biến các trình tự này để tăng cường biểu hiện gen chuyển của chúng trong loài cụ thể. Phương pháp để gây đột biến đoạn ADN mã hóa trình tự gen khởi động theo sáng chế là đã biết đối với chuyên gia trong lĩnh vực. Như được đề cập, sự thay đổi cho gen khởi động hoặc yếu tố điều chỉnh khác có thể được thực hiện ngẫu nhiên, hoặc quy trình đột biến đặc hiệu điểm. Gen khởi động và yếu tố điều chỉnh khác có thể được biến đổi bằng cách thay đổi cấu trúc của chúng thông qua việc thêm vào hoặc loại bỏ một hoặc nhiều nucleotit khỏi trình tự mà mã hóa trình tự không biến đổi tương ứng.

Gây đột biến có thể được thực hiện theo bất kỳ trong số các kỹ thuật đã biết trong lĩnh vực, như, và không giới hạn ở, tổng hợp oligonucleotit có một hoặc nhiều đột biến trong trình tự của trình tự điều chỉnh cụ thể. Cụ thể là, đột biến đặc hiệu điểm là kỹ thuật hữu dụng trong việc tạo ra đột biến của gen khởi động, thông qua đột biến đặc hiệu của

ADN nằm dưới. Endonucleaza dẫn ARN ("RGEN," ví dụ, CRISPR/Cas9) cũng có thể được sử dụng. Kỹ thuật còn đề xuất khả năng sẵn sàng để tạo ra và thử nghiệm các biến thể trình tự, ví dụ, kết hợp một hoặc nhiều các nghiên cứu trên, bằng cách tạo ra một hoặc nhiều thay đổi trình tự nucleotit trong ADN. Đột biến đặc hiệu điểm cho phép sản xuất các đột biến thông qua việc sử dụng trình tự oligonucleotit đặc hiệu mà mã hóa trình tự ADN của đột biến mong muốn, cũng như số lượng đủ nucleotit liền nhau, để tạo ra trình tự đoạn mỗi có kích thước đủ và tính phức hợp trình tự để tạo thành phức ổn định trên cả hai bên của mỗi nối loại bỏ được cắt ngang. Thông thường, đoạn mỗi có chiều dài từ khoảng 17 đến khoảng 75 nucleotit hoặc hơn là được ưu tiên, với khoảng 10 đến khoảng 25 hoặc hơn gốc trên cả hai bên của mỗi nối của trình tự được thay đổi.

Khi dòng vô tính bao gồm gen khởi động được phân lập theo sáng chế, người ta có thể mong muốn phân định các vùng gen khởi động thiết yếu trong dòng vô tính. Một phương pháp hiệu quả, nhằm đích để tạo ra gen khởi động đột biến dựa trên sự xác định chi tiết điều chỉnh giả định trong trình tự gen khởi động. Điều này có thể được khởi đầu bằng cách so sánh với trình tự gen khởi động được biết là biểu hiện trong mô tương tự đặc hiệu hoặc mẫu phát triển đơn nhất. Trình tự mà giống nhau giữa các gen khởi động với mẫu biểu hiện giống nhau là các ứng viên có khả năng để liên kết yếu tố phiên mã và do đó là các yếu tố mà thể hiện mẫu biểu hiện. Việc xác nhận các yếu tố điều chỉnh giả định này có thể đạt được bằng cách phân tích loại bỏ trình tự điều chỉnh giả định sau đó phân tích chức năng của mỗi cấu trúc loại bỏ bằng cách thử gen báo cáo mà được gắn chức năng với mỗi cấu trúc. Như vậy, khi tạo ra trình tự gen khởi động bắt đầu, bắt kỳ trong số nhiều đột biến loại bỏ khác nhau của gen khởi động bắt đầu có thể đã được điều chế.

Sáng chế bộc lộ ở đây đề xuất phân tử polynucleotit bao gồm đoạn yếu tố điều chỉnh mà có thể được dùng trong việc tạo cấu trúc các yếu tố điều chỉnh khả năng mới. Các kết hợp mới bao gồm các đoạn của phân tử polynucleotit này và ít nhất một yếu tố hoặc đoạn điều chỉnh khác có thể được tạo cấu trúc và thử ở cây và được xem là thuộc phạm vi sáng chế. Do đó thiết kế, xây dựng, và sử dụng các yếu tố điều chỉnh khả năng là một phương án của sáng chế. Gen khởi động theo sáng chế bao gồm gen tương đồng của yếu tố cis đã biết là ảnh hưởng đến sự điều chỉnh gen mà thể hiện độ tương đồng với trình tự gen khởi động theo sáng chế.

Các đoạn tương đương chức năng của một trong số các axit nucleic điều chỉnh phiên mã được mô tả ở đây bao gồm ít nhất 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, hoặc 1000 cặp bazơ của axit nucleic điều chỉnh phiên mã. Các đoạn tương đương của axit nucleic điều chỉnh phiên mã, mà thu được bằng cách loại bỏ vùng mã hóa vùng không dịch mã 5' của mARN, sau đó chỉ tạo ra vùng gen khởi động (không phiên mã). Vùng không dịch mã 5' có thể được xác định dễ dàng bằng các phương pháp đã biết trong lĩnh vực (như phân tích 5'-RACE). Theo đó, một số trong các axit nucleic điều chỉnh phiên mã, được mô tả ở đây, là các đoạn tương đương của các trình tự khác.

Như nêu trên, các đột biến loại bỏ của gen khởi động theo sáng chế còn có thể được tạo ra một cách ngẫu nhiên và sau đó được thử. Theo chiến lược này, hàng loạt cấu trúc được tạo ra, mỗi cấu trúc chứa phần khác nhau của gen khởi động (dòng phụ), và các cấu trúc này sau đó được sàng lọc về hoạt tính. Các phương tiện thích hợp để sàng lọc hoạt tính là gắn gen khởi động bị loại bỏ hoặc cấu trúc intron mà chứa đoạn loại bỏ vào gen đánh dấu lựa chọn được hoặc sàng lọc được, và phân lập chỉ các tế bào biểu hiện gen đánh dấu. Theo cách này, nhiều cấu trúc gen khởi động khác nhau, bị loại bỏ được xác định mà vẫn giữ hoạt tính mong muốn, hoặc thậm chí được tăng cường. Đoạn nhỏ nhất cần hoạt tính theo đó được nhận diện thông qua so sánh cấu trúc được lựa chọn. Đoạn này sau đó có thể được dùng để tạo cấu trúc các vector để biểu hiện gen ngoại sinh.

Cat-xet biểu hiện như được mô tả ở đây có thể bao gồm các yếu tố điều chỉnh khác. Thuật ngữ trong bối cảnh này được hiểu theo nghĩa rộng bao gồm tất cả các trình tự mà có thể ảnh hưởng đến việc tạo cấu trúc hoặc chức năng của cat-xet biểu hiện. Các yếu tố điều chỉnh có thể, ví dụ, thay đổi phiên mã và/hoặc dịch mã trong sinh vật nhân sơ hoặc nhân thực. Cat-xet biểu hiện được mô tả ở đây có thể là xuôi dòng (theo hướng 3') của trình tự axit nucleic được biểu hiện và tùy ý chứa các yếu tố điều chỉnh bổ sung, như gen tăng cường phiên mã hoặc dịch mã. Mỗi yếu tố điều chỉnh bổ sung có thể liên kết hoạt động với trình tự axit nucleic được biểu hiện (hoặc trình tự nucleotit điều chỉnh phiên mã). Các yếu tố điều chỉnh bổ sung có thể bao gồm gen khởi động bổ sung, gen khởi động nhỏ, yếu tố gen khởi động, hoặc yếu tố gen nhảy mà có thể thay đổi hoặc tăng cường đặc điểm điều chỉnh biểu hiện. Cat-xet biểu hiện còn có thể chứa một hoặc nhiều intron, một hoặc nhiều exon và một hoặc nhiều gen kết thúc.

Hơn nữa, dự tính rằng gen khởi động kết hợp các yếu tố từ nhiều hơn một gen khởi động có thể là hữu dụng. Ví dụ, Sáng chế Mỹ số 5,491,288 bộc lộ kết hợp gen khởi động virut khảm súp lơ với gen khởi động histon. Do đó, các yếu tố từ gen khởi động được bộc lộ ở đây có thể được kết hợp với các yếu tố từ gen khởi động khác. Gen khởi động mà hữu dụng để biểu hiện gen chuyển ở cây bao gồm các gen mà kích thích tạo được, của virut, tổng hợp, cơ bản (Odell Nature 313: 810–812 (1985)), được điều chỉnh tạm thời, được điều chỉnh về không gian, đặc hiệu mô, và được điều chỉnh tạm thời về không gian. Sử dụng các yếu tố điều chỉnh được mô tả ở đây, nhiều gen nông học có thể được biểu hiện trong cây biến nạp. Cụ thể hơn là, cây có thể được thiết kế di truyền để biểu hiện nhiều kiểu hình khác nhau mong muốn về nông học.

Hợp chất theo sáng chế có thể tồn tại ở các dạng đồng phân hình học hoặc đồng phân quang học khác nhau (chất đồng phân không đối quang và chất đồng phân đối ảnh) hoặc các dạng hỗn biến. Sáng chế bao gồm tất cả các chất đồng phân và chất hỗn biến này và hỗn hợp của chúng ở tất cả các tỷ lệ cũng như các dạng đồng vị như các hợp chất deuteri hóa. Sáng chế cũng bao gồm tất cả các muối, N-oxit, và phức hợp á kim của hợp chất theo sáng chế.

Mỗi gốc alkyl hoặc là một mình hoặc là một phần của nhóm lớn hơn (như alkoxy, alkoxycarbonyl, alkylcarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl) là mạch thẳng hoặc mạch nhánh và là, ví dụ, metyl, etyl, n-propyl, n-butyl, n-pentyl, n-hexyl, iso-propyl, n-butyl, sec-butyl, iso-butyl, tert-butyl hoặc neo-pentyl. Các nhóm alkyl bao gồm C1–C6 alkyl, C1–C4 alkyl, và C1–C3 alkyl.

Thuật ngữ "alkenyl," như được sử dụng ở đây, là gốc alkyl có ít nhất một liên kết đôi cacbon-cacbon, ví dụ C2–C6 alkenyl. Các ví dụ cụ thể bao gồm vinyl và allyl. Gốc alkenyl có thể là một phần của nhóm lớn hơn (như alkenoxy, alkenoxycarbonyl, alkenylcarbonyl, alkenylaminocarbonyl, dialkenylaminocarbonyl).

Thuật ngữ "axetoxy" đề cập đến $-OC(=O)CH_3$.

Thuật ngữ "alkynyl," như được sử dụng ở đây, là gốc alkyl có ít nhất một liên kết ba cacbon-cacbon, ví dụ C2–C6 alkynyl. Các ví dụ cụ thể bao gồm etynyl và propargyl. Gốc alkynyl có thể là một phần của nhóm lớn hơn (như alkynoxy, alkynoxycarbonyl, alkynylcarbonyl, alkynylaminocarbonyl, dialkynylaminocarbonyl).

Halogen là flo (F), clo (Cl), brom (Br) hoặc iot (I).

Các nhóm haloalkyl (hoặc là một mình hoặc là một phần của nhóm lớn hơn, như haloalkoxy hoặc haloalkylthio) là nhóm alkyl mà được thế bằng một hoặc nhiều trong số các nguyên tử halogen giống hoặc khác nhau và là, ví dụ, $-\text{CF}_3$, $-\text{CF}_2\text{Cl}$, $-\text{CH}_2\text{CF}_3$, hoặc $-\text{CH}_2\text{CHF}_2$.

Các nhóm hydroxyalkyl là nhóm alkyl mà được thế bằng một hoặc nhiều nhóm hydroxyl và là, ví dụ, $-\text{CH}_2\text{OH}$, $-\text{CH}_2\text{CH}_2\text{OH}$ hoặc $-\text{CH}(\text{OH})\text{CH}_3$.

Các nhóm alkoxyalkyl là nhóm alkoxy được liên kết với alkyl ($\text{R}-\text{O}-\text{R}'$), ví dụ $-(\text{CH}_2)_r\text{O}(\text{CH}_2)_s\text{CH}_3$, trong đó r là 1 đến 6 và s là 1 đến 5.

Trong bối cảnh của sáng chế, thuật ngữ "aryl" đề cập đến hệ vòng mà có thể là một, hai hoặc ba vòng. Ví dụ về các vòng này bao gồm phenyl, naphthalenyl, anthracenyl, indenyl hoặc phenantrenyl.

Trừ khi có quy định khác, alkenyl và alkynyl, trên chúng hoặc là một phần của phần tử thế khác, có thể là mạch thẳng hoặc mạch nhánh và có thể chứa từ 2 đến 6 nguyên tử cacbon, và khi thích hợp, có thể là có cấu hình (E) hoặc (Z). Các ví dụ bao gồm vinyl, allyl, etynyl và propargyl.

Trừ khi có quy định khác, cycloalkyl có thể là vòng đơn hoặc vòng đôi, có thể là tùy ý được thế bởi một hoặc nhiều nhóm $\text{C}_1\text{--C}_6$ alkyl, và chứa từ 3 đến 7 nguyên tử cacbon. Các ví dụ của cycloalkyl bao gồm cyclopropyl, 1-methylcyclopropyl, 2-methylcyclopropyl, cyclobutyl, cyclopentyl, và cyclohexyl.

Thuật ngữ "heterocyclyl" đề cập đến hệ vòng chứa từ một đến bốn dị nguyên tử được chọn từ N, O và S, trong đó nguyên tử nitơ và lưu huỳnh tùy ý được oxy hóa, và (các) nguyên tử nitơ tùy ý được tạo bậc bốn. Heterocyclyl bao gồm heteroaryl, chất tương tự bão hòa, và ngoài ra chất tương tự không bão hòa hoặc bão hòa một phần của chúng như 4,5,6,7-tetrahydro-benzothiophenyl, 9H-fluorenyl, 3,4-dihydro-2H-benzo-1,4-dioxepinyl, 2,3-dihydro-benzofuranyl, piperidinyl, 1,3-dioxolanyl, 1,3-dioxanyl, 4,5-dihydro-isoxazolyl, tetrahydrofuranlyl và morpholinyl. Ngoài ra, thuật ngữ "heterocyclyl" bao gồm heterocycloalkyl, vòng đơn vòng hoặc đa vòng không thơm có chứa nguyên tử cacbon và hydro và ít nhất một nguyên tử khác loại được chọn từ nitơ, oxy, và lưu huỳnh như oxetanyl hoặc thietanyl. Heterocycloalkyl đơn vòng có thể chứa từ 3 đến 7 cạnh.

Thuật ngữ "heteroaryl" đề cập đến hệ vòng thơm chứa từ một đến bốn dị nguyên tử được chọn từ N, O, và S, trong đó nguyên tử nitơ và lưu huỳnh tùy ý được oxy hóa, ví dụ có 5, 6, 9, hoặc 10 cạnh, và chứa hoặc là vòng đơn hoặc hai hoặc nhiều vòng dung hợp. Vòng đơn có thể chứa đến ba dị nguyên tử, và hệ hai vòng đến bốn dị nguyên tử, mà sẽ tốt hơn là được chọn từ nitơ, oxy, và lưu huỳnh. Các ví dụ của các nhóm này bao gồm pyridyl, pyridazinyl, pyrimidinyl, pyrazinyl, furanyl, thienyl, oxazolyl, isoxazolyl, oxadiazolyl, thiazolyl, isothiazolyl, thiadiazolyl, pyrrolyl, pyrazolyl, imidazolyl, triazolyl, và tetrazolyl.

Ví dụ thực hiện sáng chế

I. Xác định đột biến dịch khung ở PLA

Sáng chế xác định hàng loạt đột biến do con người độc lập thấy ở ít nhất một gen phospholipaza giống patatin AII α (pPLAII α) của ngô; cây ngô có các đột biến này ở ít nhất một trong các gen PLA của chúng; và phương pháp tạo ra và xác định đột biến tương tự và/hoặc bổ sung ở gen PLA bằng cách sàng lọc các cây lúa và ngô kết hợp và/hoặc từng cá thể. Cây lúa và ngô theo sáng chế kích thích tạo tính đơn bội do đột biến không chuyển gen ở ít nhất một trong các gen PLA của chúng.

Đặc biệt hơn nữa là, sáng chế tạo ra dòng kích thích tạo đơn bội mới của ngô. Nhiều dòng ngô kích thích tạo đơn bội đã biết tồn tại bao gồm nhưng không giới hạn ở: Stock 6, MHI (Dòng kích thích tạo đơn bội Moldovian), đột biến thể giao tử vô định ("ig"), KEMS, ZEM, ZMS, KMS, RWS, và RWK. Sáng chế đề cập đến phương pháp xác định, và/hoặc lựa chọn vật liệu di truyền mà có thể hoặc không thể kích thích tạo các thể đơn bội. Sáng chế còn đề cập đến việc làm tăng và phát triển khác vật liệu di truyền kích thích tạo thể đơn bội được chọn. Sáng chế còn đề cập đến phương pháp cải thiện vật liệu di truyền kích thích tạo đơn bội để làm tăng việc kích thích tạo thể đơn bội trên hạt tạo ra bố mẹ.

Bước khởi đầu trong việc sản xuất hạt đơn bội từ thể lai hoặc tách cây bố mẹ có từ quá trình thụ phấn với phần hoa từ dòng kích thích tạo đơn bội lên bông từ hạt tạo ra cây. Kết quả của quá trình lai này là tạo ra hạt lưỡng bội và đơn bội của mẹ (1n). Hạt đơn bội được kích thích tạo (1n) thường được phân biệt với hạt lưỡng bội bằng cách sử dụng gen đánh dấu màu sắc mà chỉ ra mức bội thể của phôi. Hạt lưỡng bội thường bị loại, trong khi hạt hoặc phôi đơn bội thường được đưa vào quá trình nhân đôi nhiễm sắc

thể để tạo ra cây đơn bội kép. Đặc biệt hơn nữa là, vật liệu di truyền đơn bội được xử lý với một hoặc nhiều tác nhân chặn giãn phân để cho phép nhiễm sắc thể đơn bội (1n) bổ sung ở một hoặc nhiều tế bào để tạo ra các cặp đồng đẳng. Sau quá trình xử lý hóa học, (các) chất hóa học nhân đôi nhiễm sắc thể được loại ra. Ngô đơn bội kép này được để chín và hạt đơn bội kép thu được khi được trồng sẽ tạo ra cây đồng nhất (còn gọi là cây hoặc dòng lai nội dòng). Dòng lai nội dòng là các vật liệu mà người lai giống sử dụng để theo đuổi chương trình phát triển giống lai của họ.

Vị trí cho đặc điểm kích thích tạo đơn bội được tạo bản đồ tỉ mỉ. Mặc dù QTL chính trên nhiễm sắc thể 1 chịu trách nhiệm đối với kích thích tạo đơn bội đã được tạo bản đồ và được công bố, Dong và cộng sự Theor. Appl. Genet (2013) 126: 1713-1720, gen/yếu tố di truyền chính xác chịu trách nhiệm đối với quy trình kích thích tạo vẫn chưa được xác định cho đến nay. Để làm rõ di truyền phát triển dưới sự kích thích tạo đơn bội, dẫn xuất Stock 6 RWK (~13% HIR) lấy từ Đại học Hohenheim năm 2006, lai với vật lai nội dòng NP2460 và NP2391, và sau đó lai ngược với RWK để tạo ra quần thể tạo bản đồ. Xem Hình 1.

HIR tăng ở cả quần thể cùng được tách với chất đánh dấu SM020SDQ trong bin 1.04, thống nhất với báo cáo gần đây về QTL gọi là *qhir1*. Xem Prigge, và cộng sự, *New Insights into the Genetics of in Vivo Induction of Maternal Haploids, the Backbone of Doubled Haploid Technology in Maize*, GENETICS (2012) 190:781–793 (bàn về QTL chính đối với HI trong Bin 1.04 [*qhir1*] và QTLs thứ yếu đối với HI trong Bins 3.02 [*qhir2*], 3.06 [*qhir3*], 4.03 [*qhir4*], 5.01 [*qhir5*], 5.04 [*qhir6*], 7.01 [*qhir7*], và 9.01 [*qhir8*]); Liu, và cộng sự, *Fine mapping of qhir8 affecting in vivo haploid induction in maize*, THEOR. APPL. GENET. (2015) 128:2507–2515 (tạo bản đồ tỉ mỉ ba mươi lăm gen cho *qhir8*); Hu, và cộng sự, *The Genetic Basis of Haploid Induction in Maize Identified with a Novel Genome-Wide Association Method*, GENETICS (2016) 202:1267–1276 (xác định rằng *qhir1* là hai QTLs: *qhir11* và *qhir12*, và tạo bản đồ tỉ mỉ *qhir6* cho vùng 1.1Mb). Tác giả thực hiện một vài vòng tạo bản đồ tỉ mỉ và thu hẹp QTL đến vùng xấp xỉ 0,57 Mb nằm trong khoảng từ 67,85 Mb đến 68,42 Mb nằm trong *qhir11*. Vùng này có bảy gen chú giải (Hình 2).

Sử dụng Illumina HiSeq2000, tác giả tạo trình tự RWK, Stock 6, và BC3F5 không kích thích tạo "RWK-NIL" mà gần đồng hợp tử với RWK nhưng có kiểu gen đơn bội NP2391 trong khoảng *qhir11*. Bằng cách so sánh vật liệu di truyền dòng kích thích

tạo và dòng không kích thích tạo, đã xác định được rằng bốn chỗ chèn nucleotit có trong dòng kích thích tạo đơn bội mà dịch khung cho mã hóa axit amin của GRMZM2G471240 không có trong vật liệu di truyền dòng không kích thích tạo. Do đó, sáng chế đã xác định gen với đột biến dịch khung trong vật liệu di truyền kích thích tạo chịu trách nhiệm đối với kích thích tạo đơn bội ở ngô. Gen ứng viên tương ứng với mô hình gen GRMZM2G471240 mã hóa phospholipaza giống patatin AII α (pPLAII α), mà tác giả gọi là *MATRILINEAL* (*MTL*) để thể hiện alen kiểu dại và alen dịch khung được gọi là *matrilineal* (*mtl*).

Trình tự ADN được tạo ra cho gen ứng viên từ hai dòng kích thích tạo (Stock6 và RWK) và một dòng không kích thích tạo (RWK-NIL). Ngoài ra, dữ liệu bộ gen B73 công bố dùng làm dòng không kích thích tạo thứ hai. Thông tin mô hình gen được so sánh với dữ liệu EST/cADN để xác nhận cấu trúc của mỗi gen. Dữ liệu trình tự chú giải được so sánh với sự khác nhau danh mục giữa bốn alen của mỗi gen.

So sánh trình tự thể hiện rằng alen B73 và RWK-NIL tương tự với nhau, và alen RWK và Stock 6 tương tự với nhau. Sự khác biệt trình tự nhất là đa hình nucleotit đơn mafkhuong thay đổi trình tự mã hóa protein. Có một số chèn và một số loại bỏ, hầu hết trong số đó là trình tự mã hóa không phải protein.

Bảng 1 dưới đây liệt kê nhận diện gen trong khoảng được thể hiện. Thông tin này là từ nhiễm sắc thể 1, xem, ví dụ, Patrick S. Schnable và cộng sự, *The B73 Maize Genome: Complexity, Diversity, and Dynamics*, 326 SCIENCE 1112–15 (2009), được kết hợp vào đây bằng cách viện dẫn toàn bộ, và liệt kê mô tả ngắn gọn các protein được mã hóa khác từ gen trong vị trí kích thích tạo đơn bội.

Bảng 1. Thông tin về Nhiễm sắc thể 1 từ Ngô.

Bản mã									
Gen_id	Bắt đầu	Kết thúc	Chiều dài truy xuất	Chiều dài đối tượng	Điểm	Tương đồng	Tương tự	Chiều dài thẳng hàng	Mô tả
GRMZM2G305400	67991172	67994092	308	362	385	33,3	53,33752	314	Xyclin D2;1
GRMZM2G082836	68107606	68110989	202	205	729	71,2	83,33333	198	Protein 1 liên kết GTP
GRMZM2G382717	68113455	68115168	396	464	489	38,77	53,17371	314	Siêu họ protein vùng Chaperone DnaJ
GRMZM2G120587	68133178	68136953	458	461	1329	55	71,23894	452	serin kiểu carboxypeptidaza 51
GRMZM2G471240	68240862	68242656	428	407	1049	51,5	72,36181	398	phospholipaza A 2A
GRMZM2G471240	68240862	68242656	401	407	961	50,15	70,0938	395	phospholipaza A 2A
GRMZM2G062320	68318898	68321409	335	334	1064	73,3	84,21053	285	Họ protein phosphoglyxerat mutaza
GRMZM5G866758	68430654	68436197	401	403	1678	80,4	90,45226	398	axetoaxetyl-CoA thiolaza 2
GRMZM5G866758	68430654	68436197	303	403	1248	78,4	89,40397	302	axetoaxetyl-CoA thiolaza 2
GRMZM2G003530	68435670	68439997	360	344	1063	60,5	76,41791	335	Vòng P chứa nucleosit triphosphat siêu họ protein hydrolaza
GRMZM2G077991	68543246	68546264	94	95	424	79,7	91,48936	94	Họ protein protein ribosomal liên kết kẽm
GRMZM2G077991	68543694	68546264	94	95	424	79,7	91,48936	94	Họ protein protein ribosomal liên kết kẽm
GRMZM2G077991	68543805	68546269	147	95	419	79,5	91,39785	93	Họ protein protein ribosomal liên kết kẽm
GRMZM2G077960	68554980	68559182	438	428	1422	65,3	79,80998	421	Họ protein 2C protein phosphataza
GRMZM2G077897	68561209	68565155	784	807	1561	48,1	65,69848	723	Protein của cây có chức năng chưa biết (DUF827)
GRMZM2G347583	68660278	68665995	1651	2156	1201	41,37	55,70954	1375	

GRMZM2G173030	68668900	68671460	626	2156	858	35,6	48,30299	586	
GRMZM2G022061	68876150	68882226	203	556	618	64,9	79,89691	194	
GRMZM2G022061	68876150	68882226	322	556	1004	66	77,47748	333	
GRMZM2G022061	68876150	68882226	142	556	547	79,6	89,84375	128	
GRMZM2G022061	68876150	68882226	322	556	1004	66	77,47748	333	
GRMZM2G022061	68876150	68882226	534	556	1802	67,7	79,81651	545	
GRMZM2G340286	68928213	68929600	378	403	570	37,83	55,75713	407	
GRMZM2G340279	68934652	68937080	746	937	3095	29,34	50,31745	2517	Siêu họ protein protein tetratricopeptit lặp (TPR) giống
GRMZM2G347808	69005208	69012612	589	455	1115	50,4	66,60178	423	Siêu họ protein metyltransferaza phụ thuộc S-adenosyl-L-methionin

Sau khi hoàn thành bản đồ tỉ mỉ về đặc điểm của dòng kích thích tạo đơn bội cho các khoảng cách chứa chỉ bảy gen, tác giả tập trung vào việc lắp ráp trình tự và phân tích. Các trình tự cho bảy gen gần tương đồng giữa B73 và RWK-NIL, nhwg RWK và Stock 6 không có GRMZM2G062320, PHOSPHOGLYXERAT MUTASE (PGM), và có sự chèn 4 cặp bazơ ("bp") trong exon thứ tư của GRMZM2G471240, PHOSPOLIPAZA GIỐNG PATATIN AII α (pPLAII α) (Hình 2). Tác giả phát hiện rằng RWK và Stock6 đều có sự chèn 4 bp giống nhau trong exon thứ tư của pPLAII α , và gen này được biểu hiện đặc hiệu trong phần hoa (xem maizegdb.org/gene_center/gene?id=GRMZM2G471240 được kết hợp vào đây bằng cách viện dẫn). GRMZM2G471240 không đột biến được thể hiện bởi SEQ ID NO: 1. GRMZM2G471240, bao gồm việc chèn 4 bp trong exon thứ tư, được thể hiện bởi SEQ ID NO: 3.

Hầu hết các thể đơn bội được xác định được tìm thấy sử dụng thử nghiệm đánh dấu taqman. Thử nghiệm đánh dấu này lợi dụng sự khác nhau ở gen pPLAII α giữa RWK x NP2222. Trong các quá trình lai mà tác giả sử dụng RWK làm giống cái, và NP2222 làm giống đực, RWK bố mẹ đồng nhất đối với alen *mtl*, trong khi NP2222 đồng nhất đối với alen *MTL*. Thể hệ con lưỡng bội là *MTL/mtl* và thể hệ con đơn bội phát triển do thể cái đơn bội là *mtl/0*. Do đó khi thử nghiệm này được thực hiện các kết quả taqman thể hiện 1 bản sao của alen *mtl* và một bản sao của alen *MTL* ở thể hệ con lưỡng bội, và 1 bản sao của *mtl* nhưng không phải bản sao của *MTL* ở thể hệ con đơn bội. Khi kiểu lai này được thực hiện, các bông được thu hoạch trong khoảng 12–21 ngày sau khi thụ phấn, phôi được chiết và mẫu phôi nhỏ ược lấy để phân tích gen đánh dấu taqman. Tùy ý phôi được trải lên môi trường rắn và nảy mầm trong tối sao cho mẫu lớn hơn của chồi hoặc rễ mọc ra có thể lấy trong khoảng 2–10 ngày sau để phân tích gen đánh dấu. Đồng thời, một số mô được cắt để phân tích mức bội thể. Trong trường hợp sau sau khi thử nghiệm phân tử được sử dụng, các mẫu lớn hơn của các thể đơn bội có thể chạy trên thiết bị phân tích mức bội thể không gian CyFlow và được xác nhận là các thể đơn bội. Trong hầu hết các trường hợp điều này tạo ra việc xác định dương tính của các thể đơn bội. Trong một vài trường hợp hiếm hơn điều này tạo ra sự lật đổ các kết quả gen đánh dấu dương tính sai và hiệu chỉnh cái gọi là lưỡng bội.

Theo cách khác tác giả thử nghiệm các thể đơn bội là qua thử nghiệm gen đánh dấu nổi trội. Trong trường hợp này, dòng đực X26 được sử dụng. Dòng này đồng nhất

cho gen đánh dấu mà đóng vai trò theo cách trội. Trong việc lai này dòng bất kỳ có thể được dùng làm dòng cái cho đến khi nó không có gen đánh dấu hoặc bất kỳ gen hoặc alen mà làm việc để ức chế kiểu hình của gen đánh dấu. Dòng X26 là dòng không kích thích tạo và đồng nhất đối với MTL. Sử dụng dòng này, thế hệ con được cắt ra trong khoảng 12 – 21 ngày sau khi thụ phấn và đánh giá sự có mặt của gen đánh dấu, hoặc chúng được kiểm tra trực tiếp trên bông, hoặc hạt được làm khô được thu hoạch và đánh giá sự có mặt của gen đánh dấu. Thế hệ con lưỡng bội thể hiện kiểu hình gen đánh dấu vì chúng có bản sao đơn của gen đánh dấu từ bố mẹ X26 đực, trong khi đó thế hệ con đơn bội phát triển từ mẹ không thể hiện kiểu hình của gen đánh dấu. Sự xâm nhập của gen đánh dấu và tỷ lệ kích thích tạo đơn bội tự phát của X26 được thử trong nhiều lần lai đối chứng. Sử dụng hệ này tác giả sàng lọc các thể đơn bội và sau đó thử nghiệm chúng trên thiết bị phân tích mức bôi thể để xác định rằng chúng là các thể đơn bội thật sự.

Tác giả đã phát triển thử nghiệm PCR để phát hiện đặc hiệu alen "kiểu đại" và "đột biến" để sàng lọc mười chín dòng kích thích tạo dẫn xuất từ Stock 6, bao gồm dòng kích thích tạo đơn bội NP2222 (NP2222-HI), đưa gen khác vào BC3 của RWK trong dòng lai nội dòng biến nạp chuẩn Syngenta NP2222. Tác giả cũng sàng lọc chín dòng đối chứng không kích thích tạo.

Phát triển thử nghiệm PCR có thể phân biệt giữa kiểu gen đơn bội RWK/Stock6 và RWK-NIL, hai cặp đoạn mồi được thiết kế: một cặp có thể khuếch đại alen dịch khung RWK/Stock6, trong khi cặp còn lại khuếch đại alen B73/RWK-NIL. Các cặp này làm việc như kỳ vọng trên RWK-NIL, RWK, và Stock6 ADN: RWK-NIL gDNA chỉ khuếch đại cặp đoạn mồi RWK-NIL. RWK và Stock6 gDNA chỉ khuếch đại cặp đoạn mồi RWK/Stock6, mà phát hiện đặc hiệu alen dịch khung. Sản phẩm PCR được tạo trình tự và trình tự tương đồng với trình tự từ việc tạo trình tự tổng thể bộ gen. SNP mà được xác định trong việc tạo trình tự tổng thể bộ gen được xác nhận trong sản phẩm PCR. Dưới đây, trên Hình 2, ADN dùng trong mỗi phản ứng là chữ in hoa. Đoạn mồi là "nil.F1/R1" và "rwk.F1/R1."

GRMZM2G471240_nil.F1: GTACGCCGTGCGCTAACA (SEQ ID NO: 5).
 GRMZM2G471240_nil.R1: TCGTACCTCCCTGTCTCCAC (SEQ ID NO: 6).
 GRMZM2G471240_rwk.F1: TACGCCGTGCGCTAACATA (SEQ ID NO: 7).
 GRMZM2G471240_rwk.R1: GTACCTCGCTCCCTGTCTCC (SEQ ID NO: 8).

Các cặp đoạn mồi "rwk.F1/R1" và "nil.F1/R1" được dùng cho kiểu gen panen của dòng kích thích tạo cao, thấp, và không kích thích tạo. Tác giả thấy rằng tất cả các 19 dòng kích thích tạo đơn bội có chèn 4 bp, bao gồm Stock6 (3% tỷ lệ kích thích tạo đơn bội ["HIR"]), RWK (dòng có nguồn gốc từ Đại học Honheim stocks, 10–15% HIR), RWS, và Z22, trong số khác. Ngược lại, alen kiểu dại được thấy ở tất cả chín dòng không kích thích tạo đơn bội (HIR trung bình bằng 0,1%). Dữ liệu chỉ ra rằng tính đồng nhất đối với alen dịch khung tương quan với khả năng kích thích tạo: Dòng kích thích tạo cao 12/12 và thấp 7/7 khuếch đại thử nghiệm dịch khung, nhưng không phải thử nghiệm kiểu dại, trong khi dòng không kích thích tạo 9/9 khuếch đại thử nghiệm kiểu dại nhưng không phải dịch khung. Điều này chỉ ra rằng khả năng kích thích tạo tương quan với đột biến GRMZM2G471240, và rằng *pPLAIIα* dưới *qhir11* và đột biến cơ bản chịu trách nhiệm đối với kích thích tạo đơn bội trong các dòng này.

Bảng 2. Các kết quả thử nghiệm PCR của GRMZM2G471240.

RWK			
Tỷ lệ kích thích tạo	Đơn vị siêu sao chép	Đơn vị siêu sao chép RWK-NIL	
<u>Đối chứng:</u>			
Stock 6 (dòng kích thích tạo thấp)	2,50%	+	-
RWK (dòng kích thích tạo cao)	12%	+	-
RWK-NIL (dòng không kích thích tạo)	<1%	-	+
<u>Các dòng kích thích tạo tốt:</u>			
ZMS	7%	+	-
Z19-PR	7%	+	-
RWS-Z86	10%	+	-
K13	9%	+	-
(ID3002/Z22)B>29-5>2-5-1-B-	7%	+	-
Z-19-//AF4031PR//Z-19-)1-1-2-3-1-3-B-	9,50%	+	-

ZR86	12%	+	-
ZR53	12%	+	-
ZR75	13%	+	-
(Z21/RWS)B(GS)-75-1-2-3-B-	~8%	+	-
Dòng kích thích tạo tốt NP2222	~9%	+	-

Các dòng kích thích tạo kém:

Stock6 R1-nj	2,50%	+	-
(Z21/RWS/[RWS]B\$)33-5-	<2%	+	-
(K-13-/(ZMS/SEW-PR)B>2>B-9//K-13-)2-4-1-	<2%	+	-
(K-13-/(ZMS/SEW-PR)B>2>B-9//K-13-)6-1-2-	<2%	+	-
(ZMS/SEW-PR)B>2>B-7-2-1-2-	<2%	+	-
NP2222 dòng kích thích tạo thấp	~3%	+	-

Các dòng không kích thích tạo và vật cho:

Stock6 R1-nj B1P1I	< 0,1%	-	+
(Z-21-/AF4031PR//Z-21-1-B-)1-1-1-1-B-	< 0,1%	-	+

FF6096	<0,1%	-	+
ID5829	<0,1%	-	+
XO5744	<0,1%	-	+
ID3002	<0,1%	-	+
AF4031PR	<0,1%	-	+
NP2222	<0,1%	-	+

Tác giả cũng xác định nhiều đa hình nucleotit đơn ("SNP") giữa alen dịch khung và của RWK-NIL. Đối với nhiều trong các SNP này, trình tự STOCK6 và RWK phù hợp với thể lai cùng dòng khác mà tác giả đã tạo trình tự, và do đó có khả năng thể hiện biến thể tự nhiên. Thật vậy, hầu hết các SNP này không thay đổi trình tự axit amin và do đó có khả năng không góp phần vào kiểu hình kích thích tạo đơn bội. Hai SNP không tạo ra thay đổi axit amin (H107Y; K232N) và không có thay đổi bảo toàn cao, do đó chúng có thể có đóng góp nhỏ vào so kiểu hình, nhưng gần như chúng không tác động đến kiểu hình vì dịch khung gây ra mất chức năng.

Tác giả đặt tên *pPLAIIα* "MATRILINEAL" (*MTL*; tức là, SEQ ID NO: 1) và alen chèn 4 bp bẩm sinh "*matrilineal*" (*mtl*; tức là, SEQ ID NO: 3). Theo trình tự protein định trước, chèn 4 bp gây ra dịch khung đọc mở của protein ở axit amin ("AA") 352 ngoài 401. Dịch khung đọc mở dẫn đến bộ ba mã kết thúc sớm.

Sau khi tìm ra đột biến bất hoạt dịch khung mà tác giả đã thử nghiệm trực tiếp hiệu quả nó có lên kích thích tạo đơn bội bằng cách bổ sung dòng kích thích tạo đơn bội với gen chuyển *pPLAIIα* kiểu dại. Bổ sung khác loại của NP2222-HI (10,2% HIR) với bản sao kiểu dại của *MATRILINEAL* loại bỏ hầu hết kích thích tạo đơn bội và thui hạt. So với đối chứng HIR làm giảm 50-lần, từ 10% đến 0,23%. Nó còn làm giảm tỷ lệ thui phôi xuống 0,65%. Dòng báo cáo chức năng chiều dài đầy đủ cũng được thực hiện sử dụng dung hợp chuyển gen của gen *MTL* kiểu dại vào GFP cũng như alen đột biến *mtl* vào GFP, để cả hai hình dung sự định vị dưới mức tế bào của *MTL* kiểu dại, nhưng cũng để thấy nếu phiên bản đột biến của protein định vị chính xác hoặc được tạo ra tất cả. Các dòng này cũng dùng làm vật liệu bổ sung để thử nghiệm sự bổ sung. Vật liệu dòng kích thích tạo đơn bội (NP2222-HI) mà đồng nhất với gen chuyển *MTL-GFP* cũng không thể hiện kiểu hình dòng kích thích tạo đơn bội. Tỷ lệ kích thích tạo của NP2222-HI rơi vào 0,60% khi nó đồng nhất cho *MTL-GFP*. Hơn nữa, gen chuyển *MTL-GFP* còn làm bất hoạt thui phôi xuống 4,86%. Cuối cùng tác giả thử nghiệm xem alen đột biến *mtl* được dung hợp vào GFP có bổ sung cho kiểu hình kích thích tạo đơn bội hay không, và nó không. Tỷ lệ kích thích tạo đơn bội và thui phôi rất tương tự ở NP2222-HI so với NP2222-HI mà đồng nhất đối với gen chuyển dung hợp đột biến *mtl-GFP*. Xem Bảng 3. Điều này thể hiện bằng chứng kết luận rằng dịch khung MATRILINEAL chịu trách nhiệm và cần cho kích thích tạo đơn bội. Để áp dụng hiểu biết này, tác giả chứng minh

rằng việc gây đột biến hoặc điều biến biểu hiện của pPLAII α trong dòng kiểu đại dẫn đến tạo ra dòng kích thích tạo đơn bội mới.

Bảng 3. Đặc điểm sinh sản của dòng kích thích tạo đơn bội, bổ sung và dòng chỉnh sửa. Bảng này thể hiện tỷ lệ kích thích tạo đơn bội và thui hạt của dòng kích thích tạo trong nền NP2222. Số lai thử nghiệm được liệt kê đầu tiên ("bông") và sau đó liệt kê thống kê hạt và phôi. Thui phôi và kích thích tạo đơn bội thường đi cùng nhau trên cùng bông. Tức là tại sao tỷ lệ thui phôi cao trong bông được lai bởi NP2222-HI đực, mà có 10,17% tỷ lệ kích thích tạo đơn bội (HIR). Cả gen chuyển kiểu đại, bao gồm một gen không có GFP và một gen được dung hợp vào GFP, bổ sung cho kiểu hình kích thích tạo đơn bội. Trong khi đó, đột biến *mtl* được dung hợp vào GFP không bổ sung.

Thử nghiệm bổ sung			Đặc điểm hạt			Phôi được thử		
Bố mẹ đực	bông	có thể tồn tại được	bị thui	% bị thui	phôi	các thể đơn bội	các lưỡng bội	HIR
NP2222-HI	4	548	498	47,61%	531	54	477	10,17%
NP2222-HI + <i>MTL</i> / <i>MTL</i>	17	4403	29	0,65%	4321	11	4310	0,25%
NP2222-HI + <i>mtl</i> -GFP/ <i>mtl</i> -GFP	3	371	298	44,54%	360	34	326	9,44%
NP2222-HI + <i>MTL</i> -GFP/ <i>MTL</i> -GFP	3	1019	52	4,86%	836	5	831	0,60%

Một vài alen kiểu *mtl* được tạo ra trong NP2222 nội dòng bằng cách tạo ra loại bỏ nhỏ trong *MTL* gần vị trí chèn 4 bp trong *mtl*, sử dụng nucleaza thực hiện kiểu hoạt hóa phiên mã (TALEN) (Boch, J. và cộng sự, *Breaking the code of DNA binding specificity of TAL-type III effectors*, SCIENCE 326:1509–1512 (2009), được kết hợp vào đây bằng cách việ dẫn). Một vài sự kiện đột biến là cây tự thụ phấn và T1 không có TALEN T-ADN được chèn nhưng đồng nhất để loại bỏ nhỏ trong *MTL* được lai xa lên NP2222. Các dòng chỉnh sửa đồng nhất để loại bỏ dịch khung trong *MTL* (sau đây gọi là *MTL*^{TAL-FS}) thể hiện HIR bằng 4,0 – 12,5% (trung bình 6,65%) (Bảng 4). Tình trạng mức bội thể của các thể đơn bội giả định 118/127 được xác nhận bởi kỹ thuật phân tích tế bào theo dòng chảy, và đánh giá kiểu hình chỉ ra rằng các cây này là các thể đơn bội. Kết quả này chứng minh rằng dịch khung trong *MTL* đủ để kích thích tạo tốc độ cao kích thích tạo đơn bội. Đóng góp khác cho kiểu hình được tạo bản đồ bao gồm vùng lân cận qh12 (xem Liu, 2016)), mà có thể tính đến sự khác nhau giữa HIR của *MTL*^{TAL-FS} và NP2222-HI. Có lý khi suy ra rằng bộ hạt, HIR và tỷ lệ thui hạt được đặt qua *mtl* bởi sự tương tác đặc hiệu kiểu gen của bố và mẹ.

Bảng 4. Đặc điểm sinh sản của dòng chính sửa *MTL* (thường gọi là dòng *MTL^{TAL-FS}*). Bảng này thể hiện kích thích tạo đơn bội và tỷ lệ thụ hạt của dòng kích thích tạo ở nền NP2222. Các sự kiện chuyển gen được thử được chờ trái sau số lai thử được tạo ("bông") và thống kê thế hệ con.

Sự kiện	ID	(Các) đột biến	Các đặc điểm hạt				Dữ liệu phân tích mức bội thể			HIR
			Các bông	Khả năng tồn tại trung bình	Thui trung bình	% bị thui	Tổng số phôi	Các thể đơn bội giả định	Các thể đơn bội được xác nhận	
39A	3954	Hai alen (loại bỏ 13 bp & 28 bp)	4	162	128	44,10%	579	37	35	6,04%
23A	3924	Hai alen (loại bỏ 8 bp & 5 bp)	2	114	116	50,40%	128	18	16	12,50%
81A	3932	Đồng nhất (13 bp del)	2	165	129	43,90%	169	18	15	8,88%
81A	3317	Đồng nhất (loại bỏ 13 bp)	2	183	108	37,10%	343	19	19	5,54%
81A	3303	Đồng nhất (loại bỏ 13 bp)	1	189	100	34,60%	176	7	7	3,98%

38A	4108	Hai alen (loại bỏ 11 bp & 5 bp)	4	147	102	40,10%	379	28	26	6,86%
18A	22807- 4016	Đồng nhất (loại bỏ 8 bp)	8	144	97	40,20%	1025	47	44	4,29%
27A	22807- 4073	Hai alen (chèn 1 bp & loại bỏ 5 bp)	2	161	92	36,40%	180	18	18	10,00%
27A	22807- 4081	Hai alen (chèn 1 bp & loại bỏ 8 bp)	6	176	116	39,80%	931	45	44	4,73%
76A	22873- 3999	Đồng nhất (chèn 2 bp)	2	175	95	35,20%	117	17	16	13,68%
32A	22873- 3991	Đồng nhất (loại bỏ 1 bp)	2	140	105	42,90%	260	14	14	5,38%
Tổng số		Tổng số	15	160	108	40%	390	24	23	7%

Hình thành hạt đơn bội ở ngô là đặc điểm sau thụ tinh được khởi động bằng thể giao tử đực khiếm khuyết. Thực tế này được phản ánh trong dữ liệu biểu hiện *MTL*. Tương quan ARN-seq công khai cho thấy bản mã *MTL* kiểu dại là đặc hiệu cho bao phấn trong giai đoạn nở hoa (xem Sekhon, R. S., và cộng sự *Genome-wide atlas of transcription during maize development*, PLANT JOURNAL, **66**, 553-563 (2011), được kết hợp vào đây bằng cách viện dẫn), thống nhất với tương quan phát triển mà thấy loại trừ trong bao phấn trước khi nứt ra (xem Zhai, J., và cộng sự *Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers*, PNAS **112**, 3146–3151 (2015), được kết hợp vào đây bằng cách viện dẫn). Tác giả phát hiện ra rằng phấn hoa kiểu dại có nhiều hơn 18x *MTL* bản mã so với túi bao phấn sau khi nở hoa, chỉ ra rằng gen là đặc hiệu thể giao tử đực (Hình 3). Để bất hoạt *MTL* bằng ARNi dẫn đến tỷ lệ tạo thành đơn bội tăng chỉ đối với *MTL^{ARNi}* (Bảng 5). Có hai biến thể nổi được chú thích của *MTL*, phản ánh vị trí nổi tùy ý 81 nucleotit trước đầu 3' của exon 2. So với NP2222, *Mtl* tăng trong NP2222-HI nhưng không phải phấn hoa *MTL^{TAL-FS}* (Hình 4), trong khi độ dồi dào của biến thể nổi được chú thích là thống nhất.

Bảng 5. Cấu trúc ARNi 22503 (SEQ ID NO: 32) để bất hoạt *Mtl* dẫn đến kích thích tạo đơn bội.

GRMZM2G471240						Phôi được thử về			
RNAi		Các đặc điểm hạt				mức bội thể			
ID Cá thể	ID biến cổ	Bông	tồn tại được	bị thui	% bị thui	Phôi	các thể đơn bội	các bội lượng	HIR
5148	001	2	701	43	5,78%	369	3	366	0,81%
5149	001	2	186	22	10,58%	166	1	165	0,60%
5153	001	2	625	61	8,89%	323	7	316	2,17%
5161	001	3	1116	87	7,23%	485	4	481	0,82%
5170	028	2	629	23	3,53%	324	1	323	0,31%
5173	028	2	551	33	5,65%	322	0	322	0,00%
5187	028	3	379	27	6,65%	333	9	324	2,70%
3731	014	2	894	23	2,51%	263	4	259	1,52%
3732	014	2	648	49	7,03%	351	0	351	0,00%

3736	007	1	277	21	7,05%	277	0	277	0,00%
3737	007	1	223	49	18,01%	175	3	172	1,71%
3751	005	1	133	6	4,32%	118	0	118	0,00%
5 sự									
TỔNG SỐ	kiện	23	6362	444	6,52%	3506	32	3474	0,91%

Dịch khung ở *mtl* xảy ra ở axit amin 380, dẫn đến 20 axit amin thay đổi sau đó bộ ba mã kết thúc sớm mà cắt cụt protein bằng 29 axit amin (Hình 5a). Protein *MTL* kiểu đại thấy ở tương quan LS-MS của phần hoa RWK-NIL và NP2222, nhưng dưới giới hạn dò trong ba ngoài ba RWK và 3 ngoài 3 mẫu NP2222-HI (Bảng 6). Điều này chứng minh rằng ngay cả khi có bản mã đột biến *mtl* tạo ra trong phần hoa, protein không phát hiện được, xác nhận điều này là mất đột biến chức năng. Cả protein *MTL* đột biến và kiểu đại tái tổ hợp thể hiện hoạt tính phospholipaza trong ống nghiệm trong thử nghiệm phân cắt liposome huỳnh quang kiểu pPLAII α (Hình 5b). Điều này chứng minh rằng sự chú thích chức năng của gen *MTL* (tức là nó mã hóa cho protein phospholipaza) là đúng.

Bảng 6. Protein đóng và mở trong mẫu phần hoa NP2222 và NP2222-HI, bao gồm *MTL*, mà thấy ở phần hoa NP2222 nhưng không thấy ở NP2222-HI.

		log2 LFQ							
		NP2222			NP2222-HI				
		rep	rep	rep	rep	rep	rep		
		1	2	3	1	2	3	ID protein chính	Mô BlastP tốt nhất (cây S sớm nhất)
Không có ở NP2222-HI		23,3	23,3	23,4	ND	ND	ND	GRMZM2G028905	L-fucoza alpha-1,3-D-xylosyltransferaza
		22,9	24,1	23,9	ND	ND	ND	GRMZM2G046743	Chất vận chuyển 1 Lysin histidin
		23,2	23,1	23,2	ND	ND	ND	GRMZM2G310362	Protein liên kết polyadenylat 5
		24,2	24,2	24,1	ND	ND	ND	GRMZM2G130121	Chaperone protein ClpB2, cloplastic
		23,5	23,5	23,6	ND	ND	ND	GRMZM2G375807	ABC chất vận chuyển D; COMATOSE
									Phospho-2-dehydro-3-deoxyheptonat
		24,1	23,7	23,9	ND	ND	ND	GRMZM2G396212	aldolaza 1
		23,7	23,8	23,9	ND	ND	ND	GRMZM2G467907	protein liên kết ARN 47
		23,8	23,8	23,9	ND	ND	ND	GRMZM2G471240	Matrilin
		ND	ND	ND	23,9	23,8	23,9	GRMZM2G013607	Feredoxin-6, cloplastic
Không có ở NP2222		ND	ND	ND	22,1	22,1	22,2	GRMZM2G030971	Phospholipaza A I
		ND	ND	ND	24,5	24,6	24	GRMZM2G064967	Mannan endo-1,4-beta-mannosidaza
		ND	ND	ND	24	24,3	24,2	GRMZM2G143613	F-box protein
									HOTHEAD (a-dicarboxylic Fas mạch dài
		ND	ND	ND	24,4	24,3	24,1	GRMZM2G166906	tổng hợp)
		ND	ND	ND	23,2	23,5	23,6	GRMZM2G181259	beta-D-xylosidaza 2
*ND, Không thấy									

*ND, Không thấy

II. Thử nghiệm nảy mầm và định vị phần hoa

Dòng báo cáo chức năng chiều dài đầy đủ được sử dụng để mô tả đặc điểm định vị *MTL*. Không có tín hiệu được thấy ở phần hoa của NP2222 hoặc NP2222 + *mtl-GFP/mtl-GFP*. Ngược lại, phần hoa NP2222 + *MTL-GFP/MTL-GFP* thể hiện tín hiệu mạnh trong tế bào chất của hai tinh tử. Tín hiệu này được thấy ở tế bào chất giao tử sợi trong ống phấn nảy mầm. Các túi phôi NP2222 được cố định 18 giờ sau khi thụ phấn với phần hoa *MTL-GFP* có tín hiệu trong vùng của tế bào kèm thoái hóa thống nhất với của SCs được phân phối trong quá trình thụ phấn. Điều này chỉ ra *MTL* là một phần của đơn vị mầm đực mà lắng phủ trong túi phôi sau khi ống phấn vỡ. *MTL-GFP* nhưng không phải *mtl-GFP* loại trừ kích thích tạo đơn bội ở NP2222-HI (Bảng 3). Kết hợp các

dữ liệu này chỉ ra rằng *MTL* là phospholipaza đặc hiệu cho tế bào chất SC, và rằng dịch khung trong *mtl* thỏa hiệp sự định vị *MTL* hoặc tính ổn định trong phân hoa của dòng kích thích tạo đơn bội. .

Việc nhận diện của *MTL* là gen nguyên nhân ở kích thích tạo đơn bội của ngô cho phép giải phẫu kiểu hình nhiều tính trạng kết hợp về mặt lịch sử với đặc điểm. Đột biến phospholipaza được kết hợp với sự nảy mầm phân hoa và phát triển ống chậm (xem Kim, H. J., và cộng sự *Endoplasmic reticulum- and golgi-localized phospholipase A2 plays critical roles in Arabidopsis pollen development and germination*, PLANT CELL 23, 94-110 (2011)), nhưng chúng là bình thường ở dòng RWK, Stock 6 và *MTL^{TAL-FS}* (Hình 6A, B). Các bông được thụ phấn với phân hoa NP2222-HI và *MTL^{TAL-FS}* thể hiện ~10-25% thụ phấn hỏng, và thử nghiệm cạnh tranh phân hoa thể hiện rằng RWK chịu sự biến dạng chia tách (SD) (Hình 6C), thống nhất với báo cáo trước đó (xem Xu, X., và cộng sự *Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize*, J. EXP. BOT. 64, 1083-1096 (2013)). Lai với phân hoa NP2222-HI + *MTL*/0 bán hợp tử tạo ra thiên hướng tỷ lệ về phía *MTL*+ thế hệ con (Hình 6D), chỉ ra rằng dòng kích thích tạo SD góp phần vào *mtl*. Sự thụ phôi, phụ phẩm tồn của việc kích thích tạo đơn bội liên quan đến hỏng sự phát triển nội nhũ, xảy ra với tỷ lệ tương tự ở dòng kích thích tạo tự nhiên và *MTL^{TAL-FS}* (Bảng 4). Kết hợp các dữ liệu này ngụ ý *mtl* trong mỗi khiếm khuyết sinh sản kết hợp với kích thích tạo đơn bội. Hai cơ chế thường được dùng để giải thích thông tin đơn bội là thụ phấn đơn và loại trừ bộ gen sau thụ tinh trứng (xem Sarkar, K. R. & Coe, E. H, *A genetic analysis of the origin of maternal haploids in maize*, GENETICS 54, 453-464 (1966); Zhang, Z., và cộng sự, *Chromosome elimination and in vivo haploid production induced by Stock 6-derived inducer line in maize (Zea mays L.)*, PLANT CELL REP. 27, 1851–1860 (2008); và Barret, P., Brinkmann, M., & Beckert, M., *A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize*, THEOR. APPL. GENET. 117, 581–594 (2008)). Trong cơ chế đầu, các thể đơn bội tạo ra từ sự thụ phấn của tế bào trung tâm nhưng không phải trứng, mà sau đó phát triển qua sự sinh sản đơn tính. Trong cơ chế sau, thụ phấn kép có trước việc loại trừ nhiễm sắc thể đực. Việc làm rõ cơ chế chính xác sẽ cần phôi thai cẩn thận sau khi thụ phấn *MTL^{TAL-FS}*, cùng với dữ liệu định lượng theo dõi sự tồn lưu ít của ADN đực ở các thể đơn bội của ngô.

Kích thích tạo đơn bội gần đây được thiết kế ở *Arabidopsis* thông qua thao tác của *CENTROMERIC HISTONE3*, mà khiến loại bỏ bộ gen một bố mẹ qua sự không cân bằng đoạn trung tâm sau khi trứng thụ tinh giữa các bộ gen được lai. Nỗ lực để tái tạo điều này ở ngô đã thành công (xem Ravi, M. & Chan, S. W. L. *Haploid plants produced by centromere-mediated genome elimination*, NATURE 464, 615-618 (2010)), nhưng tài liệu này là trường hợp đầu tiên của hệ dòng kích thích tạo đơn bội được phát động bởi protein tế bào chất mà không liên kết chất nhuộm sắc. Do đó, việc này nhấn mạnh tầm quan trọng của thành phần tinh tử không nhân trong thành công sinh sản và truyền bộ gen chính xác. Việc chuyển đổi của *MTL* ở loài thân cỏ (xem Hình 7), đặc biệt ở lúa nơi mà chất tương đồng gần nhất là đặc hiệu phân hoa và còn thấy ở tinh tử, đề xuất các phát hiện này sẽ dẫn đến sự phát triển của dòng kích thích tạo đơn bội đặc hiệu trong mới ở cây trồng quan trọng.

III. Dịch phun hóa học để kích thích tạo HI mới

Sau khi phát hiện ra rằng phospholipaza *MTL* (và đặc biệt là, đột biến dịch khung mất chức năng *mtl*) phát động kích thích tạo đơn bội, tác giả thử nghiệm xem chất ức chế phospholipaza, axit béo, hoặc các hợp chất lipid khác có thể hoạt động như dòng kích thích tạo đơn bội hóa học hay không.

Tác giả đã thu được thành công kích thích tạo đơn bội mới thông qua ứng dụng nhiều chất hóa học lên hoa, tơ, bông, núp tua, và phần hoa. Các chất hóa học này đều thuộc nhóm lipid hoặc chất ức chế phospholipaza. Bảng 7 mô tả các chất hóa học đó và tỷ lệ kích thích tạo đơn bội mà tạo ra từ việc áp dụng chúng. Cũng có thông tin về chế phẩm và thông tin về cách áp dụng trong Bảng 7. Từ bảng này rõ ràng là nhiều loại hợp chất lipid kích thích tạo các thể đơn bội mới khi được áp dụng trước, trong, hoặc sau khi thụ phấn cho tơ, mà là mô đầu nhụy đặc hiệu cho cây ngô, phần hoa, núp tua (mà chứa hoa đực chứa phần hoa trước khi phần hoa rụng), và bông. Các hợp chất này bao gồm dầu hỗn hợp, axit béo, este của axit béo, phospholipit, và chất ức chế phospholipaza.

Tác giả thử nghiệm nhiều chất ức chế phospholipaza, bao gồm manoalit, chất ức chế phospholipaza không có bất kỳ mạch axit béo. Hóa học này có khả năng kích thích tạo thông tin thể đơn bội mới. Với kết quả này trong tay, tác giả bắt đầu thử nghiệm nhiều hợp chất. Tác giả thấy rằng hợp chất hoạt tính nhất để kích thích tạo mới là methyl alpha linolenyl flophosphonat (MALFP), chất ức chế phospholipaza mạnh và chất ức

chế của enzyme biến đổi lipid khác. Hợp chất này chứa ba mạch axit béo linolenat (mười tám cacbon với ba liên kết ba) và nhóm metylol và nguyên tử flo ở vị trí nhóm đầu. Hợp chất kích thích tạo đơn bội mới hoạt tính cao khác là metyl arachidonyl fluorophosphonat (MAFP), chất ức chế phospholipaza khác với cấu trúc mạch axit béo hơi khác nhau khác (20 cacbon và bốn liên kết đôi). Không bị ràng buộc bởi lý thuyết, MALFP và MAFP ức chế phospholipaza bằng cách nằm trong túi liên kết mạch axit béo, và trong một số trường hợp xúc tác phosphoryl hóa không thuận nghịch của axit amin serin ở vị trí hoạt tính (xem Lio Y.C., Reynolds L.J., Balsinde J., và Dennis E.A. *Irreversible inhibition of Ca(2+)-independent phospholipase A2 by methyl arachidonyl fluorophosphonate*. BIOCHIM BIOPHYS ACTA 1302:55–60 (1996)). Một số axit béo như axit linoleic và axit linolenic còn có thể ức chế không cạnh tranh phospholipaza bằng cách nằm trong túi liên kết mạch axit béo mà không biến đổi cộng hóa trị vị trí hoạt tính, dù hiệu quả của việc ức chế này yếu hơn. Ballou, L.R., và Cheung, W.Y. *Inhibition of human platelet phospholipase A2 activity by unsaturated fatty acids* PROC NATL ACAD SCI 82 (2): 371-375 (1985)). Các viện dẫn này có thể giúp giải thích lý do cho phát hiện của tác giả rằng kích thích tạo đơn bội mới không chỉ phát động bằng các chất ức chế mạnh của phospholipaza như MALFP và MAFP, mà còn bởi este của axit béo bao gồm metyl este của axit arachidonic (AAME), mà có 20 cacbon và bốn liên kết đôi (este của axit béo đa bão hòa có 20 cacbon), cũng như axit linoleic (LLE), mà có 18 cacbon và hai liên kết đôi (18:2), và axit linolenic (LNA), mà có 18 cacbon và ba liên kết đôi (18:3), và etyl este của axit oleic (OAEE), và axit arachidonic (mà có 20 cacbon và bốn liên kết đôi). Lại nghĩ rằng các mạch axit béo không bão hòa này, hoặc là ở dạng axit béo tự do của chúng, dạng este của chúng (mà được tổng hợp trong phòng thí nghiệm và có thể là có khả năng xâm nhập mô) hoặc ở một số dạng khác, có thể đóng vai trò là chất ức chế phospholipaza. Trong khi thường không được công nhận rằng dạng este của các hợp chất này có thể ức chế phospholipaza (xem Ballou, nêu trên), tác giả thấy rằng metyl và etyl este của axit béo này thay vào đó dẫn đến chức năng đơn bội mới. Xem Bảng 7. Lưu ý, axit béo bão hòa và dẫn xuất, bao gồm axit stearic và etyl este của axit palmitic, đã được thử. Tình trạng kỹ thuật là các phân tử mà chứa mạch bên của axit béo bão hòa không hoặc không thể có khả năng ức chế cạnh tranh phospholipaza bằng cách nằm ở vị trí hoạt tính của nó. Xem Ballou, trước đây. Các hợp chất này được tin là không có khả năng ức chế phospholipaza. Tuy nhiên, tác giả đã có khả năng sử dụng một cách

đáng ngạc nhiên các phân tử chứa mạch bên của axit béo bão hòa để kích thích tạo thông tin của các thể đơn bội *mới*. Ví dụ, etyl este của axit palmitic (PAEE) kích thích tạo các thể đơn bội. Xem Bảng 7.

Tác giả đã quyết định thử nghiệm xem các dầu nhất định và phospholipit có thể phát động kích thích tạo đơn bội *mới* không, bao gồm dầu thực vật thông thường. Tác giả thấy rằng một hỗn hợp lipit kích thích tạo đơn bội *mới* hoạt tính cao là dầu ngô, chiết xuất cây tự nhiên có thể được mua từ cửa hàng tạp phẩm. Hỗn hợp lipit hoạt tính cao khác là dầu hạt lanh mà có được mua từ cửa hàng đồ ngũ kim. Dầu hạt cải, dầu thực vật, dầu lạc, dầu vừng, hoặc bất kỳ dầu khác, khi được áp dụng đúng cách, vẫn hoạt động như thế. Hợp chất hoạt tính cao khác là 1,2-distearyl-sn-glyxero-phosphatidylcholin (DSPC) (còn gọi là 18:0 18:0 PC), phospholipit với nhóm đầu phosphatidylcholin và hai mạch axit béo bão hòa có 18 cacbon. Phospholipit thông thường này có mặt trong nhiều vật liệu sinh học vì nó là một trong số các thành phần cốt lõi và dồi dào nhất của lớp đôi phospholipit, mà gồm màng tế bào của tất cả các loại sinh vật sống trên trái đất. Hợp chất kích thích tạo đơn bội khác là phosphatidyletanolamin với một stearyl và một chuỗi oleoyl (còn gọi là 1-stearyl-2-oleoyl-sn-phosphatidyletanolamin), phospholipit thông thường khác. Phospholipit khác, cũng như lyso-phospholipit và triaxylglyxerit khác, diaxylglyxerol, lysophospholipit, triterpenoit este, hoặc glyxerolipit sẽ đóng vai trò là dòng kích thích tạo đơn bội *mới*. Dầu này và phospholipit đã biết là ức chế phospholipaza, nhưng chúng có thể là vật liệu nguồn để tạo ra phụ phẩm có hoạt tính phospholipaza, bao gồm axit béo chính mà ức chế phospholipaza. Thực tế là các hợp chất này hoạt động là chất hóa học dòng kích thích tạo đơn bội *mới* do đó chỉ các dầu này và phospholipit có thể ức chế phospholipaza gián tiếp, hoặc có thể chúng tạo ra sự không cân bằng hoặc thay thế về lipit hoặc chế phẩm màng của phần hoa dẫn đến kích thích tạo đơn bội. Dù theo cách nào, từ công trình nghiên cứu này, rõ ràng là một phạm vi rộng lớn của các hợp chất chứa lipit, ngoài chất ức chế phospholipaza không lipit, khi được áp dụng cho hoa, có thể đóng vai trò làm dòng kích thích tạo hóa học của hạt đơn bội.

Tác giả còn thử nghiệm xem chất đệm tác giả sử dụng để áp dụng các hợp chất này cho hoa là có khả năng tự tạo ra đơn bội. Các chất đệm này kích thích tạo khác nhau các hỗn hợp chất hoạt động bề mặt nhất định và/hoặc muối và/hoặc DMSO (dimetyl sulfoxit) và/hoặc bổ sung liên quan khác. Các chất đệm này không có khả năng tạo ra

đơn bội, như có thể thấy ở dòng 121–123 của Bảng 7. Tương tự, kích thích tạo đơn bội mới không có bổ sung bất kỳ hợp chất là cực hiếm ở ngô, như có thể thấy từ dòng 1. Hai thể đơn bội được thấy trong số 3.073 thể hệ con được thử (Kelliher, và cộng sự *Maternal haploids are preferentially induced by CENH3-tailswap transgenic complementation in maize*, FRONTIERS IN PLANT SCIENCE7: 414 (2016)). Do đó, tỷ lệ kích thích tạo đôi chứng được thấy là xấp xỉ 0,065%, mà nằm trong khoảng nền được báo cáo của kích thích tạo đơn bội sinh cái (0,05 – 0,1%) (Chang, M.-T., and Coe, E. "Doubled haploids," in *Molecular Genetic Approaches to Maize Improvement*, eds A.L.Kriz and B.A. Larkins (Heidelberg: Springer), 127–142).

Cách áp dụng các lipit khác nhau này và các hợp chất ức chế lipaza có thể khá biến đổi và vẫn tạo ra các thể đơn bội. Một cách áp dụng này là hòa tan các hợp chất trong dung môi muối với bổ sung 1% DMSO. Dung môi khác này là hỗn hợp chất hoạt động bề mặt (Bảng 8). Mô áp dụng cũng có thể thay đổi, như được thể hiện rõ trong Bảng 7, trong đó áp dụng lên cả phần hoa và tơ trong khi thụ phấn, cũng như lên núp tua vài giờ hoặc ngày trước khi thụ phấn, có thể tạo ra việc hình thành các thể đơn bội. Hơn nữa, khoảng rộng nồng độ của các hợp chất có thể kích thích tạo các thể đơn bội, từ 20 uM đến 100 mM, hoặc từ 0,2 mg/mL đến 50 mg/mL. Thật vậy, khoảng rất rộng hơn khoảng này có khả năng kích thích tạo đơn bội, như có thể thấy trong Bảng 7.

Trong hai trường hợp, tác giả cũng phát triển hỗn hợp chất hoạt động bề mặt mới để giúp các nhóm lipit cụ thể có khả năng nhũ hóa và tạo thành vi nhũ (hỗn hợp nước hữu cơ với kích thước giọt của lipit nhỏ hơn kích thước nhất định – đường kính ít nhất nhỏ hơn 10 micromet và trong nhiều trường hợp đường kính nhỏ hơn 1 micromet). Xem 8 đối với chế phẩm của hai hỗn hợp chất hoạt động bề mặt. Hỗn hợp "91" được sử dụng lý tưởng với axit béo ở nồng độ bằng 6,5 phần hỗn hợp chất hoạt động bề mặt so với 1 phần axit béo. Tuy nhiên, nhiều nồng độ liên quan và hỗn hợp chất hoạt động bề mặt liên quan cũng có thể hoạt động. Hỗn hợp chất hoạt động bề mặt "92" được sử dụng lý tưởng với este của axit béo, cũng như dầu, hoặc là ở dạng tinh khiết hoặc là hỗn hợp triglycerit. Nó được trộn lý tưởng ở tỷ lệ 10 phần hỗn hợp chất hoạt động bề mặt so với 1 phần este hoặc dầu. Mỗi trong các hỗn hợp này hoặc hỗn hợp liên quan của các hợp chất có thể thích hợp để hòa tan đúng hoặc hình thành vi nhũ với nhóm phân tử lipit liên quan ngoài axit béo, este, và dầu. Chúng có thể bao gồm phospholipit, diacylglycerol, lysophospholipit, triterpenoid este, hoặc glycerolipit.

Các hỗn hợp chất hoạt động bề mặt này thường được trộn với thành phần hoạt tính và sau đó trộn ở phần trăm nhất định với chất đệm trong nước để tạo ra nhũ tương và lý tưởng là vi nhũ. Khi tạo ra nhũ tương, phần trăm của hỗn hợp chất hoạt động bề mặt + Thành phần hoạt tính trong chất đệm trong nước có thể là ở bất cứ đâu từ 0,01 - 50%. Chất đệm trong nước có thể bao gồm nhiều thứ. Tác giả đã sử dụng 1x PBS + 1% DML và PBS + 50% DML, nhưng các chế phẩm khác cũng hoạt động. "PBS" là nước muối đệm phosphat; "DML" là dimetyl lactamit.

Việc sử dụng các chất đệm này một mình không thể dẫn đến kích thích tạo đơn bội mới, như có thể thấy trong Bảng 7, dòng 121–123.

Cách áp dụng các hợp chất có thể lấy bất kỳ trong số nhiều dạng bao gồm nhúng cây, tiêm cây, vi tiêm, tưới ẩm phần hoa, phun sương phần hoa, phun phần hoa, phun sương cây, phun cây, đóng tơ, phun tơ, v.v. Tác giả thường dùng nhất là phun núm tua, phần hoa, và tơ trước, trong, hoặc sau khi thụ phấn. Tác giả đã thu được thông tin kích thích tạo *mới* với DSPC và axit linolenic khi núm tua được phun hai ngày trước khi thụ phấn. Tác giả cũng thu được thông tin kích thích tạo với DSPC khi núm tua được phun ngay vài phút trước khi gom phần hoa và sau khi thụ phấn. Tương tự, tác giả đã thu được thông tin kích thích tạo *mới* với DSPC khi phần hoa được phun ngay trước hoặc trong khi thụ phấn, và cả khi tơ được phun trước, trong, hoặc sau khi thụ phấn, bao gồm đến hai ngày trước khi thụ phấn và đến sáu giờ sau khi thụ phấn.

Tác giả cũng chứng minh thông tin kích thích tạo *mới* khi các áp dụng khác nhau được sử dụng, bao gồm bình phun bằng nhựa, giá rẻ mà có thể mua ở cửa hàng tiện lợi, và ống phun, mà là thiết bị y tế được dùng để phân phối thuốc qua đường miệng dưới dạng giọt nước rất nhỏ. Trong khi bình phun thường tạo ra kích thước giọt có đường kính nằm trong khoảng 50 – 150 micromet, ống phun có thể tạo ra kích thước giọt nhỏ hơn 10 micromet. Điều này có lợi cho việc áp dụng lên phần hoa vì kích thước phần hoa nằm trong khoảng từ 20 – 200 micromet về đường kính ở hầu hết các cây (xấp xỉ 70 micromet ở ngô; xấp xỉ 50 micromet ở lúa) và nếu một trong các giọt từ bình phun thông thường chạm hạt, giọt có thể lớn hơn hạt. Phần hoa đặc biệt nhạy cảm với độ ẩm và chạm thấm lọc. Nếu hạt phần tiếp xúc với giọt có kích thước quá lớn, hạt đó không nảy mầm thành công, làm lớn ống phấn. Điều tương tự có thể xảy ra nếu hạt nằm trên đầu nhụy hoặc tơ mà quá ướt— ví dụ, nếu tơ hoặc đầu nhụy nhận quá nhiều áp dụng lipid (nó là vi nhũ, hoặc giọt lipid đơn hoặc vi cầu được tạo ra trong dung dịch nước). Nếu tác giả

áp dụng nhiều hơn 3 mL dịch phun lipid lên phần hoa hoặc tơ qua bình phun hoặc ống phun, tác giả thường thu ược bộ hạt rất thấp.

Do đó cách áp dụng là quan trọng và cụ thể là nó quan trọng nhất để thử áp dụng là thể tích dịch phun nhỏ nhất có thể lên phần hoa hoặc đầu nhụy hoặc tơ hoặc lên hoa nói chung, và cách tốt nhất để phân bố thành phần hoạt tính, đặc biệt là lipid, với thể tích nhỏ nhất có thể là để tạo ra vi nhũ của lipid đó trong chất đệm có nước trong đó kích thước giọt lipid là ở dưới mức micromet, và sau đó phân phát vi nhũ đó với ống phun hoặc thiết bị tương tự có khả năng tạo ra kích thước hạt trong khoảng 1-2 micromet. Điều này tạo ra việc phân phối hàng triệu hạt thành phần hoạt tính ở dạng "hơi" lên môi liên quan, với kích thước giọt nhỏ hơn >100,000 lần thể tích so với chính hạt phấn (nếu giọt là 1 micromet, và hạt phấn là 50 micromet, thì đường kính của giọt là nhỏ hơn 50 lần so với hạt. Và $50^3 = 125,000$, nên thể tích của giọt là >125,000 lần nhỏ hơn hạt phấn. Với vi nhũ, kích thước giọt lipid có thể nằm trong khoảng từ xấp xỉ 20–1000 nm về đường kính, do đó phụ thuộc vào nồng độ của thành phần hoạt tính cộng hỗn hợp chất hoạt động bề mặt trong dung dịch nước, một chất có thể có một hoặc nhiều gói lipid trong mỗi giọt được phân phối đến mỗi hạt phấn hoa hoặc tơ. Hơn nữa, xem xét một số phân tử loại lipid hoặc chất ức chế phospholipaza có thể được hòa tan trong dung dịch nước thay vì được phân phối là vi nhũ, tác giả thấy rằng nhiều loại sự đa dạng lớn về nồng độ của các hợp chất có thể kích thích tạo các thể đơn bội, từ 20 $\mu\text{g/mL}$ lên đến 50 mg/mL . Các nồng độ cao hơn có thể sẽ kích thích tạo các thể đơn bội *mới*.

Bảng 7. Các hợp chất gây ra Kích thích tạo đơn bội mới.

Các chất tíc chế PLA2	Cấu trúc trên các hình vẽ	hợp chất	áp dụng	nồng độ	chế phẩm	áp dụng mô	hạt	thu	FF	lượng bội	các thể đơn bội	HIR				
	Hình 10	Metyl alpha-linolenoyl flophosphonat (MALFP)	ống phun	2%	chế phẩm 91	phấn hoa	226	36	220	173	17	8,9%				
				2%			134	5	298	128	2	1,5%				
				10%			26	54	392	22	2	8,3%				
				20%			257	5	200	176	0	0,0%				
				2%	chế phẩm 92	39	0	441	36	0	0,0%					
				10%		207	101	307	188	2	1,1%					
				20%		290	157	20	176	0	0,0%					
					Hình 11	Metyl arachidonyl flophosphonat (MAFP)	ống phun	2%	chế phẩm 91	phấn hoa	220	4	196	212	1	0,5%
								2%			437	9	20	346	13	3,6%
								20%			281	21	50	170	6	3,4%
10%	chế phẩm 92	110	4					61	101	0	0,0%					
20%		338	19					150	176	0	0,0%					
	Hình 12	Palmityl triflometylketon (PACOCF3)	bình phun					100 uM	PBS + 1% DMSO	núm tua	220	15	-	88	2	2,2%

Các chất ức chế PLA2

Hình 13	Arachidonyl triflometylketon (AACOCF3)	bình phun	100 uM	PBS + 1% DMSO	phần hoa	240	12	-	172	4	2,3%
Hình 14	Manoalit	bình phun	50 uM	PBS + 1% DMSO	phần hoa	320	35	-	176	2	1,1%
Hình 15	Etyl este của axit linoleic (LLAEE)	ống phun	2%	chế phẩm 92	núm tua + phần hoa	114	30	292	254	1	0,4%
			12%			366	36	16	256	5	1,9%
			20%			461	33	114	187	1	0,5%
		ống phun	2%	chế phẩm 92	phần hoa	401	25	30	262	2	0,8%
		ống phun	20%	chế phẩm 92		267	49	248	235	0	0,0%
		bình phun	5 mg / mL	PBS + 1% DMSO		215	83	24	176	0	0,0%
		ống phun	10 mg / mL	PBS + 1% DMSO		50	10	-	0	40 (không được kiểm tra)	N/A
		ống phun	50 mg / mL	PBS + 50% DML		241	9	344	191	1	0,5%
Hình 16		ống phun	2%			401	16	20	165	1	0,6%

Este của axit béo

	etyl este của axit linolenic (LNAEE)		10%	chế phẩm 92	núm tua + phần hoa	714	128	137	257	0	0,0%
			20%			647	39	187	88	1	1,1%
Hình 17	Metyl este của axit arachidonic (AAME)		2%	chế phẩm 92	phần hoa	274	10	374	144	2	1,4%
		ống phun	2%	chế phẩm 92		490	54	474	174	2	1,1%
			3 mg / mL	PBS + 1% DMSO		248	70	42	247	1	0,4%
			4 mg / mL	PBS + 1% DMSO		403	108	49	87	1	1,1%
		bình phun	0,1 mg / mL	PBS + 1% DMSO	phần hoa	146	20	134	139	1	0,7%
			0,5 mg / mL		núm tua	224	19	117	175	1	0,6%
Hình 18	metyl este của axit oleic (OAME)	ống phun	2%	chế phẩm 92	phần hoa	224	58	18	166	2	1,2%
			10%			79	123	474	76	1	1,3%
			20%			282	66	505	158	0	0,0%
Hình 19	Etyl este của axit oleic (OAE)	ống phun	10%	chế phẩm 92	phần hoa	15	2	459	14	0	0,0%
			2%			226	92	100	104	1	1,0%
			20%			364	57	407	344	15	4,2%
Hình 20		ống phun	2%			7	0	317	7	0	0,0%
						311	61	55	176	0	0,0%

	Etyl este của axit palmitic (PAEE)		20%	ché phẩm 92	phần hoa	40	0	564	không có dữ liệu	không có dữ liệu	N/A
Hình 21	Etyl este của axit palmitoleic (PLAEE)	ống phun	2%	ché phẩm 92	phần hoa	218	41	55	87	1	1,1%
			20%			163	19	156	172	3	1,7%
Dầu	Hỗn hợp của triglyxerit, bao gồm các loại chứa linoleic và mạch bên linolenic, cũng như nhiều loại mạch bên khác	ống phun	2%	ché phẩm 92	núm tua + phần hoa	611	27	50	86	2	2,3%
			10%			61	13	728	57	1	1,7%
			20%		phần hoa	688	38	112	595	5	0,8%
			2%			192	211	529	138	2	1,4%
			2%			144	157	371	126	6	4,5%
			10%			466	34	299	205	3	1,4%
			10%		phần hoa	36	7	1073	34	2	5,6%
			20%			78	202	632	69	2	2,8%
			20%		núm tua	26	47	539	25	0	0,0%
			2%			618	27	304	168	2	1,2%
			10%			278	25	30	149	0	0,0%
	Dầu ngô	ống phun	2%	ché phẩm 92	phần hoa	234	32	118	83	5	5,7%
			5%		núm tua	287	19	32	166	7	4,0%
			20%		phần hoa	326	94	180	217	5	2,3%

Hình 22	Axit alpha-linolenic (aLNA)	ống phun	chế phẩm 91	phấn hoa	305	11	0	255	0	0,0%
					373	4	57	343	1	0,3%
					228	32	76	198	0	0,0%
					377	25	20	166	5	2,9%
					263	48	63	244	4	1,6%
					528	8	25	170	6	3,4%
					500	40	35	432	8	1,8%
					393	87	240	85	3	3,4%
		bình phun	chế phẩm 91	phấn hoa	313	93	50	255	1	0,4%
					234	98	90	164	2	1,2%
					249	14	27	159	1	0,6%
					245	18	37	156	4	2,5%
Hình 23	Axit gamma-linolenic (gLNA)	ống phun	chế phẩm 91	phấn hoa	363	16	139	346	2	0,6%
					93	2	445	81	1	1,2%
					4	1	451	3	1	25,0%
					383	161	330	245	3	1,2%
					142	6	296	206	3	1,4%
					79	1	839	66	0	0,0%
		bình phun	chế phẩm 91	phấn hoa	363	16	139	346	2	0,6%
					93	2	445	81	1	1,2%
					4	1	451	3	1	25,0%
					383	161	330	245	3	1,2%
					142	6	296	206	3	1,4%
					79	1	839	66	0	0,0%

Axit béo

Hình 24	Axit oleic	ống phun	20%	chế phẩm 91	tơ + phần hoa	384	99	487	260	2	0,8%
						93	0	212	72	0	0,0%
						291	47	22	173	3	1,7%
						217	103	436	215	3	1,4%
Hình 25	Axit linoleic	ống phun	2%	91	phần hoa	372	31	-	325	2	0,6%
						252	9	-	252	0	0,0%
						324	24	-	252	1	0,4%
						336	12	-	295	0	0,0%
Hình 26	Axit arachidonic	ống phun	10 mg / mL	PBS + 1% DMSO	tơ	42	8	-	38	0	0,0%
						386	0	39	326	4	1,2%
						137	5	208	176	0	0,0%
						300	4	54	301	1	0,3%
Hình 27	Axit stearic	ống phun	2%	91	tơ	N/A	N/A	N/A	98	2	2,0%
						730	42	188	264	0	0,0%

			10%		phấn hoa	200	106	543	188	0	0,0%
Hình 28	Axit 9(Z)-11(E)-linoleic liên hợp	bình phun	0,1 mg / mL	PBS + 1% DMSO	tơ	N/A	N/A	N/A	98	2	2,0%
						300	2	-	88	0	0,0%
						181	63	240	82	6	6,8%
						247	73	112	82	3	3,5%
						515	16	93	175	1	0,6%
						227	95	38	170	1	0,6%
						300	12	-	87	1	1,1%
						250	60	-	176	0	0,0%
						200	0	-	160	1	0,6%
						300	14	-	88	0	0,0%
						200	18	-	175	1	0,6%
						200	20	-	174	2	1,1%
						132	10	296	88	0	0,0%
						141	8	223	87	1	1,1%
						288	12	64	176	0	0,0%

BẢNG 8. Công thức cho chế phẩm 91 ("F91") và chế phẩm 92 ("F92"), mà giúp các thành phần hoạt tính như axit béo, dầu, este, và phospholipit được nhũ hóa thành các giọt lipit rất nhỏ. Chế phẩm 91 là hỗn hợp chất hoạt động bề mặt được trộn với axit béo tối ưu ở tỷ lệ 1 phần axit béo với sáu phần rưỡi chế phẩm 91. Hỗn hợp nghĩa là để tạo ra nhũ tương với chất đệm nước ở các nồng độ bằng 0,001 – 50%. Chế phẩm 92 cũng là hỗn hợp chất hoạt động bề mặt mà có thể được trộn với este hoặc dầu tối ưu ở tỷ lệ 1 phần thành phần hoạt tính với 10 phần chế phẩm 92, và sau đó được nhũ hóa trong chất đệm nước ở các nồng độ bằng 0,001 – 50%.

Chế phẩm 91	Lượng (gam)	Tương đồng hóa học
Jeffsol AG1560	95,99	dung môi butylen carbonat
Emulpon CO360	24,02	etoxylat dầu thầu dầu (36 mol EO)
Witconate P1220EH	24,08	dodexyl (nhánh) benzen sulfonat, muối Ca. 60% dung dịch trong etyl-hexanol
Agrimer AL-22	12,03	poly(alkyl-pyrrolidon). Alkyl chính C16
	156,12	
Chế phẩm 92	Lượng (gam)	Tương đồng hóa học
Agnique AMD 3-L	104	dung môi dimetyl lactamit
Witconic 1298 mềm	15,67	axit dodexyl (thẳng) benzen sulfonic
Surfonic T-5	12,2	talowamin etoxylat (5 mol EO)
Surfonic T-15	24,21	talowamin etoxylat (15 mol EO)
	156,08	

IV. Làm tăng HIR và khôi phục bộ hạt bằng cách áp dụng hợp chất lipit

Kỹ thuật nhận diện phân tử, bao gồm nhận diện hóa chỉ tế bào và nhận diện lipit tế bào, của phần hoa dòng kích thích tạo và không kích thích tạo (trương đối) thể hiện rằng phần hoa của dòng kích thích tạo đơn bội đặc biệt thiếu ở các loại nhóm lipit nhất định và dư thừa ở các loại khác. Sự thiếu hụt này đặc biệt rõ rệt trong nhóm phân tử lipit

18 mạch cacbon, mặc dù nó cũng được thấy ở nhóm 20 mạch cacbon. Sự thiếu hụt này đặc biệt rõ rệt với lipit 18 mạch cacbon với một và hai liên kết đôi (cái gọi là oleat và linoleat) và sự dư thừa đặc biệt thấy với lipit 18 mạch cacbon với ba liên kết đôi (cái gọi là linolenat). Về loại lipit mà được thay đổi trong phần hoa của dòng kích thích tạo đơn bội, nó rất rộng, và bao gồm triglyxerit, diaxylglyxerit, axit béo tự do, lyso-phospholipit, và phospholipit. Thay đổi về hàm lượng trong phần hoa của dòng kích thích tạo đơn bội còn thay đổi qua các mức bão hòa axit béo khác nhau. Sau khi xem dữ liệu này, tác giả đã quyết định áp dụng các hợp chất lipit lên hoa trong quá trình thụ phấn với phần hoa của dòng kích thích tạo đơn bội để xem liệu việc thêm đơn giản các loại lipit nhất định có thể tác động đến tỷ lệ hình thành thể đơn bội hoặc số nhân không. Tác giả đặc biệt quan tâm đến việc áp dụng các hợp chất mà có, trong cấu trúc phân tử của chúng, một số mạch axit béo bão hòa 18 hoặc 20 cacbon mà trong số đó được thấy ít dồi dào hơn trong phần hoa của dòng kích thích tạo đơn bội (ví dụ oleat và linoleat).

Việc áp dụng lipit còn có thể dẫn đến tăng tỷ lệ hình thành thể đơn bội, bộ hạt, và giảm tần suất thui hạt, như được chỉ ra trong các Bảng 10–13. Bảng 10 thể hiện chín thử nghiệm được kiểm soát, riêng biệt sử dụng hợp chất metyl alpha linolenyl flophosphonat (MALFP). Mỗi thử nghiệm gồm hai bông được lai bởi cùng cây bố mẹ đực dòng kích thích tạo đơn bội, vào cùng ngày, ở thời điểm gần như giống hệt nhau (trong vòng 5 phút của mỗi thử nghiệm) từ cùng quần thể phần hoa giống hệt nhau được gom từ cùng bố mẹ đực. Hai bông này hoặc là nhận áp dụng đệm đối chứng, hoặc thành phần hoạt tính cộng đệm. Cả cây thử đực và cái được trồng cạnh nhau dưới cùng điều kiện trong cùng nhà kính. Trên trung bình, và so với áp dụng đối chứng của đệm không có MALFP, khi MALFP được áp dụng tỷ lệ kích thích tạo đơn bội tăng thêm 3,6%, số hạt tăng thêm gần 42 mỗi bông, và số của tổng các thể đơn bội được tái tạo mỗi bông tăng thêm 11,8 thể đơn bội mỗi bông, và tần suất thui phôi giảm đi 1,1% (Bảng 9). Tất cả điều này thể hiện sự cải thiện đối với sơ đồ tạo ra kích thích tạo đơn bội. Cải thiện đặc biệt xuất hiện khi chế phẩm 91 được dùng làm hỗn hợp chất hoạt động bề mặt để hòa tan MALFP. Cơ chế mà nhờ đó các chất ức chế này có tác dụng thêm vào với *mtl* dẫn đến tỷ lệ kích thích tạo cao hơn có thể là chúng đang phá vỡ các phospholipaza hoặc protein biến đổi lipit khác.

Cùng loại dữ liệu được gom đối với etyl este của axit béo LLAEE, dẫn đến làm tăng tỷ lệ kích thích tạo đơn bội 2,4%, tăng trên 43 hạt được tạo ở mỗi bông, tăng trên 8

thể đơn bội mỗi bông, và giảm tần suất thui phôi 3,7%. Mặc dù nhiều trong số các xử lý này liên quan đến lượng khác nhau của thành phần hoạt tính, ở phần lớn các thử nghiệm này LLAEF dẫn đến sự tăng cường về thống kê hình thành đơn bội và tạo hạt. Xem Bảng 10.

Cùng loại dữ liệu được gom đối với axit béo axit linoleic (LLA) (Bảng 11), dẫn đến tăng tỷ lệ kích thích tạo đơn bội bằng 1,5%, tăng 41 hạt sống được mỗi bông và 7,6 thể đơn bội mỗi bông, và giảm tỷ lệ thui phôi 2,9%. Đây là kết quả của 22 thử nghiệm riêng biệt, lại với mỗi thử nghiệm bao gồm một bông được thụ phấn bằng đốm đối chứng và một bông được thụ phấn bằng thành phần hoạt tính cộng đốm (LLE). Thú vị là độ nhiều bằng 18:2 giảm ở phần hoa của dòng kích thích tạo đơn bội (như được thể hiện bằng dữ liệu lipid tế bào). Không bị ràng buộc bởi lý thuyết, có lý khi giả thuyết rằng điều này có thể gây ra hoặc góp phần vào vấn đề thụ phấn và thui phôi. Khi tác giả thêm chúng lại thông qua việc áp dụng phun LLA và LLAEF, các phân tử này có thể ức chế phospholipaza dẫn đến HIR cao hơn, và bổ sung 18:2 dẫn đến nhiều hạt hơn.

Cuối cùng, dữ liệu tương tự được gom đối với MALFP, và các kết quả tương tự được thấy, bao gồm sự tăng cường về tỷ lệ kích thích tạo đơn bội bằng 1,9%, tăng số thể đơn bội được tái tạo mỗi bông bằng 6,6 và số tổng số hạt bằng 28, cũng như giảm tỷ lệ thui phôi bằng 2,2% (Bảng 12). Các hợp chất này, cũng như lipid khác, có thể được dùng để tăng cường tần suất hình thành phôi đơn bội trong vườn ươm kích thích tạo đơn bội, và còn có thể được dùng để làm tăng bộ hạt trong bối cảnh dòng kích thích tạo đơn bội tự thụ phấn để làm tăng nhiều hơn hạt dự trữ của các dòng đó.

Bảng 9. MALFP tác động lên HIR, bộ hạt, sự thui hạt, và toàn bộ các thể đơn bội tái sinh.

Thử nghiệm số	Nồng độ MALFP	Chất đốm	So sánh MALFP với đối chứng			
			Δ Các thể đơn bội	Δ HIR	Δ hạt	Δ k. a. tỷ lệ
1	2%	chế phẩm 91	+ 37	+ 16,3%	- 52	+ 16,3%
2	2%	chế phẩm 91	+ 25	+ 9,4%	+ 101	- 4,4%
3	2%	chế phẩm 92	- 9	- 6,2%	+ 68	- 3,4%
4	2%	chế phẩm 92	+ 71	+ 29,8%	+ 86	- 21,0%
5	10%	chế phẩm 92	- 20	- 7,9%	- 38	- 2,4%
6	14%	chế phẩm 91	+ 2	- 0,5%	- 18	+ 19,9%

7	14%	chế phẩm 91	+ 11	+ 3,5%	+ 70	- 19,2%
8	20%	chế phẩm 92	- 9	- 6,0%	+ 89	- 4,0%
9	20%	chế phẩm 92	- 9	- 7,6%	+ 69	-4,3%

<i>AVERAGE / EAR</i>	các thể đơn				k. a. tỷ lệ
	bội	HIR	hạt		
MAFP	34,6	17,1%	238,8		32,4%
Đối chứng	22,8	13,5%	197,3		33,5%
MAFP với đối chứng	+ 11,8	+ 3,6%	+ 41,5		- 1,1%

Bảng 10. LLAE tác động lên HIR, bộ hạt, sự thu hạt, và toàn bộ các thể đơn
bội được tái sinh.

Thử nghiệm số	Nồng độ LLAE	Chất đệm	So sánh LLAE với đối chứng			
			Δ Các thể đơn bội	Δ HIR	Δ hạt	Δ k. a. tỷ lệ
1	2%	chế phẩm 92	+ 7	+ 11,3%	- 29	- 2,4%
2	5%	chế phẩm 92	+ 3	+ 1,7%	+ 12	+ 1,8%
3	20%	chế phẩm 92	+ 10	+ 1,7%	+ 117	- 14,5%
4	0,3 mg/mL	PBS + 1%DMSO	+ 19	+ 6,5%	+ 97	+ 11,1%
5	0,3 mg/mL	PBS + 1%DMSO	+ 4	+ 1,8%	+ 14	- 8,2%
6	0,3 mg/mL	PBS + 1%DMSO	- 5	+ 11,7%	- 117	- 11,0%
7	1 mg/mL	PBS + 1%DMSO	+ 20	+ 6,8%	+ 44	- 5,3%
8	1 mg/mL	PBS + 1%DMSO	+ 14	+ 7,7%	+ 67	+ 0,2%
9	1 mg/mL	PBS + 1%DMSO	+ 8	- 1,7%	+ 87	+ 5,6%
10	3 mg/mL	PBS + 1%DMSO	=	- 6,7%	+ 99	- 17,0 %
11	3 mg/mL	PBS + 1%DMSO	+ 13	+ 8,8%	- 29	- 0,6%
12	3 mg/mL	PBS + 1%DMSO	- 1	- 5,2%	+ 110	- 9,8%
13	3 mg/mL	PBS + 1%DMSO	+ 20	+ 6,7%	+ 17	- 0,4%
14	3 mg/mL	PBS + 1%DMSO	- 7	+ 3,0%	+ 24	- 15,3%
15	10 mg/mL	PBS + 1%DMSO	+ 2	- 0,6%	+ 31	- 1,5%
16	10 mg/mL	PBS + 1%DMSO	+ 9	=	+ 51	- 14,2%
17	10 mg/mL	PBS + 1%DMSO	+ 4	+ 2%	- 2	- 1,0%
18	20 mg/mL	PBS + 1%DMSO	+ 4	+ 0,9%	+ 21	+ 4,7%
19	30 mg/mL	PBS + 1%DMSO	+ 8	- 1,1%	+ 90	- 6,5%

20	30 mg/mL	PBS + 1%DMSO	+ 5	+ 6,6%	- 64	+ 10,7%
21	30 mg/mL	PBS + 1%DMSO	+ 22	+ 0,8%	+ 165	- 6,2%

<i>TRUNG BÌNH / BÔNG</i>	các thể đơn bội	HIR	hạt	k. a. tỷ lệ
LLAEE	23,1	13,0%	194,1	29,9%
Đối chứng	14,9	10,6%	150,8	33,6%
LLAEE với đối chứng	+ 8,2	+ 2,4%	+ 43,3	- 3,7%

Bảng 11. LLA tác động lên HIR, bộ hạt, thui hạt, và toàn bộ các thể đơn bội được tái tạo.

Thứ nghiệm số	Nồng độ LLA	Chất đệm	So sánh axit linoleic (LLA) với đối chứng			
			Δ Các thể đơn bội	Δ HIR	Δ hạt	Δ k. a. tỷ lệ
1	7%	Chế phẩm 91	- 9	- 2,7%	- 9	+ 7,6%
2	0,1 mg/mL	PBS + 1%DMSO	+ 26	+ 2,2%	+ 134	+ 14,3%
3	0,1 mg/mL	PBS + 1%DMSO	- 18	+ 10,2%	- 122	+ 9,2%
4	0,1 mg/mL	PBS + 1%DMSO	+ 16	+ 8,0%	+ 8	+ 15,4%
5	0,5 mg/mL	PBS + 1%DMSO	+ 12	+ 0,2%	+ 108	- 9,4%
6	0,5 mg/mL	PBS + 1%DMSO	+ 9	+ 3,2%	+ 19	- 3,4%
7	0,5 mg/mL	PBS + 1%DMSO	- 3	- 0,5%	- 14	+ 2,3%
8	1 mg/mL	PBS + 1%DMSO	+14	+ 2,2%	+ 132	- 15,6%
9	1 mg/mL	PBS + 1%DMSO	+ 9	+ 1,7%	+ 87	- 0,7%
10	5 mg/mL	PBS + 1%DMSO	+ 23	+ 4,9%	+ 112	- 6,9%
11	5 mg/mL	PBS + 1%DMSO	- 1	- 3,1%	+ 53	- 5,5%
12	5 mg/mL	PBS + 1%DMSO	- 4	- 0,8%	- 18	+ 3,7%
13	5 mg/mL	PBS + 1%DMSO	+ 10	+ 1,6%	+ 61	- 3,0%
14	10 mg/mL	PBS + 30% DML	+ 24	+ 3,2%	+ 80	- 7,5%
15	10 mg/mL	PBS + 30% DML	+ 9	+ 1,4%	+ 44	+ 2,2%
16	10 mg/mL	PBS + 30% DML	- 2	- 3,1%	+ 32	- 0,8%
17	10 mg/mL	PBS + 30% DML	- 5	- 2,6%	- 49	- 0,2%
18	10 mg/mL	PBS + 30% DML	+ 12	+ 9,1%	+ 87	- 6,0%
19	25 mg/mL	PBS + 1%DMSO	- 15	+ 3,7%	- 92	- 5,1%
20	25 mg/mL	PBS + 1%DMSO	+ 3	+ 6,0%	- 59	+ 0,7%
21	25 mg/mL	PBS + 1%DMSO	+ 21	+ 2,5%	+ 97	- 11,8%

22	25 mg/mL	PBS + 50% DML	+ 20	+ 10,7%	+ 107	- 11,7%
		<i>AVERAGE / EAR</i>	các thể đơn bội	HIR	hạt	k. a. tỷ lệ
		LLAEE	27,9	15,4%	194,6	33,0%
		Đối chứng	20,3	13,9%	156,3	35,9%
		LLAEE với đối chứng	+ 7,6	+ 1,5%	+ 41,7	- 2,9%

Bảng 12. MAFP tác động lên HIR, bộ hạt, thui hạt, và toàn bộ các thể đơn bội được tái tạo.

Thử nghiệm số	Nồng độ MAFP	Chất đệm	<i>So sánh MAFP với đối chứng</i>			
			Δ Các thể đơn bội	Δ HIR	Δ hạt	Δ k. a. tỷ lệ
1	20 ug/mL	PBS + 1%DMSO	+ 2	+ 9,3%	- 20	+ 7,4%
2	20 ug/mL	PBS + 1%DMSO	- 4	+ 0,6%	- 27	- 5,9%
3	20 ug/mL	PBS + 1%DMSO	+ 4	- 1,9%	+ 65	+ 1,7%
4	20 ug/mL	PBS + 1%DMSO	=	- 1,4%	+ 33	- 0,6%
5	50 ug/mL	PBS + 1%DMSO	+ 16	+ 6,8%	+ 38	- 4,4%
6	50 ug/mL	PBS + 1%DMSO	+ 12	+ 4,6%	+ 60	- 8,8%
7	50 ug/mL	PBS + 1%DMSO	+ 12	+ 4,8%	+ 40	- 2,0%
8	50 ug/mL	PBS + 1%DMSO	+ 17	+ 10,5%	+ 18	- 7,7%
9	50 ug/mL	PBS + 1%DMSO	- 8	- 1,6%	- 35	+ 6,4%
10	50 ug/mL	PBS + 1%DMSO	=	+ 9,1%	- 86	+ 6,1%
11	2%	chế phẩm 92	+ 8	=	+ 109	- 1,2%
12	2%	chế phẩm 92	- 1	- 2,4%	+ 26	- 10,9%
13	10%	chế phẩm 92	+ 8	+ 0,8%	+ 44	+ 0,7%
14	10%	chế phẩm 92	- 5	- 6,9%	- 7	- 2,1%
15	20%	chế phẩm 92	- 10	- 0,9%	+ 90	- 6,7%
16	20%	chế phẩm 92	=	- 0,8%	+ 10	- 5,8%
		<i>TRUNG BÌNH / BỔNG</i>	các thể đơn bội	HIR	hạt	k. a. tỷ lệ
		MAFP	23,6	14,6%	172,5	34,6%
		đối chứng	0	12,7%	144,4	36,8%
		MAFP với đối chứng	+ 6,6	+ 1,9%	+ 28,1	- 2,2%

V. Gây đột biến/bất hoạt PLA

Trong nỗ lực để thay đổi tỷ lệ kích thích tạo đơn bội hoặc giảm tỷ lệ thụ thai trong khi lai kích thích tạo đơn bội, tác giả đã tạo ra hoặc thu được một vài dòng đột biến bằng một vài phương pháp, bao gồm dòng GM ARNi, dòng TILLING, dòng CRISPR, và dòng TALEN. Thứ nhất, tác giả tìm kiếm bằng chứng rằng đột biến nhắm đích của pPLAII α là chiến lược sống còn để tạo ra dòng kích thích tạo đơn bội mới. Do đó, tác giả đã thử cả chiến lược đột biến nhắm đích ngô CRISPR/CAS9 và TALEN nhắm đến cùng trình tự mà chứa dịch khung trong alen dòng kích thích tạo đơn bội đột biến. Điều này dẫn đến thế hệ dòng với đột biến mới, mà tác giả thử nghiệm về kích thích tạo đơn bội.

Có ba thành phần chính cho quy trình CRISPR. Xem Sáng chế Mỹ số 8,697,359 B1, được kết hợp vào đây bằng cách viện dẫn toàn bộ. Thành phần chính thứ nhất là trình tự đích. Thứ hai là Cas9, mà là endonucleaza. Thành phần chính thứ ba là ARN dẫn ("gARN"), mà bổ sung cho trình tự đích và chịu trách nhiệm tuyển chọn Cas9 cho vị trí mong muốn. Trình tự đích có chiều dài 18 đến 20bp, và tùy ý nên nằm ngay đầu 5' đến mô típ liền kề protospacer ("PAM") trong bộ gen của cây. Đối với Cas9 từ *Streptococcus pyogenes*, trình tự PAM nên là 5'-NGG-3'. Sự phiên mã của gARN có thể được dẫn bởi gen khởi động Pol III U3 (ARN bắt đầu bằng A) hoặc U6 (ARN bắt đầu bằng G). gARN nên mang trình tự đích ở đầu 5' ngay sau A (U3) hoặc G (U6). Cas9 sẽ tạo ra đoạn gãy hai đầu ("DSB") ở ba cặp bazơ 5' trình tự đích đến trình tự PAM. Trình tự axit amin của Cas9 là giống như Cas9 từ chủng SF370 *Streptococcus pyogenes*, với hai thay đổi axit amin, L1164V và I1179V trong vùng PI (1099–1368) trong thùy NUC. Hoạt tính Cas9 đã được chứng minh trong thử nghiệm biến nạp là có xấp xỉ 90% tần suất đột biến của trình tự đích được thử ở ngô. Nói chung, nên xác định nhiều PAM ứng viên và trình tự đích trong vùng đích, sau đó tìm kiếm một cái tốt nhất bằng cách tìm trình tự nào là duy nhất trong bộ gen của đích. Cây đích là ngô, lúa, hoặc cây một lá mầm bất kỳ.

Chiến lược này được theo đuổi để xác định trình tự đích CRISPR mà xếp chồng với đột biến dịch khung hiện có. Vị trí cắt chính xác là ngay hai cặp bazơ cách điểm chèn trong dịch khung. Các cấu trúc chứa cả Cas9 và gARN được biến nạp vào cây ngô. Nói chung, hai alen hoặc cây đột biến đồng nhất được tái tạo từ các sự kiện tạo ra, nhưng cây đột biến dị hợp tử cũng hữu dụng. Cây dị hợp tử tự lai, sau đó hạt T1 được trồng,

sàng lọc về tính đồng nhất của đột biến, và lai xa. Dòng biến nạp đồng nhất hoặc đột biến hai alen T0 được tự lai đơn giản và lai xa để không biến nạp NP2222. Tất cả phôi lai xa được phân lập để phân tích mức bội thể để tìm các thể đơn bội.

Ba cấu trúc đột biến nhắm đích khác nhau được tạo ra: CRISPR/CAS9 I, CRISPR/CAS9 II, và TALEN. Sự khác nhau giữa CRISPR/CAS9 I và II là nhỏ. Địa điểm vị trí đích cho tất cả ba cấu trúc là cùng vùng mà dịch khung thấy ở dòng kích thích tạo đơn bội. Đối với cấu trúc CRISPR, trình tự ARN dẫn bắt đầu ở nucleotit +1560: -GTCAACGTGGAGACAGGG- (tức là, SEQ ID NO: 20). Vị trí -AGG- PAM của SEQ ID NO: 20 gạch dưới và viết nghiêng. Chèn bốn cặp bazơ trong dòng kích thích tạo đơn bội là ở vị trí chính xác, ở nucleotit +1576. Sau khi biến nạp, một vài sự kiện CRISPR I khác nhau (bao gồm cấu trúc biểu hiện thấy ở SEQ ID NO: 34), sự kiện CRISPR II (bao gồm cấu trúc biểu hiện thấy ở SEQ ID NO: 36), và sự kiện TALEN (bao gồm cấu trúc biểu hiện thấy ở SEQ ID NO: 35) được chọn, được nuôi đến khi trưởng thành, và hình thành hạt sống. Trong thể hệ T0, tác giả thực hiện PCR ở vị trí đích và tạo trình tự sản phẩm PCR sau khi tách dòng phụ. Tác giả nhận diện nhiều đột biến duy nhất trong số các sự kiện này (và nhiều trong số các sự kiện là thể khảm hoặc nhiều alen).

Nhiều cây là khảm, như được thể hiện bởi nhiều trình tự khác nhau xuất hiện trong thể hệ T1. Sau khi tự thụ phấn T0, cây T1 được tách ra 1:2:1 cho cấu trúc đột biến đích, và nhiều cây có đột biến mới ở địa điểm đích hoặc là trong hai alen hoặc trong tình trạng đồng nhất. Tác giả sàng lọc cây con ở mức ADN sử dụng gen đánh dấu TAQMAN, nhận diện thể hai alen không có gen chuyển Cas9 hoặc TALEN, và thực hiện sắp trình tự PCR để tạo ra sản phẩm PCR đọc cặp bazơ +1494 to +1691 trong trình tự gen GRMZM2G471240. Tác giả sau đó thử nghiệm đột biến đồng nhất về khả năng kích thích tạo đơn bội. Xem SEQ ID NOs: 9–19 & 42–44 cho các trình tự của cây T1 mới ở gen *mtl*.

HIR được đo cho các dòng mới được giả định. Xem Bảng 4, ở trên. Dữ liệu HIR này là từ việc lai trong đó đực là dòng kích thích tạo đơn bội giả định và cái là dòng biến nạp nội dòng tiêu chuẩn NP2222 của tác giả. Dòng kích thích tạo đơn bội giả định được tạo ra sử dụng hoặc là đột biến nhắm đích trung gian TALEN- hoặc CRISPR/CAS9 của địa điểm pPLAIIa. Trong số các loại được thể hiện ở đây, có mười một cây dòng kích thích tạo đơn bội khác nhau bao gồm tám sự kiện khác nhau từ ba cấu trúc biến nạp riêng biệt. Sự kiện 39A là sự kiện TALEN. Sự kiện 18A và 27A là sự kiện CRISPR. Sự

iện sau là thể khảm như cây T0, và sau khi nó tự thụ phấn, nhiều đột biến được thấy ở quần thể T1, bao gồm cây "hai alen" (thuật ngữ hai alen, nghĩa là khi tác giả tạo trình tự vùng của pPLAII α mà được đột biến, tác giả thấy hai alen mới khác nhau— sao cho rõ ràng là cả hai bản sao kiểu dại của gen đã được đột biến, nhưng chúng được đột biến khác nhau, do đó có hai alen mới). Mỗi trong số mười một cây cá thể này do đó có kết hợp đột biến riêng biệt trong pPLAII α . Tất cả những gì chúng có thông thường là không có cây nào trong số mười một cây có bản sao kiểu dại của pPLAII α . Do đó, các cây này đều là "đột biến đồng nhất" đối với gen pPLAII α . Đột biến đều là dịch khung trong exon 4, bắt chước đột biến gốc trong dòng kích thích tạo đơn bội bẩm sinh. Sử dụng năm cây này làm cây đực, tác giả lai trên một hoặc một vài bông cái, tạo ra hàng nghìn phôi. Tác giả cắt ra và làm phân tích mức bội thể trên các thể hệ con đó và thấy rằng mỗi bộ thể hệ con có ít nhất 3,98% thể đơn bội với tối đa 12,5% các thể đơn bội. Điều này chứng minh rằng việc tạo ra các đột biến trong pPLAII α sẽ dẫn đến kích thích tạo đơn bội. Tác giả nghĩ rằng các kiểu đột biến khác, ngoài dịch khung, cũng sẽ dẫn đến kích thích tạo đơn bội. Các đột biến này có thể ở bất kỳ đâu trong gen, và chúng có thể là đột biến điểm hoặc chèn hoặc loại bỏ hoặc loại đột biến khác.

ARNi cũng được dùng để tạo ra dòng kích thích tạo đơn bội. Đối với ARNi, hai cấu trúc kẹp tóc được thực hiện; một cấu trúc tạo bản đồ đến ranh giới giữa exon 1 và 2, và cấu trúc còn lại tạo bản đồ đến exon 4 (các Hình 8 và 9). Các cấu trúc được biến nạp vào kiểu dại và cây T0 được tự thụ phấn. Hạt T1 từ ba sự kiện mỗi cấu trúc được trồng, sàng lọc về tính đồng nhất của gen chuyển, và lai xa trên một vài bông như thử nghiệm đối với kích thích tạo đơn bội. Sau khi kiểm tra trên 1500 hạt từ việc lai xa này, tác giả tìm thấy cả sự kiện kích thích tạo các thể đơn bội ở tỷ lệ xấp xỉ 1% đến 2%. Tỷ lệ kích thích tạo đơn bội cao nhất thu được trên bông đơn là 4,3%. bông này có khoảng 300 hạt, do đó tác giả có thể kết luận rằng tỷ lệ thui phôi cũng thấp hơn dòng kích thích tạo cao thông thường. Việc làm này chứng minh rằng chiến lược ARNi + GM có thể dùng để tạo ra dòng kích thích tạo đơn bội mới trong dòng kiểu dại thông thường khác bằng cách thay đổi biểu hiện của pPLAII α .

Phương pháp đột biến gen TILLING cũng được dùng để tạo ra và nhận diện đột biến phospholipaza và ngô theo sáng chế. Các công bố mô tả TILLING hiện có cho cây trồng như lúa: Till và cộng sự, *BMC Plant Biology* 7:19 (2007), cà chua: Rigola và cộng sự PLOS ONE 13 tháng ba 2009, và ngô: Till và cộng sự *BMC Plant Biol.* 28 tháng sáu

2004;4:12 (2004), tất cả tài liệu này được kết hợp vào đây bằng cách viện dẫn. Theo hệ phương pháp TILLING cơ bản, vật liệu cây, như hạt, chịu đột biến gen hóa học, mà tạo ra hàng loạt đột biến trong bộ gen của tế bào hạt. Các hạt được đột biến gen được trồng thành cây M1 trưởng thành và tự thụ phấn. Các mẫu ADN từ cây M2 tạo được trộn lẫn và sau đó được sàng lọc đột biến trong gen có lợi. Khi đột biến được nhận diện trong gen có lợi, hạt của cây M2 mang đột biến đó được nuôi lớn thành cây trưởng thành M3 và sàng lọc về đặc tính kiểu hình kết hợp với gen có lợi.

Bất kỳ cây trồng của ngô có ít nhất một gen phospholipaza với độ tương đồng cơ bản với SEQ ID NO: 1 có thể được dùng theo sáng chế. Như được sử dụng ở đây, "cơ bản tương đồng" nghĩa là trình tự ADN của gen đủ tương tự với SEQ ID NO: 1 ở mức nucleotit để mã hóa protein tương đương như SEQ ID NO: 1, cho phép sự khác nhau về alen giữa các cây trồng. Theo một khía cạnh của phương án ví dụ của sáng chế, "cơ bản tương đồng" có thể có mặt khi độ tương đồng giữa gen phospholipaza và SEQ ID NO: 1 là thấp khoảng 85%, với điều kiện là độ tương đồng trong vùng được bảo tồn của gen cao hơn (ví dụ, ít nhất khoảng 90%). Tốt hơn là, phần trăm tương đồng trong vùng mã hóa là 85-90%, tốt hơn nữa là 90-95%, và tối ưu, là trên 95%. Chuyên gia trong lĩnh vực có thể ưu tiên cây trồng ngô có tính phổ biến trên thị trường hoặc một loài có các đặc tính mong muốn cụ thể trong đó để tạo ra ngô đột biến phospholipaza. Tùy ý, chuyên gia trong lĩnh vực có thể ưu tiên cây trồng ngô có vài đa hình, như cây trồng nội dòng, để tạo điều kiện sàng lọc đột biến trong địa điểm phospholipaza.

Theo một khía cạnh của phương án ví dụ của sáng chế, các hạt từ lúa và ngô được đột biến gen và sau đó nuôi trồng thành cây M1. Cây M1 sau đó được để tự thụ phấn và hạt từ cây M1 được nuôi trồng thành cây M2, mà sau đó được sàng lọc đột biến ở địa điểm phospholipaza của chúng. Trong khi cây M1 có thể được sàng lọc đột biến, lợi ích của việc sàng lọc cây M2 là tất cả đột biến sinh dưỡng tương ứng với đột biến dòng tế bào sinh dục. Chuyên gia trong lĩnh vực hiểu rằng nhiều vật liệu ngô khác nhau, bao gồm, nhưng không giới hạn ở, hạt, phấn hoa, mô cây hoặc tế bào cây, có thể được đột biến gen để tạo ra ngô đột biến phospholipaza theo sáng chế. Tuy nhiên, loại vật liệu cây được đột biến có thể ảnh hưởng khi ADN của cây được sàng lọc đột biến. Ví dụ, khi phấn hoa được gây đột biến trước quá trình thụ phấn của cây không đột biến, hạt thu được từ sự thụ phấn đó được nuôi trồng thành cây M1. Mỗi tế bào của cây M1 sẽ chứa

đột biến được tạo ra ở phần hoa, do đó cây M1 này có thể sau đó được sàng lọc đột biến phospholipaza thay vì dời đến thế hệ M2.

Các tác nhân đột biến mà tạo ra đột biến điểm căn bản và xóa bỏ ngắn, chèn, chuyển đổi, và hoặc chuyển tiếp (khoảng 1 đến khoảng 5 nucleotit), như tác nhân đột biến hóa học hoặc bức xạ, có thể được sử dụng để tạo ra các đột biến theo sáng chế. Các tác nhân đột biến xác nhận với phương pháp theo sáng chế bao gồm, nhưng không giới hạn ở, etyl metansulfonat (EMS), metylmetan sulfonat (MMS), N-etyl-N-nitrosurê (ENU), triethylmelamin (TEM), N-metyl-N-nitrosourê (MNU), procarbazin, chlorambuxil, xyclophosphamit, dietyl sulfat, acrylamit monome, melphalan, nitơ mù tạc, vincristin, dimetylnitosamin, N-metyl-N'-nitro-Nitrosoguanidin (MNNG), nitrosoguanidin, 2-aminopurin, 7,12 dimetyl-benz(a)anthracen (DMBA), etylen oxit, hexametylphosphoramit, bisulfan, diepoxyalkan (diepoxyoctan (DEO), diepoxybutan (BEB), và tương tự), 2-metoxi-6-clo-9[3-(etyl-2-clo-etyl)aminopropylamino]acridin dihydroclorua (ICR-170), và formaldehyt. Đột biến tự phát trong vùng tạo hạch nhân ("NOR") mà không thể gây trực tiếp bằng tác nhân gây đột biến có thể còn được nhận diện theo nhiều phương án theo sáng chế.

Phương pháp thích hợp bất kỳ của việc điều chế ADN của cây hiện đã biết hoặc được sáng chế sau đây có thể được sử dụng để điều chế ADN của ngô để sàng lọc đột biến phospholipaza. Ví dụ, xem Chen và Ronald, *Plant Molecular Biology Reporter* 17:53-57, 1999; Stewart và Via, *Bio Techniques* 14:748-749, 1993. Hơn nữa, một vài bộ kit bán trên thị trường hiện có, bao gồm bộ kit từ Qiagen (Valencia, Calif.) và Qbiogene (Carlsbad, Calif.).

Theo một khía cạnh của phương án ví dụ của sáng chế, các mẫu ADN từ cá thể cây ngô được điều chế và sau đó được trộn lẫn để thúc đẩy việc sàng lọc đột biến ở phospholipaza của toàn bộ quần thể của cây có nguồn gốc từ mô của cây đột biến. Kích thước của nhóm trộn lẫn có thể là phụ thuộc vào độ nhạy của phương pháp sàng lọc được sử dụng. Theo một khía cạnh của phương án ví dụ của sáng chế, nhóm gồm bốn hoặc nhiều cá thể cây ngô được trộn lẫn.

Theo khía cạnh khác của phương án ví dụ, sau khi mẫu ADN được trộn lẫn, các hỗn hợp trộn lẫn được đưa qua kỹ thuật khuếch đại đặc hiệu trình tự phospholipaza, như phản ứng chuỗi polymeraza (PCR). Đối với tổng quan chung của PCR, xem nguyên tắc

PCR: A Guide to Methods and Applications (Innis, Gelfand, Sninsky, J., và White, eds.), Academic Press, San Diego, 1990, mà được kết hợp vào đây bằng cách viện dẫn. Bất kỳ đoạn mồi đặc hiệu cho địa điểm phospholipaza hoặc trình tự liền ngay với địa điểm phospholipaza có thể được sử dụng để khuếch đại trình tự phospholipaza trong mẫu ADN trộn lẫn. Tốt hơn là, đoạn mồi được thiết kế để khuếch đại các vùng địa điểm phospholipaza ở chỗ các đột biến hữu ích gần như có thể xảy ra. Tốt nhất là, đoạn mồi được thiết kế để phát hiện các đột biến trong vùng mã hóa của gen phospholipaza. Hơn nữa, tốt hơn là đối với đoạn mồi để tránh các vị trí đa hình đã biết để sàng lọc đột biến điểm dễ dàng. Để tạo điều kiện phát hiện sản phẩm PCR trên gel, đoạn mồi PCR có thể được dán nhãn sử dụng bất kỳ phương pháp đã biết hoặc phương pháp dán nhãn được sáng chế sau đây.

Theo một khía cạnh của phương án ví dụ của sáng chế, sản phẩm khuếch đại PCR có thể được sàng lọc đột biến phospholipaza sử dụng bất kỳ phương pháp mà nhận diện sự khác biệt nucleotit giữa trình tự kiểu dại và đột biến. Các phương pháp này có thể bao gồm, nhưng không giới hạn ở, tạo trình tự, sắc ký lỏng cao áp biến tính (dHPLC), điện di mao quản biến tính không đổi (CDCE), điện di mao quản gradien nhiệt độ (TGCE) (xem Li và cộng sự, *Electrophoresis* 23(10):1499-1511, 2002), hoặc bằng cách phân đoạn sử dụng phân cắt enzym, như được dùng trong phương pháp lưu lượng cao được mô tả bởi Colbert và cộng sự, *Plant Physiology* 126:480-484, 2001. Tốt hơn là, sản phẩm khuếch đại PCR được ủ với endonucleaza mà phân cắt ưu tiên sự không trùng khớp trong chuỗi kép dị hợp giữa trình tự kiểu dại và đột biến. Theo khía cạnh khác của phương án ví dụ, sản phẩm phân cắt được điện di sử dụng thiết bị gel tạo trình tự tự động hóa, và hình ảnh gel được phân tích với sự trợ giúp của chương trình xử lý hình ảnh thương mại tiêu chuẩn.

Tác giả sáng chế xác định rằng để thu được kích thích tạo đơn bội ở ngô, đột biến mà thay đổi chức năng phospholipaza là được mong muốn. Các đột biến bao gồm sai nghĩa, vô nghĩa và thay đổi mối nối, bao gồm đột biến mà cắt đứt sớm sự dịch mã của protein phospholipaza từ ARN thông tin, như các đột biến mà tạo ra bộ ba mã hóa kết thúc trong vùng mã hóa của gen phospholipaza. Các đột biến này bao gồm chèn, trình tự lặp, khung đọc mở biến đổi (ORFs) và, tốt nhất là, đột biến điểm.

Theo khía cạnh khác nữa của phương án ví dụ theo sáng chế, khi cây M2 có trình tự đột biến phospholipaza được nhận diện, đột biến được phân tích để xác định tác động

của nó lên biểu hiệu, dịch mã, và/hoặc hoạt tính của protein. Theo một phương án ví dụ, đoạn phospholipaza chứa đột biến được tạo trình tự, sử dụng kỹ thuật tạo trình tự tiêu chuẩn, để xác định vị trí chính xác của đột biến liên quan đến tổng thể trình tự phospholipaza. Mỗi đột biến được đánh giá để dự báo tác động của nó lên chức năng protein (tức là, khả năng chịu hoàn toàn mất chức năng) sử dụng dụng cụ thông tin sinh học như SIFT (Sorting Intolerant từ Tolerant; Ng và cộng sự, *Nucleic Acids Research* 31:3812-3814, 2003), PSSM (Position-Specific Scoring Matrix; Henikoff và Henikoff, *Computer Applications in the Biosciences* 12:135-143, 1996) và PARSESNP (Taylor và Greene, *Nucleic Acids Research* 31:3808-3811, 2003). Ví dụ, điểm SIFT mà ít hơn 0,05 và thay đổi lớn về điểm PSSM (ví dụ, khoảng 10 hoặc hơn) chỉ đột biến mà có khả năng có tác động có hại tới chức năng protein.

Theo khía cạnh khác của phương án ví dụ, nếu đánh giá ban đầu của đột biến ở cây M2 chỉ ra nó là tự nhiên hữu dụng và ở vị trí hữu dụng trong gen phospholipaza, thì các phân tích kiểu hình khác của cây ngô chứa đột biến này được theo đuổi. Đầu tiên, cây M2 được lai ngược hoặc lai xa hai lần để tạo ra cây BC1 để loại trừ đột biến nền. Sau đó cây BC1 được lai ngược hoặc lai xa được tự thụ phấn để tạo ra cây BC1F2 mà đồng nhất cho đột biến phospholipaza.

Một vài đặc tính vật lý của cây đột biến phospholipaza đồng nhất này được đánh giá để xác định xem đột biến tạo ra thay đổi kiểu hình hữu dụng ở ngô hay không. Ngô phospholipaza đột biến được đánh giá về kích thích tạo đơn bội so với ngô bố mẹ thường (ví dụ, kiểu đại) hoặc so với ngô đối chứng anh em kiểu đại. Bảng 13 thể hiện đột biến mới thu được bởi TILLING.

Bảng 13. Đột biến pPLAII α mới thu được bởi TILLING & HIR của nó. Cột thay đổi nucleotit thể hiện vị trí từ bắt đầu của trình tự cADN (TRÌNH TỰ SỐ 1), và nucleotit được thay đổi được viết in hoa trong bối cảnh bộ ba mã của nó. Thay đổi axit amin sau đó được chỉ ra tiếp đó tác động của thay đổi này (Hấp thu hoặc Không hấp thu). Hai alen mà không hấp thu, một chứa các thể đơn bội với tỷ lệ bằng 1,04% (3/288).

Dòng	Thay đổi nucleotit	Exon	Thay đổi AA	Hấp thu?	Các thể		PA được xác nhận	HIR
					lượng bội	Các thể đơn bội		
1139	bp+128 tGt / tAt	1	C13Y	Có	389	0	0	0,00%

3594	bp+167	cCc / cTc	1	P26L	Có	381	0	0	0,00%
0505	bp+431	cCg/ cTg	1	P114L	Không	235	0	0	0,00%
2658	bp+718	Gcg / Acg	4	A237T	Có	379	0	0	0,00%
1983	bp+1077	atG / atA	4	M356I	Không	285	3	3	1,04%
2732	bp+1163	aCt / aTt	4	T385I	Có	383	0	0	0,00%
2414	bp+1226	aGa / aAa	4	R406K	Có	392	0	0	0,00%

Danh pháp dùng trong Bảng 13 chỉ nucleotit hoặc axit amin kiểu đại, sau đó vị trí của nó theo SEQ ID NO1 viện dẫn, sau đó nucleotit hoặc axit amin được thay đổi ở vị trí sử dụng hệ thuật ngữ mã hóa di truyền tiêu chuẩn.

Đối với ngô, gen pPLAII α của ngô TILLING tạo ra alen mới mà có tỷ lệ kích thích tạo đơn bội thấp. Điều này cho phép tạo ra chuỗi alen, bao gồm bất hoạt, của GRMZM2G471240. Trình tự của hai đoạn của gen này (tối đa 1,5 kb, mà bằng 20 đơn vị siêu sao chép mỗi gen) được sàng lọc cho đột biến. Các trình tự này gồm trình tự bộ gen bao gồm intron, cộng với trình tự cADN dự báo và trình tự mã hóa đối với hai biến thể nổi. trình tự tăng và đơn vị siêu sao chép đơn được thiết kế dựa trên các trình tự đó, và sàng lọc đột biến được thực hiện trong quần thể ngô M2 khối có sẵn. Các đột biến được xác định khác biệt về trình tự ADN và kết quả trên trình tự protein dịch mã. Hạt M3 được trồng và tự lai và để tạo ra dòng M4 kèm theo tách các các thể đồng nhất đột biến giả định. Các cá thể này được xác định bằng cách sắp trình tự PCR và lai xa và tự lai để thử khả năng tạo ra đơn bội của dòng đột biến.

Để thực hiện lai thử nghiệm, các dòng mới được trồng dọc theo dòng gen đánh dấu mà lặn đồng nhất đối với gen đánh dấu màu không gây chết người. Lai qua lại được dùng để thử tính đặc hiệu của kích thích tạo chuyển đổi đực với cái bằng cách đánh giá các cây tạo ra về các thể đơn bội, mà biểu hiện kiểu hình màu sắc. Kết quả dương được xác nhận bằng phân tích mức bội thể như được mô tả ở trên.

Các cá thể mà đồng nhất cho đột biến SNP được lai là đực với dòng gen đánh dấu cái và dẫn đến hình thành tỷ lệ thể đơn bội thấp trong một số trường hợp (xem Bảng 18). Kết quả dương được xác nhận bằng phân tích mức bội thể như được mô tả ở trên. Cụ thể là, dòng mà dẫn đến hình thành thể đơn bội có đột biến G đến A ở cặp bazơ 1077 của trình tự cADN. Đột biến này gây ra thể axit amin của methionin (M) cho isoleuxin

(I) ở axit amin 356. Đây là thay đổi axit amin không bảo tồn mà có thể phá vỡ hoạt tính của protein dẫn đến hình thành tỷ lệ thấp các thể đơn bội. Trong số 288 thể hệ con được thử, tác giả tìm thấy ba thể đơn bội, cho tỷ lệ kích thích tạo bằng 1% (3/288).

VI. Tạo dòng kích thích tạo đơn bội ở lúa

Ở lúa, thể tương đồng gần nhất với pPLAII α của ngô là Os03g27610, phospholipaza giống patatin của lúa (*OspPLAII ϕ*) với chú thích tương tự, cấu trúc gen và mẫu biểu hiện, tức là, được biểu hiện trong phần hoa và không có mặt ở đầu khác (Hình 10). SEQ ID NO: 21 bao gồm trình tự ADN bộ gen của Os03g27610, SEQ ID NO: 22 bao gồm trình tự cADN, và SEQ ID NO: 23 bao gồm trình tự axit amin. Mỗi quan hệ chặt chẽ giữa các dấu hiệu, cùng với khoảng cách tiến hóa ngắn giữa hai loại thân cỏ này, đề xuất rằng đột biến ở gen lúa cũng có thể tạo ra sự tăng dòng kích thích tạo đơn bội. Trong công bố gần đây protein của lúa được phát hiện trong hạt tinh tử của hạt phấn hoa (Abiko và cộng sự, 2013), đề xuất sự tham gia của protein trong thụ phấn và/hoặc phát triển hợp tử.

Để tăng tỷ lệ kích thích tạo đơn bội ở ngô và tạo ra dòng kích thích tạo đơn bội thứ nhất ở lúa, nghiên cứu TILLING di truyền ngược được sử dụng để thu được đột biến mới ở gen GRMZM2G471240 của ngô và gen Os03g27610 của lúa. Xem McCallum CM và cộng sự (2000) *Targeting induced local lesions IN genomes (TILLING) for plant functional genomics*, PLANT PHYSIOL. 123: 439–42, được kết hợp vào đây bằng cách viện dẫn. TILLING đề xuất nghiên cứu để tạo ra đột biến mới do không có kiểm soát bởi các nhà nghiên cứu trong đó đột biến gen etylmetansulfonat (EMS) sẽ tạo ra đột biến mới. Sự đa dạng và sự dồi dào của alen mới được tạo ra và được thử về tỷ lệ kích thích tạo đơn bội.

Mười ba dòng TILLING M3 khác nhau được thu. Xem Bảng 14. Cột PosGenomic chỉ vị trí nucleotit của đột biến và thay đổi (ví dụ, G803A chỉ ra rằng cặp bazơ G ở vị trí 803 được đổi thành A). Tác động là thay đổi axit amin hoặc thay đổi protein khác mà tạo ra từ đột biến (ví dụ A209T chỉ ra rằng Alanin ở axit amin 209 được đổi thành Threonin). Điểm BLOSUM là dự báo độ mạnh của tác động của việc thay đổi axit amin sẽ có lên cấu hình hoặc gập của protein (càng âm, tác động càng xấu). "Loại" chỉ loại thay đổi axit amin ("NSM" nghĩa là đột biến không câm; "PSM" nghĩa là đột biến làm câm một phần; "làm câm" nghĩa là đột biến làm câm; "nổi" nghĩa là đột biến

vị trí nối tạo ra nối nằm chỗ; "intron" nghĩa là đột biến là ở intron). Cuối cùng GSOR# là dòng ID cho tập hợp Genetic Stocks – Oryza ở USDA.

Mười ba dòng này được tự lai để tạo ra M4 và hạt M4 được trồng và thử về tính đồng nhất. Các cá thể đột biến đồng nhất được tự lai và lai xa để thử khả năng kích thích tạo đơn bội. Thế hệ con tạo thành được kiểm tra hàm lượng ADN mỗi tế bào (mức bội thể) sử dụng thiết bị phân tích mức bội thể.

Các thay đổi không bảo tồn, như thay đổi vị trí nối thể và thay đổi với điểm BLOSUM âm có tiềm năng kích thích tạo đơn bội lớn nhất. Điều này có nhiều tác động bất ổn định lên sản phẩm protein, và do đó là alen TILLING kích thích tạo đơn bội cao cấp so với loại khác, làm tăng nhiều thể đơn bội mỗi sự lai kích thích tạo đơn bội và có thể tạo ra bộ hạt chuyển tiếp một phần. Thật vậy, tác giả đã bắt đầu thấy rằng trong một số trong các dòng tự thụ phấn T4. Dòng với bộ hạt thấp nhất là đột biến điểm nối G153A, với chỉ 29 hạt được tái tạo mỗi 12 cây M4 đột biến đồng nhất được lai. Các dòng khác có nhiều hơn 100 được tái tạo.

Bảng 14. Alen TILLING ở lúa Os03g27610.

Gen	PosGen	PosTIL	Hiệu	BLOSUM	Loại	GSOR#
			quả			
Os03g27610	G803A	G590A	A209T	0	NSM	406317
Os03g27610	G761A	G548A	D195N	1	NSM	405490
Os03g27610	G1163A	G950A	G293E	-2	PSM	403403
Os03g27610	G1189A	G976A	G302R	-2	PSM	406250
Os03g27610	T374C	T161C	intron	NA	intron	403453
Os03g27610	G1026A	G813A	K247=	NA	câm	406338
Os03g27610	C738T	C525T	P187L	-3	PSM	405205
Os03g27610	G1149A	G936A	Q288=	NA	câm	405898
Os03g27610	G366A	G153A	nối	NA	nối	403878
Os03g27610	G366A	G153A	nối	NA	nối	405549
Os03g27610	C792T	C579T	T205I	-1	PSM	404794

Os03g27610	A1021G	A808G	T246A	0	NSM	404534
Os03g27610	G558A	G345A	V156M	1	NSM	404675

Tùy ý, gen phospholipaza ở lúa thấy ở Os03g27610 có thể được chỉnh sửa bằng phương pháp CRISPR/Cas9. Như được nêu ở trên, có ba thành phần chính cho quy trình CRISPR. Thành phần chính thứ nhất là trình tự đích. Thứ hai là Cas9, mà là endonucleaza. Thành phần chính thứ ba là ARN dẫn ("gARN"), mà bổ sung cho trình tự đích và chịu trách nhiệm tuyển chọn Cas9 cho vị trí mong muốn. Các ARN dẫn có thể dưới dạng ARN dẫn đơn (sgRNA) hoặc ARN dẫn kép (dgRNA). Đối với lúa, tác giả tạo ra ba cấu trúc nhắm đích gen phospholipaza của lúa. SEQ ID NO: 38 bao gồm cat-xet biểu hiện mà tạo ra dgRNA nhắm đích Os03g27610, trong exon 4 gần với vị trí mà có đột biến của bốn cặp bazơ tự nhiên trong thể tương đồng của ngô. Trong gen của lúa, vị trí đích của ARN dẫn là GAGACCGGCAGGTACGTCGAGG. SEQ ID NO: 39 bao gồm cat-xet biểu hiện mà tạo ra sgARN nhắm đích Os03g27610, exon 4, ở cùng vị trí đích của gARN như nhắm đích trong SEQ ID NO: 38. Đột biến dịch khung cho cả SEQ ID NO 38 và 39 được kỳ vọng xảy ra ở thanh dọc nằm giữa G và T trong trình tự CAGGTACG | TCGAGG (ở cặp bazơ +1150 của trình tự gADN trong SEQ ID NO 22). Do đó, cả hai trong số các cấu trúc này được kỳ vọng để tạo ra các đột biến dòng kích thích tạo đơn bội mà chỉ là bảy cặp bazơ xuôi dòng từ vị trí có sự chèn dòng kích thích tạo đơn bội của ngô. Các đột biến này trong hầu hết các trường hợp sẽ là đột biến dịch khung mà kích thích tạo sự chèn hoặc loại bỏ nhỏ, ví dụ loại bỏ G hoặc T ở vị trí cắt, hoặc bất kỳ đột biến tương tự khác. SEQ ID NO: 40 bao gồm cat-xet biểu hiện mà tạo ra dgARN nhắm đích Os03g27610. SEQ ID NO: 41 bao gồm cat-xet biểu hiện mà tạo ra sgARN nhắm đích Os03g27610. Cả hai trong số các ARN dẫn lân cận mà nhắm đích trình tự CCTCGCCGATTACTTCGACTGCA trong Exon 1. Điều này có thể tạo ra sự bất hoạt phần lớn trình tự mã hóa của gen. Đột biến mà được tạo ra nên xảy ra ở vị trí cắt nơi mà thanh dọc nằm giữa C và C trong trình tự CCTCGC | CGATTAC (ở cặp bazơ +215 của trình tự cADN trong SEQ ID NO 22). Do đó cả cấu trúc 40 và 41 được kỳ vọng để tạo ra tần suất cao của cây chứa đột biến bất hoạt của gen, mà cũng nên dẫn đến tỷ lệ kích thích tạo đơn bội cao ở lúa.

Cây lúa được biến nạp với cấu trúc biến nạp bao gồm trình tự được chọn từ nhóm bao gồm SEQ ID Nos: 38-41. Qua cơ cấu CRISPR/Cas9 được mã hóa trong cấu trúc biến nạp, alen phospholipaza mới được tạo ra trong dòng biến nạp, tức là, cây lúa T0. Cây lúa T0, được trồng và lai (tức là, tự thụ phấn) để tạo ra cây T1. Cây lúa T1 được thử về tính đồng nhất ở alen phospholipaza mới. Cây lúa T1 đồng nhất được lai với dòng lúa, và tạo ra thế hệ con được thử về tính đơn bội sử dụng thiết bị phân tích mức bội thể. Các phôi đơn bội chứa không ADN T1 phát hiện được được xác định và đếm, và HIR được đo. Ít nhất một phôi đơn bội được tạo ra từ việc lai, và HIR tăng. Tốt hơn là, HIR là ít nhất 5%. Ít nhất một phôi đơn bội được xử lý với tác nhân nhân đôi nhiễm sắc thể, ví dụ colchixin, và cây đơn bội kép phát triển từ đó.

YÊU CẦU BẢO HỘ

1. Phương pháp kích thích tạo phôi đơn bội trong phép lai giữa các cây, phương pháp này bao gồm:
 - (a) biểu hiện phospholipaza giống patatin đột biến ở cây, trong đó phospholipaza giống patatin được mã hóa bởi trình tự axit nucleic của SEQ ID NO: 1, và đột biến này dẫn đến sự biểu hiện giảm của phospholipaza giống patatin; hoặc
 - (b) dùng cho cây phân tử ARN can thiệp kích thước nhỏ bao gồm 18–21 nucleotit của gen mã hóa phospholipaza giống patatin, trong đó phospholipaza giống patatin được mã hóa bởi trình tự axit nucleic của SEQ ID NO: 1; hoặc
 - (c) biến nạp cây với phospholipaza giống patatin đột biến, trong đó phospholipaza giống patatin được mã hóa bởi trình tự axit nucleic của SEQ ID NO: 1, và đột biến này dẫn đến sự biểu hiện giảm của phospholipaza giống patatin; hoặc
 - (d) gây đột biến trình tự phospholipaza giống patatin của cây sử dụng chỉnh sửa gen, trong đó phospholipaza giống patatin được mã hóa bởi trình tự axit nucleic của SEQ ID NO: 1, và đột biến này dẫn đến sự biểu hiện giảm của phospholipaza giống patatin;

trong đó cây này được dùng làm cây bố mẹ trong phép lai này, sao cho phép lai này tạo ra ít nhất một phôi đơn bội.
2. Phương pháp theo điểm 1, trong đó phospholipaza giống patatin đột biến được mã hóa bởi trình tự axit nucleic của SEQ ID NO: 3.
3. Phương pháp theo điểm 1, trong đó việc chỉnh sửa gen của bước (d) được thực hiện bằng đột biến định hướng điểm.

4. Phương pháp theo điểm 3, trong đó đột biến định hướng điểm được thực hiện bằng kỹ thuật được chọn từ nhóm bao gồm CRISPR/Cas9, TALENs, ngón tay kẽm, và meganucleaza.
5. Phương pháp theo điểm 1, trong đó cây là cây một lá mầm.
6. Phương pháp theo điểm 5, trong đó cây một lá mầm là cây ngô, cây lúa, cây lúa mỳ, hoặc cây lúa mạch.
7. Phương pháp theo điểm 1, trong đó cây là cây ngô hoặc cây lúa.
8. Phương pháp theo điểm 7, trong đó cây ngô hoặc cây lúa tạo ra phần hoa dùng trong phép lai.

DANH MỤC TRÌNH TỰ

- <110> Syngenta Participations, AG
 Kelliher, Tim
 Starr, Dakota
 Wendeborn, Sebastian
 Delzer, Brent
 Chintamanani, Satya
 Skibbe, David
 Chen, Zhongying
- <120> PHƯƠNG PHÁP KÍCH THÍCH TẠO PHÔI ĐƠN BỘI TRONG PHÉP LAI GIỮA CÁC CÂY
- <130> 80906-WO-REG-ORG-P-1
- <150> US 62/256,902
 <151> 2015-11-15
- <150> US 62/300,507
 <151> 2016-02-16
- <160> 81
- <170> PatentIn phiên bản 3.5
- <210> 1
 <211> 1371
 <212> ADN
 <213> Zea mays
- <400> 1
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgctcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgctc ggggtctcat cccgggaacc 240
 atcctcgctt tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggtggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcccgcg cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggt gtggcggcca acaaccgcg gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcaggggc tctacacggc gcggcagtgc tcccgggtgg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020

gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggagggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
 ccaagagggt ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatittta ataaaacata aaaatatata t 1371

<210> 2
 <211> 401
 <212> PRT
 <213> Zea mays

<400> 2

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 3
 <211> 1375
 <212> ADN

<213> Zea mays

<400> 3

```

agttcatcac taatcacact tattgtgccc tcgacgagta tctagctagc tcattaatcg      60
atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc      120
aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg      180
cagaggggtga cgggtctgac ggtggacggc ggcggcgctc ggggtctcat cccgggaacc      240
atcctcgcct tcctggaggc caggctgcag gagctggacg gaccggaggc gaggtggcg      300
gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc      360
gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacta cttttacatg      420
gagaactgcc cgcgcattct cctcagaag agcaggcttg cgcccgccat gtccgcgctg      480
aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag      540
acgagggcca agagcacgcc tctgaagaac gcgctgctct cggacgtgtg cattggcacg      600
tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag      660
gagcgcgaat acaacctcat cgacggcggg gtggcgggcca acaaccgac gatggttgcg      720
atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca      780
gtgaaccctg cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc      840
gagcagggcc tctacacggc gcggcagtgc tcccggtggg gcattctgcg gtgggtccgc      900
aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac      960
atccacgtcg ccgcgatggt ccagtcgctc cacagcgacg gcgactacct acgcatccag     1020
gacaactcgc tccgtggcgc cgcggaacc gtggacgcgg cgacgccgga gaacatgcgg     1080
acgctcgtcg ggatcgggga gcggatgctg gcacagcggg tgtccagggt caacgtggag     1140
acagggagcg aggtacgaac cggtgaccgg agaaggaagc aatgccgatg cctcggtg     1200
gctcgttagg cagctctccg aggagaggag aacaaggctc gcgcgcgcg tctctgccat     1260
caacccaga agctctagat gtgcgcccta cgatatctaa gacaagtggc tttactgtca     1320
atcacatgct tgtaaataag tagactttat tttataaaaa tataaaaata tatat         1375

```

<210> 4

<211> 372

<212> PRT

<213> Zea mays

<400> 4

```

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
1           5           10           15

```

```

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
20           25           30

```


Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn Tyr Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

Ser Glu Val Arg Thr Gly Asp Arg Arg Arg Lys Gln Cys Arg Cys Pro
 355 360 365

Arg Trp Ala Arg
 370

<210> 5
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi GRMZM2G471240_nil.F1

<400> 5
 gtacgccgtg cgctaaca

18

<210> 6
 <211> 20
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi GRMZM2G471240_nil.R1

<400> 6
 tcgtacctcc ctgtctccac

20

<210> 7
 <211> 19
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi GRMZM2G471240_rwk.F1

<400> 7
 tacgccgtgc gctaacata

19

<210> 8
 <211> 20
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi GRMZM2G471240_rwk.R1

<400> 8

gtacctcgct ccctgtctcc

20

<210> 9

<211> 1358

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do TALEN ở Sự kiện 39A cá thể ID 22808-3954 alen 1

<400> 9

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg	60
atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc	120
aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg	180
cagaggggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc	240
atcctcgctt tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg	300
gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc	360
gcgcccggca aggacaagcg gcctctctac gctgccaaagg acatcaacca cttttacatg	420
gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg	480
aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag	540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg	600
tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag	660
gagcgcgaat acaacctcat cgacggcggg gtggcgggca acaaccgac gatggttgcg	720
atgacgcaga tcacaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca	780
gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtgcacgtcc	840
gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgccg gtggctccgc	900
aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctgggtggac	960
atccacgtcg ccgcgatggt ccagtcgctc cacagcgacg gcgactacct gcgcatccag	1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg	1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag	1140
aaccggtgac tggcgaagga agcaatgcg atgcoctcgg tgggctcgct aggcagctct	1200
ccgaggagag gagaacaagg ctgcgcgcgc gcgtctctgc catcaacca agaggctcta	1260
gatgtgcgtc gtacgatatc taagacaagt ggctttactg tcagtcacat gottgtaaat	1320
aagtagactt tattttaata aaacataaaa atatatat	1358

<210> 10
 <211> 1366
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đột biến MTL do TALEN ở Sự kiện 23A cá thể ID 22808-3924 alen 1

<400> 10
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgctt tcctcgaggc caggtctgag gagctggacg gaccggaggc gaggtggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcagc cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcgggcca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggagaa ccggtgactg gcgaagggaag caatgccgat gccctcggtg ggctcgtag 1200
 gcagctctcc gaggagagga gaacaaggct cgcgcgcgcg gtctctgccg tcaacccaag 1260
 aggctctaga tgtgcgtcgt acgatatcta agacaagtgg ctttactgtc agtcacatgc 1320
 ttgtaaataa gtagacttta ttttaataaa acataaaaat atatat 1366

<210> 11
 <211> 1358
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đột biến MTL do TALEN ở Sự kiện 81A

<400> 11
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgctcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcacgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcgggca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtg tcccggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccatt catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 aaccggtgac tggcgaagga agcaatgccg atgccctcgg tgggctcgtc aggcagctct 1200
 ccgaggagag gagaacaagg ctgcgcgcgc gcgtctctgc catcaacca agaggctcta 1260
 gatgtgcgtc gtacgatatc taagacaagt ggctttactg tcagtcacat gcttgtaaatt 1320
 aagtagactt ttttttaata aaacataaaa atatatat 1358

<210> 12

<211> 1343

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do TALEN ở Sự kiện 39A cá thể ID 22808-3954 alen 2

<400> 12
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180

cagaggggtga cggtgctgac ggtggacggc ggcggcgctcc ggggtctcat cccgggaacc 240
 atcctcgccct tcctcgagggc caggctgcag gagctggacg gaccggagggc gaggctggcg 300
 gactacttgg actacatcgc cggaaccagc accggcggtc tcataccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcattg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcccgcg cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggt gtggcgggca acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccgggtgg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcg cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccaggg caacgtggag 1140
 acggaagcaa tgccgatgcc ctcggtgggc tcgctaggca gctctccgag gagaggagaa 1200
 caaggctcgc gcgcccgtc tctgccatca acccaagagg ctctagatgt gcgtcgtacg 1260
 atatctaaga caagtggctt tactgtcagt cacatgcttg taaataagta gactttat 1320
 taataaaaca taaaaatata tat 1343

<210> 13

<211> 1363

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do TALEN ở Sự kiện 23A ID 22808-3924 alen 2

<400> 13

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcgggc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgctcc ggggtctcat cccgggaacc 240
 atcctcgccct tcctcgagggc caggctgcag gagctggacg gaccggagggc gaggctggcg 300
 gactacttgc actacatcgc cggaaccagc accggcggtc tcataccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420

gagaactgcc cgcgcacatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcacg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcgcgcg cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgcac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtg ccccggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccac catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcacccag 1020
 gacaactcgc tccgtggcg cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acaggaaccg gtgactggcg aaggaagcaa tgccgatgcc ctcggtgggc tcgctaggca 1200
 gctctccgag gagaggagaa caaggctcgc gcgccgcgtc tctgccatca acccaagagg 1260
 ctctagatgt gcgtcgtacg atatctaaga caagtggctt tactgtcagt cacatgcttg 1320
 taaataagta gactttatct taataaaaca taaaaatata tat 1363

<210> 14

<211> 1360

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do TALEN ở Sự kiện 38A cá thể ID 22808-4108 alen 1

<400> 14

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgagge caggtgcag gagctggacg gaccggaggc gaggtggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccgcca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcacatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcacg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcgcgcg cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660

gagcgcggaat acaacctcat cgacggcggt gtggcggcc acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccgggtgg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acaaccggtg actggcgaag gaagcaatgc cgatgccctc ggtgggctcg ctaggcagct 1200
 ctccgaggag aggagaacaa ggctcgcgcg ccgcgtctct gccatcaacc caagaggctc 1260
 tagatgtgcg tcgtacgata tctaagacaa gtggctttac tgtcagtcac atgcttgtaa 1320
 ataagtagac tttattttta taaaacataa aaatatatat 1360

<210> 15

<211> 1363

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do CRISPR ở Sự kiện 18A ID 22807-4016

<400> 15

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcgga atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctc gcgagcccgt cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcgcgctcc ggggtctcat cccgggaacc 240
 atcctcgctt tctcagaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cgcccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcggaat acaacctcat cgacggcggt gtggcggcc acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccgggtgg gtatctgccg gtggctccgc 900

aacaacggca tggcccccacat catcgacatc ttcattggcgg ccagctcgga cctggtggac 960
atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
acaggaaccg gtgactggcg aaggaagcaa tgccgatgcc ctcggtgggc tcgctaggca 1200
gctctccgag gagaggagaa caaggctcgc gcgccgcgtc tctgccatca acccaagagg 1260
ctctagatgt gcgtcgtacg atatctaaga caagtggctt tactgtcagt cacatgcttg 1320
taaataagta gactttatatt taataaaaca taaaaatata tat 1363

<210> 16

<211> 1372

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do CRISPR ở Sự kiện 27A cá thể ID 22807-4801 alen 1

<400> 16

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
aatacctgta gcacgaaggc gatggccggg agcgtggctc gcgagcccg cgtgctgggg 180
cagaggggtga cgggtgctgac ggtggacggc ggccggcgtc ggggtctcat cccgggaacc 240
atcctcgcct tctcagaggc caggctgcag gagctggacg gaccggaggc gaggtggcg 300
gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgac gatggttgcg 720
atgacgcaga tcacaaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
gagcagggcc tctacacggc gcggcagtg ccccggtggg gtatctgccg gtggtccgc 900
aacaacggca tggcccccacat catcgacatc ttcattggcgg ccagctcgga cctggtggac 960
atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140

acaagggagg tacgaaccgg tgactggcga aggaagcaat gccgatgcc tcggtgggct 1200
 cgctaggcag ctctccgagg agaggagaac aaggctcgcg cgccgcgtct ctgccatcaa 1260
 cccaagaggc tctagatgtg cgtcgtacga tatctaagac aagtggcttt actgtcagtc 1320
 acatgcttgt aaataagtag actttatttt aataaaacat aaaaatatat at 1372

<210> 17

<211> 1372

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến do CRISPR ở Sự kiện 27A cá thể ID 22807-4081 alen 1

<400> 17

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccg cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgctc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcg cgaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggt gtggcgggca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtgcacgtcc 840
 gagcagggcc tctacacggc gcggcagtg tcccggtggg gtatctgccg gtgggtccgc 900
 aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acaagggagg tacgaaccgg tgactggcga aggaagcaat gccgatgcc tcggtgggct 1200
 cgctaggcag ctctccgagg agaggagaac aaggctcgcg cgccgcgtct ctgccatcaa 1260
 cccaagaggc tctagatgtg cgtcgtacga tatctaagac aagtggcttt actgtcagtc 1320
 acatgcttgt aaataagtag actttatttt aataaaacat aaaaatatat at 1372

<210> 18
 <211> 1373
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đột biến MTL do CRISPR ở Sự kiện 76A cá thể ID 22873-3999

<400> 18
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccg cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcg cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcgggca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtcg tcccggtggg gtatctgcgc gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acaaagggag gtacgaaccg gtgactggcg aaggaagcaa tgccgatgcc ctcggtgggc 1200
 tcgctaggca gctctccgag gagaggagaa caaggctcgc gcgcgcgctc totgccatca 1260
 acccaagagg ctctagatgt gcgtcgtaag atatctaaga caagtggctt tactgtcagt 1320
 cacatgcttg taaataagta gactttatct taataaaaca taaaaatata tat 1373

<210> 19
 <211> 1370
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đột biến do CRISPR ở Sự kiện 32A cá thể ID 22873-3991

<400> 19
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcgggca acaaccgac gatggttgcg 720
 atgacgcaga tcacaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtg tcccggtggg gtatctgccg gtgggtccgc 900
 aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg ggcactacct ggcattccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgcggga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acgggaggta cgaaccggtg actggcgaag gaagcaatgc cgatgccctc ggtgggctcg 1200
 ctaggcagct ctccgaggag aggagaacaa ggctcgcgcg ccgcgtgtct gccatcaacc 1260
 caagaggctc tagatgtgcg tcgtacgata tctaagacaa gtggctttac tgtcagtcac 1320
 atgcttgtaa ataagtagac tttattttta taaaacataa aaatatatat 1370

<210> 20
 <211> 18
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> trình tự nucleotit mã hóa ARN dẫn

<400> 20
 gtcaacgtgg agacaggg 18

<210> 21
 <211> 1823
 <212> ADN

<213> Oryza sativa

<220>

<221> 5'UTR

<222> (1)..(78)

<220>

<221> exon

<222> (79)..(444)

<223> exon 1

<220>

<221> Intron

<222> (445)..(536)

<223> intron 1

<220>

<221> exon

<222> (537)..(717)

<223> exon 2

<220>

<221> Intron

<222> (718)..(803)

<223> intron 2

<220>

<221> exon

<222> (804)..(967)

<223> exon 3

<220>

<221> Intron

<222> (968)..(1074)

<223> intron 3

<220>

<221> exon

<222> (1075)..(1662)

<223> exon 4

<220>

<221> 3'UTR

<222> (1663)..(1823)

<400> 21

acagtgacta gtgacaaacg atcgatcgat cccatccatcc acaaaccctc ctcgatctca 60

tcttccttcg tctcgtca atg gcg gcg agc tac tcg tgc cgg cgg aca tgc 111

Met Ala Ala Ser Tyr Ser Cys Arg Arg Thr Cys
1 5 10

gag gcg tgc agc acg agg gcg atg gcc ggg tgc gtg gtg ggc gag ccg 159

Glu Ala Cys Ser Thr Arg Ala Met Ala Gly Cys Val Val Gly Glu Pro
15 20 25

gcg tcg gcg ccg ggg cag cgg gtg acg ttg ctg gcg atc gac ggc ggc 207

Ala Ser Ala Pro Gly Gln Arg Val Thr Leu Leu Ala Ile Asp Gly Gly
30 35 40

ggc atc agg ggc ctc atc ccg ggc acc atc ctc gcc ttc ctc gag gcc 255

Gly Ile Arg Gly Leu Ile Pro Gly Thr Ile Leu Ala Phe Leu Glu Ala
45 50 55

agg ctg cag gag ctg gat ggc ccc gac gcg cgc ctc gcc gat tac ttc 303
 Arg Leu Gln Glu Leu Asp Gly Pro Asp Ala Arg Leu Ala Asp Tyr Phe
 60 65 70 75

gac tgc atc gcc ggg acc agc acc ggc ggc ctc atc acc gcc atg ctc 351
 Asp Cys Ile Ala Gly Thr Ser Thr Gly Gly Leu Ile Thr Ala Met Leu
 80 85 90

gcc gcg ccc ggc gac cac ggc cgc ccg ctc ttc gcc gcc agc gac atc 399
 Ala Ala Pro Gly Asp His Gly Arg Pro Leu Phe Ala Ala Ser Asp Ile
 95 100 105

aac cgc ttc tac ctc gac aac ggc cca ctc atc ttc cca caa aag 444
 Asn Arg Phe Tyr Leu Asp Asn Gly Pro Leu Ile Phe Pro Gln Lys
 110 115 120

taactgatca cctcgaattc gatctcctct cttcgatctc tgcattatatt gatttgattg 504

gggattgtgg gcggcgtggc gtggcgtcca gg agg tgc ggc atg gcg gcg gcc 557
 Arg Cys Gly Met Ala Ala Ala
 125

atg gcg gcg ctg acg agg ccg agg tac aac ggc aag tac ctg cag ggg 605
 Met Ala Ala Leu Thr Arg Pro Arg Tyr Asn Gly Lys Tyr Leu Gln Gly
 130 135 140 145

aag atc agg aag atg ctg ggc gag acg agg gtg cgc gac acg ctg acg 653
 Lys Ile Arg Lys Met Leu Gly Glu Thr Arg Val Arg Asp Thr Leu Thr
 150 155 160

aac gtc gtc atc ccc acg ttc gac gtc agg ctg ctc cag cca acc atc 701
 Asn Val Val Ile Pro Thr Phe Asp Val Arg Leu Leu Gln Pro Thr Ile
 165 170 175

ttc tcc aca tac gac g tgcgtgcgtt gattccatcc gcattggcgt tggaatcagc 757
 Phe Ser Thr Tyr Asp
 180

tgattgtttg attgatcgaa caattgatcg gttaaaattt tgcagg cg aag agc 811
 Ala Lys Ser
 185

atg ccg ctc aag aac gcg ctc ctc tcc gac atc tgc atc agc aca tcc 859
 Met Pro Leu Lys Asn Ala Leu Leu Ser Asp Ile Cys Ile Ser Thr Ser
 190 195 200

gcg gcg ccg acc tac ctc ccc gcg cac tgc ttc cag acc acc gac gac 907
 Ala Ala Pro Thr Tyr Leu Pro Ala His Cys Phe Gln Thr Thr Asp Asp
 205 210 215

gcc acc ggc aag gtc cgc gag ttc gac ctc atc gac ggc ggc gtc gcc 955
 Ala Thr Gly Lys Val Arg Glu Phe Asp Leu Ile Asp Gly Gly Val Ala
 220 225 230

gcc aac aac ccg gtaactaatc aatcaagcaa tccatcaaac gaagatccac 1007
 Ala Asn Asn Pro
 235

atgtgcattc ctgtggtaca aatgctgac gatcgatgga tggatcgatt ttcgcgagaa 1067

cgtacag acg atg gtg gcc atg acg cag atc acc aag aag ata atg gtg 1116
 Thr Met Val Ala Met Thr Gln Ile Thr Lys Lys Ile Met Val
 240 245 250

aag gac aag gag gag ctg tac ccg gta aag ccg tcg gac tgc ggt aag 1164
Lys Asp Lys Glu Glu Leu Tyr Pro Val Lys Pro Ser Asp Cys Gly Lys
255 260 265

ttc ctg gtg ctg tcc gtg ggc acc ggg tcg acg tcg gac cag ggg atg 1212
Phe Leu Val Leu Ser Val Gly Thr Gly Ser Thr Ser Asp Gln Gly Met
270 275 280

tac acg gcg agg cag tgc tcg cgg tgg ggg atc gtc cgg tgg ctg cgc 1260
Tyr Thr Ala Arg Gln Cys Ser Arg Trp Gly Ile Val Arg Trp Leu Arg
285 290 295

aac aag ggg atg gcg ccc atc atc gac atc ttc atg gcg gcc agc tcc 1308
Asn Lys Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser
300 305 310 315

gac ctc gtc gac atc cac gcc gcc gtc atg ttc cag tcg ctg cac agc 1356
Asp Leu Val Asp Ile His Ala Ala Val Met Phe Gln Ser Leu His Ser
320 325 330

gac ggc gac tac ctc cgc atc cag gac aac acg ctc cac ggc gac gcc 1404
Asp Gly Asp Tyr Leu Arg Ile Gln Asp Asn Thr Leu His Gly Asp Ala
335 340 345

gcc acg gtg gac gcc gcc acc agg gac aac atg cgg gcg ctc gtc ggg 1452
Ala Thr Val Asp Ala Ala Thr Arg Asp Asn Met Arg Ala Leu Val Gly
350 355 360

atc ggc gag cgg atg ctg gcg cag cgg gtg tcg agg gtc aac gtc gag 1500
Ile Gly Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu
365 370 375

acc ggc agg tac gtc gag gtg ccc ggc gcc ggc agc aac gcc gac gcg 1548
Thr Gly Arg Tyr Val Glu Val Pro Gly Ala Gly Ser Asn Ala Asp Ala
380 385 390 395

ctg agg ggc ttc gcc agg cag ctc tcc gag gag agg agg gcg agg cta 1596
Leu Arg Gly Phe Ala Arg Gln Leu Ser Glu Glu Arg Arg Ala Arg Leu
400 405 410

ggt cgg cga aac gcc tgc ggc ggc ggc ggc gaa gga gag ccc agc ggc 1644
Gly Arg Arg Asn Ala Cys Gly Gly Gly Glu Gly Glu Pro Ser Gly
415 420 425

gtg gcg tgc aag cgt tag taactgtaca cgcacatgc tgacgcgac 1692
Val Ala Cys Lys Arg
430

ttttttatatt ttctttttttt ttttttacct ttctagcgga catggggaat aacaagacgt 1752

gacagtagtg caatcggttt gtaacgtgcg tataccaaca ttgatccatt tcttcatcac 1812

agtttcagtt c 1823

<210> 22
<211> 1299
<212> ADN
<213> Oryza sativa

<400> 22
atggcggcga gctactcgtg ccggcggaca tgcgaggcgt gcagcacgag ggcgatggcc 60

```

gggtgcgtgg tgggcgagcc ggcgtcggcg ccggggcagc gggtgacgtt gctggcgatc 120
gacggcggcg gcatcagggg cctcatcccg ggcaccatcc tcgccttcct cgaggccagg 180
ctgcaggagc tggatggccc cgacgcgcgc ctgcgccatt acttcgactg catcgccggg 240
accagcaccg gcggcctcat caccgccatg ctgcgcgcgc ccggcgacca cggccgcccg 300
ctcttcgcgc ccagcgacat caaccgcttc tacctcgaca acggcccact catcttccca 360
caaaagaggt gcggcatggc ggcggccatg gcggcgctga cgaggccgag gtacaacggc 420
aagtacctgc aggggaagat caggaagatg ctgggcgaga cgagggtgcg cgacacgctg 480
acgaacgtcg tcatccccac gttcgacgtc aggtctgtcc agccaaccat cttctccaca 540
tacgacgca agagcatgcc gctcaagaac gcgtcctct ccgacatctg catcagcaca 600
tcgcgggcgc cgacctacct cccgcgcac tgcttcaga ccaccgacga cgccaccggc 660
aaggtccgcg agttcgacct catcgacggc ggcgtcgccg ccaacaacc gacgatggtg 720
gccatgacgc agatcaccaa gaagataatg gtgaaggaca aggaggagct gtaccggta 780
aagccgtcgg actgcggtaa gttcctggtg ctgtccgtgg gcaccgggtc gacgtcggac 840
caggggatgt acacggcgag gcagtgtcgc cggtggggga tcgtccggtg gctgcgcaac 900
aaggggatgg cgcccatcat cgacatcttc atggcggcca gctccgacct cgtcgacatc 960
cacgccgcgc tcatgttcca gtcgtgcac agcgacggcg actacctccg catccaggac 1020
aacacgctcc acggcgacgc cgccacggtg gacgccgcca ccagggacaa catgcgggcg 1080
ctcgtcggga tcggcgagcg gatgctggcg cagcgggtgt cgagggtcaa cgtcgagacc 1140
ggcaggtacg tcgaggtgcc cggcgccggc agcaacgccg acgcgctgag gggcttcgcc 1200
aggcagctct ccgaggagag gagggcgagg ctaggtcggc gaaacgcctg cggcgggcgc 1260
ggcgaaggag agcccagcgg cgtggcgtgc aagcgttag 1299

```

<210> 23
 <211> 432
 <212> PRT
 <213> *Oryza sativa*

<400> 23

Met	Ala	Ala	Ser	Tyr	Ser	Cys	Arg	Arg	Thr	Cys	Glu	Ala	Cys	Ser	Thr
1				5					10					15	

Arg	Ala	Met	Ala	Gly	Cys	Val	Val	Gly	Glu	Pro	Ala	Ser	Ala	Pro	Gly
			20					25					30		

Gln	Arg	Val	Thr	Leu	Leu	Ala	Ile	Asp	Gly	Gly	Gly	Ile	Arg	Gly	Leu
		35					40					45			

Ile	Pro	Gly	Thr	Ile	Leu	Ala	Phe	Leu	Glu	Ala	Arg	Leu	Gln	Glu	Leu
	50					55					60				

Asp Gly Pro Asp Ala Arg Leu Ala Asp Tyr Phe Asp Cys Ile Ala Gly
65 70 75 80

Thr Ser Thr Gly Gly Leu Ile Thr Ala Met Leu Ala Ala Pro Gly Asp
85 90 95

His Gly Arg Pro Leu Phe Ala Ala Ser Asp Ile Asn Arg Phe Tyr Leu
100 105 110

Asp Asn Gly Pro Leu Ile Phe Pro Gln Lys Arg Cys Gly Met Ala Ala
115 120 125

Ala Met Ala Ala Leu Thr Arg Pro Arg Tyr Asn Gly Lys Tyr Leu Gln
130 135 140

Gly Lys Ile Arg Lys Met Leu Gly Glu Thr Arg Val Arg Asp Thr Leu
145 150 155 160

Thr Asn Val Val Ile Pro Thr Phe Asp Val Arg Leu Leu Gln Pro Thr
165 170 175

Ile Phe Ser Thr Tyr Asp Ala Lys Ser Met Pro Leu Lys Asn Ala Leu
180 185 190

Leu Ser Asp Ile Cys Ile Ser Thr Ser Ala Ala Pro Thr Tyr Leu Pro
195 200 205

Ala His Cys Phe Gln Thr Thr Asp Asp Ala Thr Gly Lys Val Arg Glu
210 215 220

Phe Asp Leu Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val
225 230 235 240

Ala Met Thr Gln Ile Thr Lys Lys Ile Met Val Lys Asp Lys Glu Glu
245 250 255

Leu Tyr Pro Val Lys Pro Ser Asp Cys Gly Lys Phe Leu Val Leu Ser
260 265 270

Val Gly Thr Gly Ser Thr Ser Asp Gln Gly Met Tyr Thr Ala Arg Gln
275 280 285

Cys Ser Arg Trp Gly Ile Val Arg Trp Leu Arg Asn Lys Gly Met Ala
290 295 300

Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile
305 310 315 320

His Ala Ala Val Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu
325 330 335

Arg Ile Gln Asp Asn Thr Leu His Gly Asp Ala Ala Thr Val Asp Ala
340 345 350

Ala Thr Arg Asp Asn Met Arg Ala Leu Val Gly Ile Gly Glu Arg Met
355 360 365

Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Tyr Val
370 375 380

Glu Val Pro Gly Ala Gly Ser Asn Ala Asp Ala Leu Arg Gly Phe Ala
385 390 395 400

Arg Gln Leu Ser Glu Glu Arg Arg Ala Arg Leu Gly Arg Arg Asn Ala
405 410 415

Cys Gly Gly Gly Gly Glu Gly Glu Pro Ser Gly Val Ala Cys Lys Arg
420 425 430

<210> 24

<211> 1371

<212> ADN

<213> Zea mays

<400> 24

agttcatcac taatcacact tattgtgccc tcgacgagta tctatagcta gctcattaat	60
cgattcgggg gtgtgttgtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc	120
aatacctgta gcacgaaggc gatggccggg agcgtggctcg gcgagcccgt cgtgctgggg	180
cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc	240
atcctcgctt tcttgaggc caggctgcag gagctggacg gaccggaggc gaggtggcg	300
gactacttcg actacatcg cggaaccagc accggcggtc tcatcaccgc catgctcacc	360
gcgcccggca aggacaagcg gcctctctac gctgccaaagg acatcaacca cttttacatg	420
cagaactgcc cgcgcattct tctcagaag agcaggcttg cgcccgccat gtccgcgctg	480
aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag	540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg	600
tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag	660
gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgac gatggttgcg	720
atgacgcaga tcacaaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca	780
gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc	840

gagcaggggcc tctacacggc gcggcagtgcc tcccgggtggg gtatctgccg gtgggtccgc 900
 aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctgggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggaggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtctc tgccatcaac 1260
 ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 25
 <211> 1375
 <212> ADN
 <213> Zea mays

<400> 25
 agttcatcac taatcacact tattgtgccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgccct tcctggaggc caggctgcag gagctggacg gaccggaggc gaggtggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaaagg acatcaacta cttttacatg 420
 gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt gcgagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc totgaagaac gogctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaaccctg cyaaactgcc caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgcc tcccgggtggg gcattctgcc gtgggtccgc 900
 aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctgggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct acgcatccag 1020
 gacaactcgc tccgtggcgc cgcggcaacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagcggg tgtccagggt caacgtggag 1140
 acagggagcg aggtacgaac cggtgaccgg agaaggaagc aatgccgatg ccctcgggtg 1200

gctcgctagg cagctctccg aggagaggag aacaaggctc gcgcgccgcg tctctgccat 1260
 caaccccaga agctctagat gtgcgcccta cgatatctaa gacaagtggc tttactgtca 1320
 atcacatgct tgtaaataag tagactttat tttaataaaa tataaaaata tatat 1375

<210> 26
 <211> 1375
 <212> ADN
 <213> Zea mays

<400> 26
 agttcatcac taatcacact tattgtgcc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctggaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaaagg acatcaacta cttttacatg 420
 gagaactgcc cgcgcattct ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gcgctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgcgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaaccctg cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtgcagctcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggg gcattctgcc gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct acgcatccag 1020
 gacaactcgc tccgtggcgc cgcggcaacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagcggg tgtccagggt caacgtggag 1140
 acagggagcg aggtacgaac cggtgaccgg agaaggaagc aatgccgatg ccctcggtgg 1200
 gctcgctagg cagctctccg aggagaggag aacaaggctc gcgcgccgcg tctctgccat 1260
 caaccccaga agctctagat gtgcgcccta cgatatctaa gacaagtggc tttactgtca 1320
 atcacatgct tgtaaataag tagactttat tttaataaaa tataaaaata tatat 1375

<210> 27
 <211> 401
 <212> PRT

<213> Zea mays

<400> 27

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

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

28

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
210 215 220

Ala Glu Glu Leu Tyr Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu
225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
340 345 350

Ser Glu Val Arg Thr Gly Asp Arg Arg Arg Lys Gln Cys Arg Cys Pro
355 360 365

Arg Trp Ala Arg
370

<210> 30

<211> 15463

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 22466

<220>

<221> vùng khởi động

<222> (214)..(2339)

<220>

<221> gen

<222> (2341)..(3018)

<220>

<221> trình tự kết thúc

<222> (3024)..(4036)

<220>

<221> trình tự tăng cường

<222> (4044)..(4237)

<223> eFMV

<220>

<221> trình tự tăng cường

<222> (4244)..(4536)

<223> e35S

<220>

<221> vùng khởi động

<222> (4544)..(8537)

<223> Ubi

<220>

<221> trình tự kết thúc

<222> (10009)..(11008)

<223> Ubi

<220>

<221> vùng khởi động

<222> (11028)..(12618)

<223> pPLA2

<220>

<221> exon

<222> (12619)..(12977)

<223> PLA2 exon1

<220>

<221> Intron

<222> (12978)..(13096)

<223> PLA2 intron1

<220>

<221> exon

<222> (13097)..(13193)

<223> PLA2 exon2

<220>

<221> Intron

<222> (13194)..(13388)

<223> PLA2 intron2

<220>

<221> exon

<222> (13389)..(13550)

<223> PLA2 exon3

<220>
 <221> Intron
 <222> (13551)..(13660)
 <223> PLA2 intron3

<220>
 <221> exon
 <222> (13661)..(14248)
 <223> PLA2 exon4

<220>
 <221> trình tự kết thúc
 <222> (14249)..(15244)
 <223> tPLA2

<400> 30
 cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagtttaaac tggcactagc ctaacgggtg tgactaacta ggccgcttcc 180
 ctaattagct aaccggggg cgcgccggga cccgaaagta gcaaacaaca ggttcatgtg 240
 cactataaaa agacaaaatt ctcgagtttc atcttttatt ccacataagc cttatatattt 300
 ccattttcat atgattttta gtttaagttt gtgtcttaac tttttcgta atacgtaatt 360
 ctatgcatta tggatgctg aagtattttt gtttaaaaa atgaaatgtc aaaatacgtt 420
 ttgtgatcta tttccatgtt ttcacctaac aggtggtttt tactatatat tctgccataa 480
 ctctagcctt agatgtaaat cgaaaaaaaa tgagagatga gctggagata gccttagatg 540
 aagcgtctga aatataaaag aaagagtaat gttgaacgca gtaggtgtag cagctgtagt 600
 tccatctcta ggaaaggga ctgcaatccg ggctccgggc ctgcgcgaat ctggcctgtc 660
 gtgtagatgc agccctgtcc atgacggccc aagcaacgcc cgcggtctc gatccaccac 720
 ggaaccact ccgacacaca ctgacacaca catgctggat gtggatgtgc tgtccaatta 780
 ttagtagcaa ttcggtaggc acaggcacgt actggcgggt gtttttagctg taagtaccga 840
 accaatcacg gttaagaacc gattaatccg tgcccagccg ccgagtgcgt tcgtacgtgc 900
 atcggatgca ctgcatgaat tgagagcatc atcatatcat acgcaggagt agtacgacgc 960
 cgctgctgtc ttgtccggct aatgctttgc tcacagatta gtccatcgcc caccggtcgg 1020
 gtggtgtgga tcgctgatgc cactgctttt tgtttggttt ttattccctt gataatcctc 1080
 cgcgtccttg aatgtatcta tttattttca ttccgaaatc cttttcacga aaaagaaaac 1140
 gaataaaaag agagttacga atacgcttcc ggcgggccac atcaccttcc agcgaacatc 1200
 gcgcgcgcgt gacgtgtgc ccacgcggc cgtccatata gccatccgac gaccgtggaa 1260
 gctggcagcg gccgctccgt tccgtcgaag gggcagggtca gtcaggtcac ccacacggcc 1320
 acaccgcgcg gggggatacg cgggtgaaaa cccggcgacc acatcaaaac acgaggcgtc 1380
 tccgcagga ctggtcactc ggcacgcagg cagaggcagc acagcagcag ccagctccat 1440

ccatcctctt	tccccctctc	gcttcgcttc	ctcggcggat	tcctcctccc	tcggccgtcc	1500
ccgtccccct	cttcgccgcg	ccagctcgcc	cgagttggta	cgcccatcct	ctgctgtact	1560
ccccccgttc	ctgctgtgcg	aaaccgagca	cccaaaccct	agcctaagct	attcgcatcg	1620
caaaactctat	tgagcgacgg	atcgcgaaac	gcgcgccgcc	gcattcggac	gagacctcgt	1680
ggccccgtacc	ttcctactct	accgtcgtcg	ccggagacca	gctcaccggg	taggggtttcg	1740
ggattagggg	cttggggcct	ggattcccgg	gggactccat	acccgggtgg	cttagaaggg	1800
gaagggggcc	cagcgggtgt	ggtggttcga	ctgtcagcga	gcgagacggc	ggggcaccgc	1860
cgatcggggc	tcgctggaca	tttgattgag	ctgggcaggc	gaaggcgtgg	agcttgctct	1920
tcgattggga	ttagaggagg	gcggaggtgg	tatggtgggc	ggctcggctg	gctcaatgga	1980
gccgtgggcc	ggtggtgggg	caggcgagat	cacgctctct	catagagcgt	aatgggttgc	2040
gtaacacact	cctgtttgta	tttcgatccg	atctaataat	taaagtttat	ttacgtggtg	2100
cacatgtaga	atthttcccat	gccactgct	aacttggatt	ggatttttgc	caaaaaaacg	2160
ggtgtgcacc	caatcactta	tagtgaaacc	gttgtgtgtt	tgatgtgctt	ctaaccaagt	2220
acaaatccag	tgaaaacaca	cctaccattc	aacaatccac	atthttggtg	caattatcag	2280
cattctagaa	aggtgcatgt	gccattgat	acacttgcta	ttggtgcagg	caaaccacc	2340
atggcctcct	ccgagaacgt	catcaccgag	ttcatgcgct	tcaagggtgc	catggagggc	2400
accgtgaaag	gccacgagtt	cgagatcgag	ggcgagggcg	agggccgccc	ctacgagggc	2460
cacaacaccg	tgaagctgaa	ggtgaccaag	ggcgggcccc	tgcccttcgc	ctgggacatc	2520
ctgtcccccc	agttccagta	cggctccaag	gtgtacgtga	agcaccgccg	cgacatcccc	2580
gactacaaga	agctgtcctt	ccccgagggc	ttcaagtggg	agcgcgtgat	gaacttcgag	2640
gacggcgggc	tggcgaccgt	gaccaggac	tcctccctgc	aggacggctg	cttcatctac	2700
aagggtgaagt	tcacgcggct	gaacttcccc	tccgacggcc	ccgtgatgca	gaagaagacc	2760
atgggctggg	aggcctccac	cgagcgctg	tacccccgcg	acggcgtgct	gaagggcgag	2820
accacaaagg	ccctgaagct	gaaggacggc	ggccactacc	tggtggagtt	caagtccatc	2880
tacatggcca	agaagcccgt	gcagctgccc	ggctactact	acgtggacgc	caagctggac	2940
atcacctccc	acaacgagga	ctacaccatc	gtggagcagt	acgagcgcac	cgagggccgc	3000
caccacctgt	tcctgtagga	gctcgcatca	tgatcatgca	tcattggactc	ggcctactac	3060
tgtggatttg	tatgccatta	tagacttggg	gctgtgaaag	actgcttgat	gatttgcggg	3120
tttggtgctg	tgtaaaaaaa	ggtcccttgg	ctcccagaag	accatgaagg	ttcggatcta	3180
tcattgtaatt	ccttgttatc	tgccaattat	gtatggacta	tggacatgtg	ttgcgctgtt	3240
caacttacta	ctacaaataa	gtaatcgata	tgthccccct	ccatgtctcg	gtgacaattg	3300
tctggagaag	cttaggggtc	gtttgtttgg	gattatgtct	ggagaaactt	atthttaaact	3360

aagtgtgagt tcaagttaag ttagattata taatctaggc agattataat tccaagcgaa 3420
caggtcctta gtgttttttg aaaatcctag gtgttctttt ggctacattg ttgtgtgtgc 3480
agatcccttg ttggtctgta agcgtgggga agtaagaatc gtccgtttct actgaagacc 3540
tgctcgagtt aggcaccgag gatgccggt accaaacaga gcaatagtgt ctctgtgggc 3600
acagtggagt gtgaatctgt gtgatgcaaa tccgtcattt gtttagcaaa atttccagcg 3660
ttgcatgatg cagtttcttt aacacggact taagggaagg gaaaaaatg ttgagccagg 3720
agatccttca atgtgttaga ctgacgtgat agccaactaa accacgacgc aatgttgtcg 3780
ttaatgacaa aaaaactatt tgttcctaaa tccttggcga cattgcatgg ctgtctcatg 3840
agataatggc ctcatctctt atttatctct tatttatagc cggaagtggg agtgaccctt 3900
gcttgattgc tcgtatgcc tctcaagttc tcaaccgtgt cgagcagcca ttttcccatc 3960
tcaagcgcac catcgtttcg tttgacctca tctgctatcc tgctcctagt gcaaatcaca 4020
tgcgacagaa agtgtgcgga ccagctgct tgtggggacc agacaaaaaa ggaatggtgc 4080
agaattgtta ggcgaccta ccaaagcaa ctttgccttt attgcaaaga taaagcagat 4140
tcctctagta caagtggga acaaaataac gtggaaaaga gctgtcctga cagccactc 4200
actattgctt ttgacgaacg cagtgcgac caaaaactc gagacttttc aacaaagggg 4260
attatccgga aacctcctcg gattccattg ccagctatc tgtcacttta ttgtgaagat 4320
agtggaaaag gaaggtggct cctacaaatg ccatcattgc gataaaggaa aggctatcgt 4380
tgaagatgcc tctgccgaca gtggtcccaa agatggacc ccaccacga ggagcatcgt 4440
ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg atatctccac 4500
tgacgtaagg gttgacgaac aatcccacta tccttcaagc tttggaaaac caaaacaacg 4560
aataagcaaa ctgcaggaaa agtatgcagt ggaaaccaac ccagattcgg acgataggaa 4620
agtatcaagt gaatgatttg ccaggaaaag gagaggggta aaaaggggag aagattttaga 4680
agatctaaag cacaagaacc agagattaga ttgaacaata ggggaacttg agcatccttt 4740
ttttcttcag ggaaaaactg aaaatccaaa ccatgttgag caaaaccgag tgggattgga 4800
aaccaaaaaa ccgagataa agaaactcga gaaaaagcat gaaatcgaaa ccaacttcag 4860
taaaacaaaa ggaggacaga aaagaaagtc ggaagctata agaatacat taacattcag 4920
tgaaacagca tgctgtcttc ttcttttttt atgcacaaca gagcatacat atataccttc 4980
ccaggctgag gacttggcgg aggagagccg cggatagggt ggcggtgcag acggtctgga 5040
cgggcccga gacggagacg aacagcgggc ccttctgcc caggcaccac gottggaacg 5100
ccaagcacgc gccaacgcg gcccgcgga ggaacgacat ccccgcgaca agcgtggcgt 5160
cgatcctggg cgaccccaag ccgaggaacc tccttccag gacgagccgt aggacggcgg 5220
tgagagaggc acccgtcgag gaggtggcgc agcacaaggt gagcagcgcg ggaaggcgg 5280

cgagcgtggc ggccctgcagg acggtgacga ggcggaagac ggtgacgccg ggcgacgaggc 5340
 agcagcagcc gaggatccag tcgtaggagg aagccggacc aaaccgggca atgcaacctg 5400
 cagatgcact agacggaggt aacgaggagg aggagaaaac agagcaagag caggcggaga 5460
 gaagatagag caaaacacga gtgaggcaca gcgtaagcac tcggtagaag tctccagagg 5520
 cgaggtgcmc acaggagaac agatgagtaa agtcagccaa ggatccacga tccaacggct 5580
 acgaatTTTT ggagtgcgt ggataggctc aaaggcgcca tttccatccg gctttatagt 5640
 attttaaaaa aattcatttt cctccctcta gtgtgtgcgg aggcgtgagc ccgtttaacg 5700
 gcgttgagaa gtctaacgga caccaaccac aaccaggaac cagcgccggc cgcgccggccg 5760
 agtgaagcag actgcatacg gcacggcgcg gcatctctct ggctgcctct cgagagttcc 5820
 gccccacct tcccgcgga gcgtgggtgt ttcgctttcc gctgtcggca tccggaagtt 5880
 gcgtggcaga gtggacggag acgaggccgg gtcctccagc tcctctcaaa cgtcacggca 5940
 ccggcatccg gcagccagcg cggtccttcc caaccactcg ttcccaacct atcccccttc 6000
 ctgccccgc gtcataaata gccagcccca tcccagctt ctttcccaa cctcatcttc 6060
 tctccttttg ctctgaacgc acacaccgcc cggctctcga tctccgatcc ccgatccct 6120
 cgtcgatcct aggtacggcg accatcctac ccccccccc cccccctctc tctctgcctt 6180
 ctctagatcg gcgatccgat ccatgcttac ttggttaggg cctgctaact atgttcatgt 6240
 ttgcgttaga tccgtgcatg gacgcgatct gtacacacca gacgcgttct gattgctagc 6300
 taactcgcca gtacctggga atcctgggat ggctgtagcc ggccccgcac gcagacggga 6360
 ccgatttcat gattctctat ttttttcttt gtttcgttgc ctagggtttc gttcgatoga 6420
 tccgcgttat tctttatttc catatattct ggtacgatgt tgatacggtt cgaccgtgct 6480
 gcttacgttc tgtgcgcttg tttgcgggt cattttttacc ttgccttttt tgtatggttt 6540
 ggttgtggcg atgtggtctg gtcgggctgt cgttctagat cggagtagag tgctgtttca 6600
 aactgtctag cggatctatt agatttggat ctgcatgtgt gacatatatc ttcgtagtta 6660
 agatgatgca tctgtatgtg tgacatgcgg atctattaga tttggatctg tatgtgtgac 6720
 atatatcttc gtagttgaga tgatgcatct gtatgtgtga catatatctt cgtagttaag 6780
 attatgcatg gaaatatcaa tccttttagat aaggacgggt atacttgttg ctgtgggttt 6840
 tactggtact tcgatagatg catatacatg atctaactg cttagataga tgaagtaaca 6900
 tgctgctacg gtttaataat tcttgagttg atttttactg gtacttagat agatgtatat 6960
 acatgcttag atacatgaag taacatgctc ctacagttcc tttaatcatt attgagtacc 7020
 tatatattct aataaatcag tatgttttaa attattttga ttttactggg acttagatag 7080
 atgtatatat acatgctcaa acatgcttag atacatgaag taacatgctg ctacgggttta 7140
 gtcattattg agtgccata tattctaata aatcagtatg ttttaaatta ttttgatttt 7200

actggtactt agatagatgt atatatacat gctcaaacat gcttagatac atgaagtaat 7260
 atgctactac ggtttaattg ttcttgagta cctatatatt ctaataaatc agtatgtttt 7320
 aaattatttc gattttactg gtacttagat agatgtatat atacatgctt agatacatga 7380
 agtaacatgc tactacggtt taattgttct tgaataccta tatattctaa taaatcagta 7440
 tgttttaaat tatttcgatt ttactggtac ttagatagat gtatatatac atgctcgaac 7500
 atgcttagat acatgaagta acatgctaca tatatattat aataaatcag tatgtcttaa 7560
 attattttga ttttactggt acttagatag atgtatatac atgctcaaac atgcttagat 7620
 acatgaagta acatgctact acggtttaat cattattgag tacctatata ttctaataaa 7680
 tcagtatgtt ttcaattgtt ttgattttac tggtagcttag atatatgtat atatacatgc 7740
 tcgaacatgc ttagatacgt gaagtaacat gctactatgg ttaattgttc ttgagtacct 7800
 atatatctta ataaatcagt atgtttttaa ttatttcgat ttactggta cttagataga 7860
 tgtatatata catgctcgaa catgcttaga tacatgaagt aacatgctac tacggtttaa 7920
 tcgttcttga gtacctatat attctaataa atcagtatgt cttaaattat cttgatttta 7980
 ctggtactta gatagatgta tatacatgct tagatacatg aagtaacatg ctactatgat 8040
 ttaatcggtc ttgagtacct atatatctta ataaatcagt atgtttttaa ttattttgat 8100
 ttactggta cttagataga tgtatatata catgctcgaa catgcttaga tacatgaagt 8160
 aacatgctac tacggtttaa tcattcttga gtacctatat attctaataa atcagtatgt 8220
 ttttaattat ttgatatta ctggtactta acatgttttag atacatcata tagcatgcac 8280
 atgctgctac tgtttaaatca ttctggaata cctatatatt ctaatatatc agtatgtctt 8340
 ctaattatta tgattttgat gtacttgtat ggtggcatat gctgcagcta tgtgtagatt 8400
 ttgaataccc agtgtgatga gcatgcatgg cgccttcata gttcatatgc tgtttatttc 8460
 ctttgagact gttctttttt gttgatagtc accctgttgt ttggtgattc ttatgcagat 8520
 ccagatcttc gtaaaccatg gctctggtgg ctaggccagt gctgtctgct agagtggctg 8580
 cttctaggcc aagggtggcc gctagaaagg ctgtgagggg gagcgctgcc agtcacggcg 8640
 ctagctctag gccagccacc gctaggaagt ccagcggcct gagcggcacc gtgcgcatcc 8700
 caggcgacaa gagcatcagt caccgcagct tcatgttcgg cggcctggcc agcggcgaga 8760
 ctaggatcac cggcctgctc gagggcgagg acgtgatcaa caccggcaag gccatgcagg 8820
 ctatgggcgc tcgcatccgc aaggagggcg acacctggat catcgacggc gtgggtaatg 8880
 gcggcctgct ggccccagag gcccactgg acttcggcaa cgcgcgccacc ggctgcaggc 8940
 tgactatggg cctgggtggg gtgtacgact tcgacagcac ctttatcggc gacgccagcc 9000
 tgaccaagag gccgatgggc aggggtgctga acccgctgcg cgagatgggc gtgcaggcca 9060
 agagcgagga cggcgacagg ctgccggtga cctgagggg cccaaagacc ccgaccccca 9120

tcacctaccg cgtgccaatg gccagcgccc aggtgaagtc agccgtgctg ctggccggcc	9180
tgaacacccc aggcatacc accgtgatcg agccgatcat gacccgcat cacaccgaga	9240
agatgctgca gggcttcggc gccaacctga ccgtcgagac tgacgctgac ggctgagga	9300
ccatcaggct cgagggcagg ggcaagctga ccggccagggt gatcgacgtg ccgggcgacc	9360
caagcagcac cgccttccca ctggtggccg ccctgctggt gccaggctct gacgtgacca	9420
tcctgaacgt gctgatgaac ccgagccgca ccggcctgat cctgaccctg caggagatgg	9480
gcgccgatat cgaggtgatc aacccgaggc tggctggcgg cgaggacgtc gccgacctga	9540
gggtgaggtc cagcaccctg aagggcgtga ccgtgccaga ggacagggcc ccgagtatga	9600
ttgacgagta ccgatacctg gccgtggccg ctgccttcgc cgagggcgcc accgtgatga	9660
acggcctcga ggagctgcgc gtgaaggaga gcgacaggct gagcgctgtg gccaacggcc	9720
tgaagctgaa cggcgtggac tgcgacgagg gcgagactag cctgggtggtg aggggcaggc	9780
cagacggcaa gggcctgggc aacgcttccg gcgctgccgt ggccaccac ctggatcaca	9840
ggatcgccat gtcgttcctc gtgatgggccc tcgtgagcga gaaccgggtg accgtggacg	9900
acgccaccat gatcgccacc agcttccccg agttcatgga cctgatggcc ggctggggcg	9960
ctaagatcga gctgagcgac accaaggccg cctgatcacc tagagctcgc caaggttcaa	10020
ttaagctgct gctgtacctg ggtatctgcg tcgtctggtg ccctctggtg tacctctata	10080
tggatgtcgt cgtctaataa acatctgtgg tttgtgtgtc atgaatcgtg gttgtggctt	10140
cgttggttta atggacctgt tgtgtcctct gtgttgatcc caaaactctt ctgcagcagt	10200
atggcttgaa tccttatgaa gtttgatatt tgaacttaaa agtctgctca ttatgttttt	10260
ttctggttat atctcctaataa taactgcctg ggatcaaatt tgattcgctg gtgtttattg	10320
gacccctccc aggttcttgc tttctaccgt ttcttgctga atgttaactt gattctgtca	10380
ggctcagttt ccactatgg cttacagctt aacgtgtttg gtttgttgaa tggttaacttg	10440
gttttggtcaa gctcagtttt ttactctggc ttacagcata acatgtttga cttttggttt	10500
tgctgctttg ttattgggtt ctgggtagtt cttgatgaat ccaaaagatc atgtgcacag	10560
ccatattatc tatttaagcg atccagggtta ttactatgaa aggatgcctt ctagctaagg	10620
agtagttagg tttttcttcc aagggttaaat tttctcgatg ctctagtgtt cctgtgacca	10680
taatcataat aattcctttg aaagctctat ggtccctgga agcagggcat acaatgcaag	10740
acagcaactt gatcacatca actgaagtat acagggttct cttactctt ggtgacttcg	10800
gtttaatgga ccggttgtag tcgtgttcta tccgtaaccg ttgtgatgtc ttgtgtgttt	10860
ggttgcccga tagctgggac cagcagcttt ccgtctaatt ctgatggata gctatagacg	10920
gcactgagat ggttatatta taacctctga tcctgaactc tacgagatcg tctcatccgt	10980
cattgccacc aaatacacca tttaaattact aattagctaa cggaccctat ttgtactcat	11040

tccatgtctc ataaactttg ggcaccatcc atccaacaca tccaatctaa acacaccaaa 11100
 cgatgggggaa tggaaagagc agtattcgat tcaacaatgg caaacaata tcaactgaatt 11160
 agaccaagaa taaacctaat tagacaacga cctcccaacc atcattcgtc aggctgtaaa 11220
 gaagataaag ctgccatggg gcatggatca agcagaacac cagagatgaa tccaaacaca 11280
 cagaaaatca cgcgcgctgt ctacaatgac aacaagcccc acatttcatt gcagtacact 11340
 gggctacaaa ggcacgtaca acaaagagct agggaaacat tgcggagggc acgagagagc 11400
 agctaacttg acaatatagc agactgagct tgcactgtta gcaggcgagg aagggaatca 11460
 tggggacgga gaatggggtc catgcccgcg aaggagaagg cggacgccgc cacggtggca 11520
 ccggcgcacg cgcacacagg gaaccgcac aggcagccat ggatgctgcc tcgccattgc 11580
 gccggtcgtc tctgccacgc tctctctctc tctccgctgc atcgccgtgg atggggcaag 11640
 cagagagcag ggactgacg gatctgggcg gaggactcgc cttggagagc gcggacgcag 11700
 acgggattct agggagagag cgaagacggg gcgcgcgcgc cgctcgcgcg gcgtgggtggc 11760
 ggcgagatta gcgggggtgg ggggagggcg gagccgtggg gaggggtgtg acgccctcct 11820
 taccctctta agtagtagta gagatataat ccgttcctaaa atatccatcc gttcaattta 11880
 tatttcgttt gatcttttta ccctaaattt gattgactca tcttattaaa aaagttcata 11940
 actattatta atctttattg agatatcatt tagcatataa tatactttta gtgtgggttt 12000
 agattttttt taaaaaaaaa aattcgcaaa aattaaatga aacgaccaa tcaaacttga 12060
 aaagtaaaac taattataaa tttgaacgga aggagtaaga ggatgtttga atgtactaga 12120
 gctaatagtt ggttgcttta aaatttgcta gtagaattag ctagctaata aatatctaga 12180
 taactattag ctaatttgct aaaacagcta atagttgaac tattagctag attgtttgga 12240
 tgtattcggc taattttaat ggctaactat tagctatagt acaatattca aacacctcct 12300
 aattaaaatg gacaaatata tcttcttttg gtcccttgcg ttagattttt catatctcct 12360
 tatttagtat aaaagaatca tcaaaaagtg gacaaccctc agtggaacac catttttagta 12420
 gtggttgcat gaaacctttc gcgcaccagt ttctatgtgt cactctaaaa atgggacagc 12480
 atgtacgtag tgcctatata tatacaagtc atctatcggt gcctcctcag ttcactacta 12540
 atcacactta ttgtgccctc gaagagtata tatagctagc tcattaatcg attcgggggt 12600
 gtgttgctga aggcggca atg gcg agc tac tcg tcg cgg cgt cca tgc aat 12651
 Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn
 1 5 10
 acc tgt agc acg aag gcg atg gcc ggg agc gtg gtc ggc gag ccc gtc 12699
 Thr Cys Ser Thr Lys Ala Met Ala Gly Ser Val Val Gly Glu Pro Val
 15 20 25
 gtg ctg ggg cag agg gtg acg gtg ctg acg gtg gac ggc ggc ggc gtc 12747
 Val Leu Gly Gln Arg Val Thr Val Leu Thr Val Asp Gly Gly Gly Val
 30 35 40

cgg ggt ctc atc ccg gga acc atc ctc gcc ttc ctg gag gcc agg ctg Arg Gly Leu Ile Pro Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu 45 50 55	12795
cag gag ctg gac ggg ccg gag gcg agg ctg gcg gac tac ttc gac tac Gln Glu Leu Asp Gly Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr 60 65 70 75	12843
atc gcc gga acc agc acc ggc ggt ctc atc acc gcc atg ctc acc gcg Ile Ala Gly Thr Ser Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala 80 85 90	12891
ccc ggc aag gac aag ccg cct ctc tac gct gcc aag gac atc aac cac Pro Gly Lys Asp Lys Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His 95 100 105	12939
ttt tac atg cag aac tgc ccg cgc atc ttt cct cag aa gtgagtccga Phe Tyr Met Gln Asn Cys Pro Arg Ile Phe Pro Gln Lys 110 115 120	12987
tgctgccgcc attgttcttg catccatcca gcatcgtacg tacgtcctct atacatctgc	13047
ggatcatcat gtgcgcatgt ttgtggcatg catgcatgca tgtgagcag g agc agg Ser Arg	13103
ctt gcg gcc gcc atg tcc gcg ctg agg aag cca aag tac aac ggc aag Leu Ala Ala Ala Met Ser Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys 125 130 135	13151
tgc atg cgc agc ctg att agg agc atc ctc ggc gag acg agg Cys Met Arg Ser Leu Ile Arg Ser Ile Leu Gly Glu Thr Arg 140 145 150	13193
gtaagcgaga cgctgaccaa cgatcatcatc cctgccttcg acatcaggct gctgcagcct	13253
atcatcttct ctacctacga cgtacgtacg tcgtcacgaa tgattcatct gtacgtcgtc	13313
gcatgcgaat ggctgcctac gtacgccgtg cgctaacata ctcagctctt tcctatctgc	13373
tgcgccaatt tgcag gcc aag agc acg cct ctg aag aac gct ctg ctc tcg Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser 155 160	13424
gac gtg tgc att ggc acg tcc gcc gcg ccg acc tac ctc ccg gcg cac Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His 165 170 175 180	13472
tac ttc cag act gaa gac gcc aac ggc aag gag cgc gaa tac aac ctc Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu 185 190 195	13520
atc gac ggc ggt gtg gcg gcc aac aac ccg gtaactgact agctaactgg Ile Asp Gly Gly Val Ala Ala Asn Asn Pro 200 205	13570
aaaacggacg cacagactcc atgtocatgg cggcccacaa ggtcgatgct aattgttgc	13630
tatgtatgtc gcccgattgc acatgcgtag acg atg gtt gcg atg acg cag atc Thr Met Val Ala Met Thr Gln Ile 210	13684
acc aaa aag atg ctt gcc agc aag gac aag gcc gag gag ctg tac cca	13732

Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr Pro	
215 220 225 230	
gtg aag ccg tcg aac tgc cgc agg ttc ctg gtg ctg tcc atc ggg acg	13780
Val Lys Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly Thr	
235 240 245	
ggg tcg acg tcc gag cag ggc ctc tac acg gcg cgg cag tgc tcc cgg	13828
Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser Arg	
250 255 260	
tgg ggt atc tgc cgg tgg ctc cgc aac aac ggc atg gcc ccc atc atc	13876
Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile Ile	
265 270 275	
gac atc ttc atg gcg gcc agc tcg gac ctg gtg gac atc cac gtc gcc	13924
Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val Ala	
280 285 290	
gcg atg ttc cag tcg ctc cac agc gac ggc gac tac ctg cgc atc cag	13972
Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile Gln	
295 300 305 310	
gac aac tcg ctc cgt ggc gcc gcg gcc acc gtg gac gcg gcg acg ccg	14020
Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr Pro	
315 320 325	
gag aac atg cgg acg ctc gtc ggg atc ggg gag cgg atg ctg gca cag	14068
Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala Gln	
330 335 340	
agg gtg tcc agg gtc aac gtg gag aca ggg agg tac gaa ccg gtg act	14116
Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Tyr Glu Pro Val Thr	
345 350 355	
ggc gaa gga agc aat gcc gat gcc ctc ggt ggg ctc gct agg cag ctc	14164
Gly Glu Gly Ser Asn Ala Asp Ala Leu Gly Gly Leu Ala Arg Gln Leu	
360 365 370	
tcc gag gag agg aga aca agg ctc gcg cgc cgc gtc tct gcc atc aac	14212
Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg Arg Val Ser Ala Ile Asn	
375 380 385 390	
cca aga ggc tct aga tgt gcg tcg tac gat atc taa gacaagtggc	14258
Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp Ile	
395 400	
tttactgtca gtcacatgct tgtaaataag tagactttat ttttaataaaa cataaaaaata	14318
tatatatggt cttgaatata aaattgataa ccaaattaaa attogaacca tcaattatac	14378
ataatttttac tttatttttt ataaaacgtg aaoggggaagg actaccgtga atgactatag	14438
aaccaatcat actagtataa aatatatgat gacactacgg gagagacaaa ctttgtctgg	14498
cgctaaatat tttgccgagt gtgaattcac gggcactagg caaagatcctt ctttgccgag	14558
tgttacgctg ggcaaagtaa gacactaggt aaatcagtca tttgccgagt gtccgccact	14618
aggcaaagca aaacactggc aaatcaaaag tttacctagt gccagacact aggcaaaaaa	14678
aaaacgctcg gcaaatecga agtttcccta gtgccagaca ctagacaaaag aaaaacactt	14738
gataaactag cgtcgtcagc taacaccatc caccaaccgt taacgttgcc gagtatctga	14798

cttcgacact cggcaaagaa ggtctctttg cctagtgtcg gtctggaaca ctaggcaaag 14858
 aggcacttta cctagtgtcg tatTTtgaca ctCagtaaaa taatTTTTTT tctttctgct 14918
 tccaaacttt ttatgatgtg ttCctatagc acctagaact acatgtcaag ttttggtaaa 14978
 atttttgaag tttttgctat atttacttaa tttatTTtat ttaattgaat ttcttttgat 15038
 aattcaaatt tgaactcggc aaggtaaagaa gcgagggtag cctggaaaca cactttgcct 15098
 agtggtacac tcggtacagg agcctccct gcctagtgtc gcactcgaca aaagattcgc 15158
 ctttgcttag cgctgcactc ggcacaggag tcgcctttgc ctagtgctgc actaggcaaa 15218
 gcctccgtta ccgtgccttc catcgtcgga ccgcctgcag gcccgggggc gcgccctaatt 15278
 tagctaacgg ccaggatcgc cgcgtgagcc tttagcaact agctagatta attaacgcaa 15338
 tctgttatta agttgtctaa gcgtaattt gtttacacca caatatatcc tgccaccagc 15398
 cagccaacag ctccccgacc ggcagctcgg cacaaaatca ccactcgata caggcagccc 15458
 atcag 15463

<210> 31

<211> 15757

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 22467

<220>

<221> vùng khởi động

<222> (214)..(2339)

<220>

<221> gen

<222> (2341)..(3018)

<220>

<221> trình tự kết thúc

<222> (3024)..(4036)

<220>

<221> trình tự tăng cường

<222> (4044)..(4237)

<223> eFMV

<220>

<221> trình tự tăng cường

<222> (4244)..(4536)

<223> e35S

<220>

<221> vùng khởi động

<222> (4544)..(8537)

<223> Ubi

<220>

<221> gen

<222> (8538)..(9995)
<223> EPSPS

<220>
<221> trình tự kết thúc
<222> (10009)..(11008)
<223> Ubi

<220>
<221> gen
<222> (11028)..(15538)
<223> GRMZM2G062320

<220>
<221> vùng khởi động
<222> (11028)..(12596)
<223> pPGM

<220>
<221> exon
<222> (12597)..(12834)
<223> PGM exon1

<220>
<221> Intron
<222> (12835)..(12974)
<223> PGM intron1

<220>
<221> exon
<222> (12975)..(13198)
<223> PGM exon2

<220>
<221> Intron
<222> (13199)..(13295)
<223> PGM intron2

<220>
<221> exon
<222> (13296)..(13382)
<223> PGM exon3

<220>
<221> Intron
<222> (13383)..(13718)
<223> PGM intron3

<220>
<221> exon
<222> (13719)..(13787)
<223> PGM exon4

<220>
<221> Intron
<222> (13788)..(14102)
<223> PGM intron4

<220>
<221> exon
<222> (14103)..(14204)
<223> PGM exon5

<220>

<221> Intron
 <222> (14205)..(14304)
 <223> PGM intron5

<220>
 <221> exon
 <222> (14305)..(14424)
 <223> PGM exon6

<220>
 <221> Intron
 <222> (14425)..(14492)
 <223> PGM intron6

<220>
 <221> exon
 <222> (14493)..(14540)
 <223> PGM exon7

<220>
 <221> trình tự kết thúc
 <222> (14541)..(15538)
 <223> tPGM

<400> 31
 cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagttttaaac tggcactagc ctaacggtgt tgactaacta ggccgcttcc 180
 ctaattagct aaccgcgggg cgcgccggga cccgaaagta gcaaacaaca ggttcatgtg 240
 cactataaaa agacaaaatt ctcgagtttc atcttttatt ccacataagc cttatatattt 300
 ccatttttcat atgattttta gtttaagttt gtgtcttaac tttttcgta atacgtaatt 360
 ctatgcatta tggatgcgtg aagtattttt gtttaaaaaa atgaaatgtc aaaatacgtt 420
 ttgtgatcta ttccatggtt ttcacctaac aggtgggttt tactatatat tctgccataa 480
 ctctagcctt agatgtaaat cgaaaaaaaa tgagagatga gotggagata gccttagatg 540
 aagcgtctga aatataaaag aaagagtaat gttgaacgca gtaggtgtag cagctgtagt 600
 tccatctcta ggaaaggga ctgcaatccg ggctccgggc ctgcgcgaat ctggcctgtc 660
 gtgtagatgc agccctgtcc atgacggccc aagcaacgcc cgcggctctc gatccaccac 720
 ggaaccact ccgacacaca ctgacacaca catgctggat gtggatgtgc tgtccaatta 780
 ttagtagcaa ttcggttaggc acaggcacgt actggcgggt gtttttagctg taagtaccga 840
 accaatcacg gttaagaacc gattaatccg tgcccagccg ccgagtgcgt tcgtacgtgc 900
 atcggatgca ctgcatgaat tgagagcatc atcatatcat acgcaggagt agtacgacgc 960
 cgctgctgtc ttgtccgggt aatgctttgc tcacagatta gtccatcgcc cacggtcgggt 1020
 gtggtgtgga tcgtgatgc cactgctttt tgtttggttt ttattccctt gataatcctc 1080
 cgcgtccctg aatgtatcta tttattttca ttccgaaatc cttttcacga aaaagaaaac 1140
 gaataaaaag agagttacga atacgcttcc ggcggccac atcaccttcc agcgaacatc 1200

ggcgcgcgcgt gacgtgtcgc ccacgcggc cgtccatata gccatccgac gaccgtggaa 1260
 gctggcagcg gccgctccgt tccgtcgaag gggcagggtca gtcagggtcac ccacacggcc 1320
 acacccgcgc ggggggatacg cgggtggaaaa cccggcgacc acatcaaaac acgaggcgctc 1380
 tccgcagga ctgggtcactc ggcacgcagg cagaggcagc acagcagcag ccagctccat 1440
 ccacctcttt tccctctctc gcttcgcttc ctggcgcat tctctctccc tcggccgtcc 1500
 ccgtcccttt ctgcgcgcg ccagctcgcc cgagttggta cgcctatcct ctgctgtact 1560
 ccccccgttc ctgctgtgcg aaaccgagca cccaaaccct agcctaagct attcgcatcg 1620
 caaactctat tgagcgacgg atcgcgaaac gcgcgccgc gcattcggac gagacctcgt 1680
 ggcccgtaac ttctactct accgtcgtcg ccggagacca gtcaccgggt taggggttctg 1740
 ggattagggg cttgggggctt ggattcccg gggactccat acccggtggc cttagaaggg 1800
 gaagggggcc cagcgggtgtt ggtgggttca ctgtcagcga gcgagacggc ggggcaccgc 1860
 cgatcgggcg tcgctggaca tttgattgag ctgggcaggc gaaggcgtgg agcttgctct 1920
 tcgattggga ttagaggagg gcggaggtgg tatggtgggc ggctcggctg gctcaatgga 1980
 gccgtggcc ggtggtggg caggcgagat cacgctctct catagagcgt aatgggttgc 2040
 gtaacacact cctgtttgta tttcgatccg atctaataat taaagtttat ttacgtgggtg 2100
 cacatgtaga attttcccat gccactgct aacttgatt ggatttttgc caaaaaaacg 2160
 ggtgtgcacc caatcactta tagtgaaacc gttgtgtgtt tgatgtgctt ctaaccaagt 2220
 acaaatccag tgaaaacaca cctaccattc aacaatccac attttggttg caattatcag 2280
 cattctagaa aggtgcatgt gccattgat acattgcta ttggtgcagg caaaccacc 2340
 atggcctcct ccgagaacgt catcaccgag ttcatgcgt tcaagggtgc catggagggc 2400
 accgtgaacg gccacgagtt cgagatcgag ggcgagggcg agggccgccc ctacgagggc 2460
 cacaacaccg tgaagctgaa ggtgaccaag ggcggccccc tgccttcgc ctgggacatc 2520
 ctgtccccc agttccagta cggctccaag gtgtacgtga agcaccgcc cgacatcccc 2580
 gactacaaga agctgtcctt ccccgagggc ttcaagtggg agcgcgtgat gaacttcgag 2640
 gacggggcg tggcgaccgt gaccaggac tctccttcgc aggacggctg ctcatctac 2700
 aaggtgaagt tcatcggcgt gaacttcccc tcgcagggc ccgtgatgca gaagaagacc 2760
 atgggctggg aggcctccac cgagcgctg taccgccgcg acggcgtgct gaagggcgag 2820
 acccacaagg cctgaagct gaaggacggc ggccactacc tgggtggagt caagtccatc 2880
 tacatggcca agaagccgt gcagctgcc ggctactact acgtggacgc caagctggac 2940
 atcacctccc acaacgagga ctacaccatc gtggagcagt acgagcgac cgagggccgc 3000
 caccacctgt tctgttagga gctcgcata tgatcatgca tcatggactc ggctactac 3060
 tgtggatttg tatgccatta tagacttggg gctgtgaaag actgcttgat gatttgccgg 3120

tttgttgctg tgtaaaaaaa ggtcccttgg ctcccagaag accatgaagg ttcggatcta 3180
 tcatgtaatt ccttggtatc tgccaattat gtatggacta tggacatgtg ttgcgctgtt 3240
 caacttacta ctacaaataa gtaatcgata tgttcccttc ccatgtctcg gtgacaattg 3300
 tctggagaag cttaggggtc gtttgtttgg gattatgtct ggagaaactt attttaaact 3360
 aagtgtgagt tcaagttaag ttagattata taatctaggc agattataat tccaagcgaa 3420
 caggtcotta gtgttttttg aaaatcctag gtgttctttt ggctacattg ttgtgtgtgc 3480
 agatcccttg ttggtctgta agcgtgggga agtaagaatc gtccgtttct actgaagacc 3540
 tgctcgagtt aggcaccgag gatgccgga accaaacaga gcaatagtgt ctctgtgggc 3600
 acagtggagt gtgaatctgt gtgatgcaa tccgtcattt gtttagcaa atttccagcg 3660
 ttgcatgatg cagtttcttt aacacggact taagggaagg gaaaaaatg ttgagccagg 3720
 agatccctca atgtgttaga ctgacgtgat agccaactaa accacgacgc aatgttgtcg 3780
 ttaatgacaa aaaaactatt tgttcctaaa tccttggcga cattgcatgg ctgtctcatg 3840
 agataatggc ctcatctctt atttatctct tatttatagc cggaagtggc agtgaccctt 3900
 gcttgattgc tcgtatgcca tctcaagttc tcaaccgtgt cgagcagcca ttttcccatc 3960
 tcaagcgcat catcgtttcg tttgacctca tctgctatcc tgctcctagt gcaaatcaca 4020
 tgcgacagaa agtgtgcgga ccagctgct tgtggggacc agacaaaaaa ggaatgggtgc 4080
 agaattgtta ggcgcaccta ccaaagcaa ctttgccttt attgcaaaga taaagcagat 4140
 tcctctagta caagtggga acaaaataac gtggaaaaga gctgtcctga cagccactc 4200
 actattgcgt ttgacgaacg cagtgcgac cacaaaactc gagacttttc aacaaagggc 4260
 attatccgga aacctcctcg gattccattg ccagctatc tgtcacttta ttgtgaagat 4320
 agtgaaaaag gaaggtggct cctacaaatg ccatcattgc gataaaggaa aggctatcgt 4380
 tgaagatgcc tctgccgaca gtggtcccaa agatggacc ccacccacga ggagcatcgt 4440
 ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg atatctccac 4500
 tgacgtaagg gttgacgaac aatccacta tccttcaagc tttggaaaac caaaacaacg 4560
 aataagcaaa ctgcaggaaa agtatgcagt ggaaaccaac ccagattcgg acgataggaa 4620
 agtatcaagt gaatgatttg ccaggaaaag gagaggggta aaaaggggcg aagattttaga 4680
 agatctaaag cacaagaacc agagattaga ttgaacaata gggaacttgg agcatccttt 4740
 ttttcttcag ggaaaaactg aaaatccaaa ccatgttgag caaaaccgag tgggattgga 4800
 aacaaaaaaa cccgagataa agaaactcga gaaaaagcat gaaatcgaaa ccaacttcag 4860
 taaaacaaaa ggaggacaga aaagaaagtc ggaagctata aagaatacat taacattcag 4920
 tgaaacagca tgctgtcttc ttcttttttt atgcacaaca gagcatacat atataccttc 4980
 ccaggctgag gacttggcgg aggagagccg cggatagggt ggcggtgcag acggtctgga 5040

cgggcccga gacggagacg aacagcgggc ctttcctgcc caggcaccac gcttgaacg 5100
 ccaagcacgc gccaaaccgc gccccgccga ggacgacgat ccccgcgaca agcgtggcgt 5160
 cgatcctggg cgaccccaag ccgaggaacc tcccttccag gacgagccgt aggacggcgg 5220
 tgagagaggc acccgtcgcg gaggtggcgc agcacaaggt gagcagcgcg gggaaggcgg 5280
 cgagcgtggc ggccctgcagg acggtgacga gcgcgaagac ggtgacgccg gcgacgaggc 5340
 agcagcagcc gaggatccag tcgtaggagg aagccggacc aaaccgggca atgcaacctg 5400
 cagatgcact agacggaggt aacgaggagg aggagaaaac agagcaagag caggcggaga 5460
 gaagatagag caaaacacga gtgaggcaca gcgtaagcac tcggtagaag tctccagagg 5520
 cgaggtgcmc acaggagaac agatgagtaa agtcagccaa ggatccacga tccaacggct 5580
 acgaattttt ggagtgcgt ggataggctc aaaggcgcca tttccatccg gctttatagt 5640
 attttaaaaa aattcatttt cctccctcta gtgtgtgcgg aggcgtgagc ccgtttaacg 5700
 gcgttgagaa gtctaacgga caccaaccac aaccaggaac cagcgccggc cgcgccgccg 5760
 agtgaagcag actgcatacg gcacggcgcg gcctctctct ggctgcctct cgagagttcc 5820
 gccccacct tcccgcgta gcgtggtggg ttgcgtttcc gctgtcggca tccggaagtt 5880
 gcgtggcaga gtggacggag acgaggccgg gtccctccagc tcctctcaaa cgtcacggca 5940
 ccggcatccg gcagccagcg cggtccttcc caaccactcg tcccaaccc atccccctc 6000
 ctgccccgcc gtcataaata gccagcccca tcccagctt ctttcccca cctcatcttc 6060
 tctccttttg ctctgaacgc acacaccgcc cggctctccga tctccgatcc ccgatccct 6120
 cgtcgatcct aggtacggcg accatcctac cccccccccc cccccctctc tctctgcctt 6180
 ctctagatcg gcgatccgat ccatgcttac ttggttaggg cctgctaact atgttcatgt 6240
 ttgcgttaga tccgtgcatg gacgcgatct gtacacacca gacgcgttct gattgctagc 6300
 taactcgcca gtacctggga atcctgggat ggctgtagcc ggccccgcac gcagacggga 6360
 ccgatttcat gattctctat ttttttcttt gtttgcgttg ctagggtttc gttcgatcga 6420
 tccgcgttat tctttatttc catatattct ggtaogatgt tgatacgggt cgaccgtgct 6480
 gcttacgttc tgtgcgcttg tttgcgggt catttttaac ttgccttttt tgtatggttt 6540
 ggttggtggc atgtggtctg gtcgggctgt cgttctagat cggagtagag tgctgtttca 6600
 aactgtctag cggatctatt agatttggat ctgcatgtgt gacatatatc ttcgtagtta 6660
 agatgatgca tctgtatgtg tgacatgcgg atctattaga tttggatctg tatgtgtgac 6720
 atatatcttc gtagttgaga tgatgcatct gtatgtgtga catatatott cgtagttaag 6780
 attatgcatg gaaatatcaa tcctttagat aaggacgggt atacttggtg ctgtggggtt 6840
 tactggtact tcgatagatg catatacatg atctaacatg cttagatata tgaagtaaca 6900
 tgctgctacg gtttaataat tottgagttg atttttactg gtacttagat agatgtatat 6960

acatgcttag atacatgaag taacatgctc ctacagttcc tttaatcatt attgagtacc	7020
tatatattct aataaatcag tatgttttaa attattttga ttttactggg acttagatag	7080
atgtatatat acatgctcaa acatgcttag atacatgaag taacatgctg ctacggttta	7140
gtcattattg agtgctata tattctaata aatcagtatg ttttaaatta ttttgatttt	7200
actggtactt agatagatgt atatatacat gctcaaacat gcttagatac atgaagtaat	7260
atgctactac ggtttaattg ttcttgagta cctatatatt ctaataaatc agtatgtttt	7320
aaattatttc gattttactg gtacttagat agatgtatat atacatgctt agatacatga	7380
agtaacatgc tactacggtt taattgttct tgaataccta tatattctaa taaatcagta	7440
tgttttaaata tatttcgatt ttactggtag ttagatagat gtatatatac atgctcgaac	7500
atgcttagat acatgaagta acatgctaca tatatattat aataaatcag tatgtcttaa	7560
attattttga ttttactggg acttagatag atgtatatatac atgctcaaac atgcttagat	7620
acatgaagta acatgctact acggtttaata cattattgag tacctatata ttctaataaa	7680
tcagtatgtt ttcaattgtt ttgattttac tggtagcttag atatatgtat atatacatgc	7740
tcgaacatgc ttagatacgt gaagtaacat gctactatgg ttaattgttc ttgagtacct	7800
atatattcta ataaatcagt atgtttttaa ttatttcgat tttactggta cttagataga	7860
tgtatatata catgctcgaa catgcttaga tacatgaagt aacatgctac tacggtttaa	7920
tcgttcttga gtacctatat attctaataa atcagtatgt cttaaattat cttgatttta	7980
ctggtagctta gatagatgta tatacatgct tagatacatg aagtaacatg ctactatgat	8040
ttaatcggtc ttgagtacct atatattcta ataaatcagt atgtttttaa ttattttgat	8100
tttactggta cttagataga tgtatatata catgctcgaa catgcttaga tacatgaagt	8160
aacatgctac tacggtttaa tcattcttga gtacctatat attctaataa atcagtatgt	8220
ttttaattat tttgatatta ctggtagctta acatgttttag atacatcata tagcatgcac	8280
atgctgctac tgtttaatca ttctggaata cctatatatt ctaatatatc agtatgtctt	8340
ctaattatta tgattttgat gtacttgtat ggtggcatat gctgcagcta tgtgtagatt	8400
ttgaataccc agtgtgatga gcatgcatgg cgccttcata gttcatatgc tgtttatttc	8460
ctttgagact gttctttttt gttgatagtc accctgttgt ttggtgattc ttatgcagat	8520
ccagatcttc gtaaaccatg gctctgggtg ctaggccagt gctgtctgct agagtggctg	8580
cttctaggcc aagggtggcc gctagaaagg ctgtgagggt gagcgctgcc agtcacggcg	8640
ctagctctag gccagccacc gctaggaagt ccagcggcct gagcggcacc gtgcgcatcc	8700
caggcgacaa gagcatcagt caccgcagct tcatgttcgg cggcctggcc agcggcgaga	8760
ctaggatcac cggcctgctc gagggcgagg acgtgatcaa caccggcaag gccatgcagg	8820
ctatggggcg tcgcatccgc aaggagggcg acacctggat catcgacggc gtgggtaatg	8880

gcggcctgct ggccccagag gccccactgg acttcggcaa cgccgccacc ggctgcaggg 8940
 tgactatggg cctgggtgggc gtgtacgaact tcgacagcac ctttatcggc gacgccagcc 9000
 tgaccaagag gccgatgggc aggggtgctga acccgctgcg cgagatgggc gtgcaggtca 9060
 agagcgagga cggcgacagg ctgccgggtga ccctgagggg cccaaagacc ccgacccccga 9120
 tcacctaccg cgtgccaatg gccagcgcgc aggtgaagtc agccgtgctg ctggccggcc 9180
 tgaacacccc aggcatacc accgtgatcg agccgatcat gacccgcgat cacaccgaga 9240
 agatgctgca gggcttcggc gccaacctga ccgtcgagac tgacgctgac ggcgtgagga 9300
 ccatcaggct cgagggcagg ggcaagctga ccggccagggt gatcgacgtg ccgggcgacc 9360
 caagcagcac cgccttccca ctgggtggccg ccctgctggt gccaggctct gacgtgacca 9420
 tcctgaacgt gctgatgaac ccgagccgca ccggcctgat cctgaccctg caggagatgg 9480
 gcgccgatat cgaggtgatc aacccgaggc tggctggcgg cgaggacgtc gccgacctga 9540
 ggggtgaggtc cagcaccctg aagggcgtga ccgtgccaga ggacagggcc ccgagtatga 9600
 ttgacgagta cccgatcctg gccgtggccg ctgccttcgc cgagggcgcc accgtgatga 9660
 acggcctcga ggagctgcgc gtgaaggaga gcgacaggct gagcgctgtg gccaacggcc 9720
 tgaagctgaa cggcgtggac tgcgacgagg gcgagactag cctggtggtg aggggcaggg 9780
 cagacggcaa gggcctgggc aacgcttccg gcgctgccgt ggccaccac ctggatcaca 9840
 ggatcgccat gtcgttccctc gtgatgggcc tcgtgagcga gaaccgggtg accgtggacg 9900
 acgccaccat gatcgccacc agcttccccg agttcatgga cctgatggcc ggccctggcg 9960
 ctaagatcga gctgagcgac accaaggccg cctgatcctc tagagctcgc caaggttcaa 10020
 ttaagctgct gctgtacctg ggtatctgcg tcgtctggtg ccctctggtg tacctctata 10080
 tggatgtcgt cgtctaataa acatctgtgg tttgtgtgtc atgaatcgtg gttgtggctt 10140
 cgttggttta atggacctgt tgtgtccctc gtgttgtaac caaaactctt ctgcagcagt 10200
 atggcttgaa tccttatgaa gtttgatatt tgaacttaaa agtctgctca ttatgttttt 10260
 ttctggttat atctcctaataa taactgacct ggatcaaatt tgattcgctg gtgtttattg 10320
 gacccctccc aggttcttgc tttctaccgt ttcttgctga atgttaactt gattctgtca 10380
 ggctcagttt cccactatgg cttacagctt aacgtgtttg gtttggtgaa tgttaacttg 10440
 gttttgtcaa gctcagtttt ttactctggc ttacagcata acatgtttga cttttggttt 10500
 tgctgctttg ttattgggtt ctgggtagtt cttgatgaat ccaaaagatc atgtgcacag 10560
 ccataattatc tatttaagcg atccaggtta ttactatgaa aggatgcctt ctagctaagg 10620
 agtagttagg tttttctctc aagggttaaata tttctcgatg ctctagtgtt cctgtgacca 10680
 taatcataat aattcctttg aaagctctat ggtccctgga agcagggcat acaatgcaag 10740
 acagcaactt gatcacatca actgaagtat acagggttct ctttaactctt ggtgacttcg 10800

gtttaaatgga ccggttgtac tcgtgttota tccgtaaccg ttgtgatgtc ttgtgtgttt 10860
 gggtgcggga tagctgggac cagcagcttt cgtctaat ctgatggata gctatagacg 10920
 gcaatgagat gggttatatta taacctctga tcctgaactc tacgagatcg tctcatccgt 10980
 cattgccacc aaatacacca ttaaattact aattagctaa cggacccctt gcctcggtta 11040
 caattgagtt tgtcataaga aagaatgaga aagaaaagaa gcaatccaag cgcaagagct 11100
 caaaggaaca caagtcactc tctcactagt cactattgat tggaattgaa ctagggactt 11160
 gggagaggat tttatctctt tgggtgtgtc tgtattgaat gctatagctc ttgtaatgtg 11220
 tagaatgttg gaaacttga tgccattgaa tgtggggtgg ttgggggtatt tatagcccca 11280
 accaccacaa tgtggtcgtt ggaagtctgc tgtcgcattg cgcaccggat agtccgatgc 11340
 gccaccggac actgtccggt gcgccagcca cgtcagccag ccgttggggg tcgaccgttg 11400
 gagctctgac ttgtggggcc tctgggctgt ccggtggtgc accggacagg tcctgtagac 11460
 tgtccggtgc accaactgcg cgtgctctga cctctgcgcg cgcaggcgcg cattaaatgc 11520
 gttgcagtcg accgttgcgc ttgaagtagt cgttgcctcg ctggcacacc ggacagtgtc 11580
 cgggtgaatta tagcggagcg gcctccattt tcccgaagg agcgagttca gcgtcgagtt 11640
 cctggtgca ccggacactg tccggtggca caccggacag tccggtgcgc cagaccaggg 11700
 cacacttcgg ttgtctttgg ctctctttgt ttgaaccctt tcttgggtct tttattggtc 11760
 tattgtgaac ctttggcacc tgtaaaactt ataactaga gtaaaactagt tagtccaatt 11820
 atttgtgttg ggcaattcaa ccacccaaat caattaggaa aaggtgtaag cctatttccc 11880
 tttcaggagg cgtacgtgag agggagaggt gaaaaggaac aacgcgtata ccagataagg 11940
 tcccacagcc taagtaggta gccttcta atctctacta actattaagg agagagtgt 12000
 gactgcccc gctccctacc caacgcccc cgctacctgt tcaccgcgcg ccagcgaaac 12060
 ctccgcacgc ccaactgcca tctgttcccc gtgcgccagc gaaacatccg cagccccgcg 12120
 gccgcctgt tccccgcga tcccgctgca cgacttctgc taccgcaacg gccaccacg 12180
 cagccccgc tgttcaccgc gcatcccgct gacctccct tcacgctcgc acacgctccg 12240
 ttccccacc ccaccgcaat ccccgacgct ataagagcgg taaccaactc catctccctg 12300
 gtgccacgca ttgttgagtt ctttaagggtgc gtttcgttga ggacttggtc atttttgttg 12360
 gtcattgtatt ccattttact gctctaccat tttgtggaat aaagggagga atgttttcac 12420
 tagaagagtt catcaatctt atgttggttt cttggatcag ttttgcctta tggctaaatg 12480
 gtcgaattga gcctatttca ttataaagtt agcgagcgaa taattgttca gcctcttcct 12540
 agaactcatt accagtagaa tcagttacta actgcttttc tttttcttgg attaga atg 12599
 Met
 1
 got ggg got atc tct cac cat gcg cta gca ttt tca caa tcc cac tgg 12647

Ala Gly Ala Ile Ser His His Ala Leu Ala Phe Ser Gln Ser His Trp
5 10 15

tgc agt gcg aag aac tct aga ttc gga aag agg acg ggc aat gct cgc 12695
Cys Ser Ala Lys Asn Ser Arg Phe Gly Lys Arg Thr Gly Asn Ala Arg
20 25 30

ctg gtt tat cta aaa gga aga tgt ggt tca ggc agc aga aaa ctg ggt 12743
Leu Val Tyr Leu Lys Gly Arg Cys Gly Ser Gly Ser Arg Lys Leu Gly
35 40 45

ttg atg tgg gcc tcg agc tcg cag tct tct gtc atg gag ccg acg cac 12791
Leu Met Trp Ala Ser Ser Ser Gln Ser Ser Val Met Glu Pro Thr His
50 55 60 65

cta cca tct gat ggc aac agc agc cac acc cca aaa aaa tca a 12834
Leu Pro Ser Asp Gly Asn Ser Ser His Thr Pro Lys Lys Ser
70 75

gtaatttttaa cgacctccta tgggtggttat ttgttttttaa tttgagaaaa ctatccattt 12894

gacacattta actttgggct tctcagaatt tgggggcata taataagatc tgctaattctg 12954

ttatctctat gtcgtttag gt gaa agc gct ctt ata ttg att tgg cat ggt 13006
Ser Glu Ser Ala Leu Ile Leu Ile Trp His Gly
80 85 90

gaa tcc ctg tgg aac gag aaa aat cta ttt cct ggc tgc atc gat gta 13054
Glu Ser Leu Trp Asn Glu Lys Asn Leu Phe Pro Gly Cys Ile Asp Val
95 100 105

ccc ctg aca ccg aag ggt gtt gag gag gcc att gag gca ggt aaa agg 13102
Pro Leu Thr Pro Lys Gly Val Glu Glu Ala Ile Glu Ala Gly Lys Arg
110 115 120

ata tgc aat atc cca atc gat gtg ata tat act tca tca ctg att tgt 13150
Ile Cys Asn Ile Pro Ile Asp Val Ile Tyr Thr Ser Ser Leu Ile Cys
125 130 135

gct cag atg acc gca atg ctt gcc atg atg cag cat cga cgc aag aag 13198
Ala Gln Met Thr Ala Met Leu Ala Met Met Gln His Arg Arg Lys Lys
140 145 150

gtttgtgtct ttcctttgaa attccagtaa tttcttctag catttgtag aacttgccgg 13258

agaaatcatg ctttgcctgt gatatatgta tttatag atc cta gtt atc acg cat 13313
Ile Leu Val Ile Thr His
155 160

aat gag agt gaa caa gct cac agg tgg agt cag ata tac agt gag gag 13361
Asn Glu Ser Glu Gln Ala His Arg Trp Ser Gln Ile Tyr Ser Glu Glu
165 170 175

aca atg aaa cag tcc att cct gtcacacag cttggcaatt gaatgaacgg 13412
Thr Met Lys Gln Ser Ile Pro
180

atgtaatact ttctccatac tctttgattt gctaattact ccctctgtct caaaatagta 13472

ttaatttttag ctcttgattt ttatgtctat attcaaatag atgatgataa atctagattc 13532

tagacacaaa tataaaacat atacatcaag tattatatga atctattaat ttactaagac 13592

caatttttaat ttggggacaga gggagtatac gattataata gttgtttgac tgtgcttctc 13652

50

```

tttcataatc catttgtaat cctggccata cataaatttt gttattttta ttttagctgtt 14800
tcactactac attgcaacca tcagtatcat gcagacttcg atatatgtca cgatttgcat 14860
agtcttatca ttgaagagca cgtgtcacac ctaccggtag aagttccctc gtacattgtc 14920
agggggacct catggctgtc gctaaagccg gcgactctct ggttggtctg ggcaccgcat 14980
ccgacgacaa caccgtcacg ctgtccagct catcatccac ctgaagatca acctgccacg 15040
taagtcggta gtgaccggcc atgaattctt gtgcgctggg gcaacgggcg ttgcattccg 15100
ttgtgtgagt gttttagaac aagatgtatg taaaggagta gcacatccgg tggtgcaatg 15160
gctagggtga ctacctgct gatgagctcg cggtgcaact cgtttagcag cccagccaag 15220
agtcacgggt acttgccatg tcgtgctgtg tcgacgacta ctatatcaag gactgtggcg 15280
tcctcttggc gccagagggt aggagtaccg acagcaatga cctctcgcca tcgtcggggg 15340
agcttttgtg ctggcccgc ttaattctc ttcgtaaaca catacacttg cttttttgt 15400
ttaagcaagt gagtatagt gtgcgtacct gatgactgac aatgtacgag gaaacttcta 15460
ctggcagggt tggcacgtgc tcttcaatga tgagaccatg taaatcgtga catatatcga 15520
agtctgcatg atactgatcg gaccgcctgc agggccgggg gcgcgcccta attagctaac 15580
ggccaggatc gccgcgtgag cctttagcaa ctagctagat taattaacgc aatctgttat 15640
taagttgtct aagcgtcaat ttgtttacac cacaatatat cctgccacca gccagccaac 15700
agctccccga ccggcagctc ggcacaaaat caccactcga tacaggcago ccatcag 15757

```

<210> 32
 <211> 15355
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Casset biểu hiện của cấu trúc 22503

<220>
 <221> trình tự tăng cường
 <222> (214)..(407)
 <223> eFMV

<220>
 <221> trình tự tăng cường
 <222> (414)..(706)
 <223> e35S

<220>
 <221> vùng khởi động
 <222> (714)..(4707)
 <223> Ubi

<220>
 <221> gen
 <222> (4708)..(6165)
 <223> EPSPS

<220>
 <221> trình tự kết thúc
 <222> (6179)..(7178)
 <223> Ubi

 <220>
 <221> vùng khởi động
 <222> (7198)..(9285)
 <223> pPLA2m

 <220>
 <221> vòng gấp
 <222> (9298)..(10212)
 <223> exon2 đích của PLA2

 <220>
 <221> trình tự kết thúc
 <222> (10223)..(11218)
 <223> tPLA2

 <220>
 <221> vùng khởi động
 <222> (11226)..(13349)

 <220>
 <221> gen
 <222> (13352)..(14131)

 <220>
 <221> đặc điểm hỗn tạp
 <222> (14039)..(14128)

 <220>
 <221> trình tự kết thúc
 <222> (14138)..(15136)

<400> 32
 cctgtggttg gcatgcacat acaaattggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taaccgocaa tatatcctgt 120
 caaacactga tagttttaaac tggcactagc ctaacgggtg tgactaacta ggccgcttcc 180
 ctaattagct aaccggggg cgcgccggga ccagctgct tgtggggacc agacaaaaaa 240
 ggaatggtgc agaattgtta ggcgcaccta caaaagcaa ctttgccttt attgcaaaga 300
 taaagcagat tcctctagta caagtgggga aaaaataac gtggaaaaga gctgtcctga 360
 cagcccactc actattgcgt ttgacgaacg cagtgcgcac cacaaaactc gagacttttc 420
 aacaaagggt attatccgga aacctcctcg gattccattg ccagctatc tgtcacttta 480
 ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccattcattgc gataaaggaa 540
 aggctatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc ccaccacga 600
 ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg 660
 atatctccac tgacgtaagg gttgacgaac aatcccacta tccttcaagc tttggaaaac 720
 caaaacaacg aataagcaaa ctgcaggaaa agtatgcagt ggaaaccaac ccagattcgg 780

acgataggaa agtatcaagt gaatgatttg ccaggaaaag gagaggggta aaaaggggcg 840
 aagattttaga agatctaaag cacaagaacc agagattaga ttgaacaata gggaacttgg 900
 agcatccttt ttttcttcag ggaaaaactg aaaatccaaa ccatgttgag caaaaccgag 960
 tgggattgga aaccaaaaaa cccgagataa agaaactcga gaaaaagcat gaaatcgaaa 1020
 ccaacttcag taaaacaaaa ggaggacaga aaagaaagtc ggaagctata aagaatacat 1080
 taacattcag tgaacacagca tgctgtcttc ttcttttttt atgcacaaca gagcatacat 1140
 atataccttc ccaggctgag gacttggcgg aggagagccg cggatagggtt ggcggtgcag 1200
 acggtctgga cgggcccga gacggagacg aacagcgggc ccttcctgcc caggcaccac 1260
 gcttgaacg ccaagcacgc gccaaaccgc gccccgccga ggacgacgat ccccgcgaca 1320
 agcgtggcgt cgatcctggg cgacccaag ccgaggaacc tcccttcag gacgagccgt 1380
 aggacggcgg tgagagaggc acccgtcgcg gaggtggcgc agcacaaggt gagcagcgcg 1440
 gggaaggcgg cgagcgtggc ggctgcagg acggtgacga gcgcgaagac ggtgacgccg 1500
 gcgacgaggc agcagcagcc gaggatccag tcgtaggagg aagccggacc aaaccgggca 1560
 atgcaacctg cagatgcact agacggaggt aacgaggagg aggagaaaac agagcaagag 1620
 caggcggaga gaagatagag caaaacacga gtgaggcaca gcgtaagcac tcggtagaag 1680
 tctccagagg cgaggtgcgc acaggagaac agatgagtaa agtcagccaa ggatccacga 1740
 tccaacggct acgaattttt ggagtgcgt ggataggctc aaaggcgcca tttccatccg 1800
 gctttatagt attttaaaaa aattcatttt cctccctcta gtgtgtgcgg aggcgtgagc 1860
 ccgtttaacg gcgttgagaa gtctaacgga caccaaccac aaccaggaac cagcgccggc 1920
 cgcgccgcgg agtgaagcag actgcatacg gcacggcgcg gcattctctt ggctgcctct 1980
 cgagagtacc gccccacct tcccgcggta gcgtgggtgg ttcgctttcc gctgtcggca 2040
 tccggaagtt gcgtggcaga gtggacggag acgaggccgg gtccctccagc tcctotcaaa 2100
 cgtcacggca ccggcatccg gcagccagcg cggtccttcc caaccaactcgt ttcccaaccc 2160
 atcccccttc ctgcgccgcc gtcataaata gccagcccca tcccagctt ctttcccaa 2220
 cctcatcttc tctccttttg ctctgaacgc acacaccgcc cggctctccga tctccgatcc 2280
 ccgatccctt cgtcgatcct aggtacggcg accatcctac ccccccccc cccctctc 2340
 tctctgcctt ctctagatcg gcgatccgat ccatgcttac ttggttaggg cctgctaact 2400
 atgttcatgt ttgcgttaga tccgtgcatt gacgcgatct gtacacacca gacgcgttct 2460
 gattgctagc taactcgcca gtacctggga atcctgggat ggctgtagcc ggccccgcac 2520
 gcagacggga ccgatttcat gattctctat tttttctttt gtttcgttgc ctagggtttc 2580
 gttcgatcga tccgcgttat tctttatttc catatattct ggtacgatgt tgatacggtt 2640
 cgaccgtgct gcttacgttc tgtgcgcttg tttgccgggt catttttacc ttgccttttt 2700

tgtatggttt ggttgtggcg atgtggtctg gtcgggctgt cgttctagat cggagtagag	2760
tgctgtttca aactgtctag cggatctatt agatttggat ctgcatgtgt gacatatatc	2820
ttcgtagtta agatgatgca tctgtatgtg tgacatgcgg atctattaga tttggatctg	2880
tatgtgtgac atatatcttc gtagttgaga tgatgcatct gtatgtgtga catatatctt	2940
cgtagttaag attatgcatg gaaatatcaa tcctttagat aaggacgggt atacttgttg	3000
ctgtggggttt tactgggtact tcgatagatg catatacatg atctaacatg cttagatata	3060
tgaagtaaca tgctgctacg gtttaataat tcttgagttg atttttactg gtacttagat	3120
agatgtatat acatgcttag atacatgaag taacatgctc ctacagttcc tttaatcatt	3180
attgagtacc tatatatctt aataaatcag tatgttttaa attattttga ttttactggt	3240
acttagatag atgtatatat acatgctcaa acatgcttag atacatgaag taacatgctg	3300
ctacgggttta gtcattattg agtgcctata tattctaata aatcagtatg ttttaaatta	3360
ttttgatttt actggtactt agatagatgt atatatacat gctcaaacat gcttagatac	3420
atgaagtaat atgctactac ggtttaattg ttcttgagta cctatatatt ctaataaatc	3480
agtatgtttt aaattatttc gattttactg gtacttagat agatgtatat atacatgctt	3540
agatacatga agtaacatgc tactacggtt taattgttct tgaataccta tatattctaa	3600
taaatcagta tgttttaaat tatttcgatt ttactggtag ttagatagat gtatatatac	3660
atgctcgaac atgcttagat acatgaagta acatgctaca tatatattat aataaatcag	3720
tatgtcttaa attattttga ttttactggt acttagatag atgtatatatac atgctcaaac	3780
atgcttagat acatgaagta acatgctact acgggttaat cattattgag tacctatata	3840
ttctaataaaa tcagtatgtt ttcaattgtt ttgattttac tggtagcttag atatatgtat	3900
atatacatgc tcgaacatgc ttagatacgt gaagtaacat gctactatgg ttaattgttc	3960
ttgagtacct atatattcta ataaatcagt atgtttttaa ttatttcgat tttactggta	4020
cttagataga tgtatatata catgctcgaa catgcttaga tacatgaagt aacatgctac	4080
tacgggttaa tcgttcttga gtacctatat attctaataa atcagtatgt cttaaattat	4140
cttgatttta ctggtactta gatagatgta tatacatgct tagatacatg aagtaacatg	4200
ctactatgat ttaatcgttc ttgagtacct atatattcta ataaatcagt atgtttttaa	4260
ttattttgat tttactggta cttagataga tgtatatata catgctcgaa catgcttaga	4320
tacatgaagt aacatgctac tacgggttaa tcattcttga gtacctatat attctaataa	4380
atcagtatgt ttttaattat tttgatatta ctggtactta acatgcttag atacatcata	4440
tagcatgcac atgctgctac tgtttaatca ttogtgaata cctatatatt ctaatatatc	4500
agtatgtctt ctaattatta tgattttgat gtacttgat ggtggcatat gctgcagcta	4560
tgtgtagatt ttgaataccc agtgtgatga gcatgcatgg cgccttcata gttcatatgc	4620

tgtttatttc ctttgagact gttctttttt gttgatagtc accctgttgt ttggtgattc	4680
ttatgcagat ccagatcttc gtaaaccatg gctctgggtg ctaggccagt gctgtctgct	4740
agagtggctg cttctaggcc aaggggtggc gctagaaagg ctgtgagggg gagcgctgcc	4800
agtcacggcg ctagctctag gccagccacc gctaggaagt ccagcggcct gagcggcacc	4860
gtgcgcattcc caggcgacaa gagcatcagt caccgcagct tcatgttcgg cggcctggcc	4920
agcggcgaga ctaggatcac cggcctgctc gagggcgagg acgtgatcaa caccggcaag	4980
gccatgcagg ctatgggctc tcgcatccgc aaggagggcg acacctggat catcgacggc	5040
gtgggtaatg gcggcctgct ggccccagag gcccactgg acttcggcaa cgccgccacc	5100
ggctgcaggc tgactatggg cctgggtggc gtgtacgact tcgacagcac ctttatcggc	5160
gacgccagcc tgaccaagag gccgatgggc aggggtgctga acccgctgcg cgagatgggc	5220
gtgcaggtca agagcgagga cggcgacagg ctgccggtga ccctgagggg cccaaagacc	5280
ccgaccccga tcacctaccg cgtgccaatg gccagcgccc aggtgaagtc agccgtgctg	5340
ctggccggcc tgaacacccc aggcattacc accgtgatcg agccgatcat gaccgcgat	5400
cacaccgaga agatgctgca gggcttcggc gccaacctga ccgtcgagac tgacgctgac	5460
ggcgtgagga ccatcaggct cgagggcagg ggcaagctga ccggccagggt gatcgacgtg	5520
ccgggcgacc caagcagcac cgccttccca ctgggtggcg ccctgctggt gccaggctct	5580
gacgtgacca tcctgaacgt gctgatgaac ccgagccga ccggcctgat cctgacctg	5640
caggagatgg gcgccgatat cgaggtgatc aaccgaggc tggctggcgg cgaggacgtc	5700
gccgacctga gggtgaggtc cagcaccctg aagggcgtga ccgtgccaga ggacagggcc	5760
ccgagtatga ttgacgagta cccgatcctg gccgtggcg ctgccttcgc cgagggcgcc	5820
accgtgatga acggcctcga ggagctgcgc gtgaaggaga gcgacaggct gagcgctgtg	5880
gccaacggcc tgaagctgaa cggcgtggac tgcgacgagg gcgagactag cctggtggtg	5940
aggggcaggc cagacggcaa gggcctgggc aacgcttcgc gcgctgccgt ggccaccac	6000
ctggatcaca ggatcgccat gtcgttcctc gtgatgggccc tcgtgagcga gaacccgggtg	6060
accgtggacg acgccaccat gatcgccacc agcttccccg agttcatgga cctgatggcc	6120
ggcctgggcg ctaagatcga gctgagcgac accaaggccg cctgatcacc tagagctcgc	6180
caaggttcaa ttaagctgct gctgtacctg ggtatctgcg tcgtctggtg ccctctgggtg	6240
tacctctata tggatgtcgt cgtctaataa acatctgtgg tttgtgtgtc atgaatcgtg	6300
gttgtggctt cgttggttta atggacctgt tgtgtcctct gtgttggtacc caaaaactctt	6360
ctgcagcagt atggcctgaa tccttatgaa gtttgatatt tgaacttaaa agtctgctca	6420
ttatgttttt ttctggttat atctcctaataa taactgcctg ggatcaaatt tgattcgctg	6480
gtgtttattg gacccctccc aggttcttgc tttctaccgt ttcttgctga atgttaactt	6540

gattctgtca ggctcagttt cccactatgg cttacagctt aacgtgtttg gtttgttgaa 6600
 tgttaacttg gttttgtcaa gctcagtttt ttactctggc ttacagcata acatgtttga 6660
 cttttggttt tgctgctttg ttattgggtt ctgggtagtt cttgatgaat ccaaaagatc 6720
 atgtgcacag ccatattatc tatttaagcg atccaggta ttactatgaa aggatgcctt 6780
 ctagctaagg agtagttagg ttttttcttc aagggttaa tttctcgatg ctctagtgtt 6840
 cctgtgacca taatcataat aattcctttg aaagctctat ggtccctgga agcagggcat 6900
 acaatgcaag acagcaactt gatcacatca actgaagtat acagggttct cttaactctt 6960
 ggtgacttcg gtttaatgga ccggttgtagc tcgtgttcta tccgtaaccg ttgtgatgac 7020
 ttgtgtgttt ggttgcgga tagctgggac cagcagctt cgtctaatt ctgatggata 7080
 gctatagacg gcaactgagat gggtatatta taacctctga tcctgaactc tacgagatcg 7140
 tctcatcgt cattgccacc aaatacacca ttaaattact aattagctaa cggaccctat 7200
 ttgtactcat tccatgtctc ataaactttg ggcaccatcc atccaacaca tccaatctaa 7260
 acacacaaaa cgatggggaa tggaaagagc agtattcgat tcaacaatgg caaacaata 7320
 tcactgaatt agaccaagaa taaacctaat tagacaacga cctcccaacc atcattcgtc 7380
 aggctgtaaa gaagataaag ctgccttggg gcatggatca agcagaacac cagagatgaa 7440
 tccaaacaca cagaaaatca cgcgcgctgt ctacaatgac aacaagcccc acatttcatt 7500
 gcagtacact gggctacaaa ggcacgtaca acaaagagct agggaaacat tgcggagggc 7560
 acgagagagc agctaacttg acaatatagc agactgagct tgcaactgta gcaggcgagg 7620
 aagggaatca tggggacgga gaatggggtc catgccgcg aaggagaagg cggacgccgc 7680
 cacggtggca ccggcgcacg cgcacacagg gaaccgcac aggcagccaa ggatgctgcc 7740
 tcgccattgc gccggtcgtc tctgccacgc tcctctctct ccccgctgc atogccgtgg 7800
 atggggcaag cagagagcag ggactgcgac gatctgggcg gaggactcgc cttggagagc 7860
 gcggacgcag acgggattct agggagagag cgaagacggg gcgcgcgcgc cgctcgcgcg 7920
 gcgtggtggc ggcgagatta gcgggggttg ggggagggcg gagccgtggt gaggggtgtg 7980
 acgccctcct taccctctta agtagtagta gagatataat ccgttccaaa atatccatcc 8040
 gttcaattta tatttcgttt gatcttttta ccctaaattt gattgactca tottattaaa 8100
 aaagttcata actattatta atctttattg agatatcatt tagcatataa tatactttaa 8160
 gtgtgggttt agattttttt taaaaaaaaa aattcgcaaa aattaaatga aacgacccaa 8220
 tcaaacttga aaagtaaaac taattataaa tttgaacgga aggagtaaga ggatgtttga 8280
 atgtactaga gctaatagtt ggttgcttta aaatttgcta gtagaattag ctagctaata 8340
 aatatctaga taactattag ctaatttgct aaaacagcta atagttgaac tattagctag 8400
 attgtttgga tgtattcggc taattttaat ggctaactat tagctatagt acaatattca 8460

aacacctcct aattaaaatg gacaaatata ttttcttttg gtcccttgcg ttagattttt	8520
catatctcct tatttagtat aaaagaatca tcaaaaagtg gacaaccctt agtggaacac	8580
cattttagta gtggttgcat gaaacctttc ggcaccagtg ttctatgtgt cactctaaaa	8640
atgggacagc atgtacgtag tgcctatata tatacaagtc atctatcgtt gcctcctcag	8700
ttcatcacta atcacactta ttgtgccttc gacgagtatc tatagctagc tcattaatcg	8760
attcgggggt gtgttgctga aggcggcatt ggcgagctac tcgtcgcggc gtccaagcaa	8820
tacctgtagc acgaaggcga tcgccgggag cgtggtcggc gagcccgtcg tgctggggca	8880
gagggtgacg gtgctgacgg tggacggcgg cggcgtccgg ggtctcatcc cgggaacat	8940
cctcgccttc ctggaggcca ggctgcagga gctggacgca ccggaggcga ggctggcgga	9000
ctacttcgac tacatcgccg gaaccagcac cggcgggtctc atcaccgctt tgctgaccgc	9060
gccgggcaag gacaagcggc ctctctagtc tgccaaggac atcaaccact tttacatcca	9120
taactgcccg cgcattcttc ctccagaagt agtcggatgc tgccgccatt gttcttgcat	9180
ccatccagca tcgtacgtac gtctctata catctgcgga tcatcatgtg cgcattgttg	9240
tggcatgcat gcatgcatgt gagcaggagc aggccttgca aaaccccatg gagatctgaa	9300
caggcttgcg gccgccgtgt ccgcgctgag gaagccaaag tacaacggca agtgcatgcg	9360
cagcctgatt aggagcatcc tcggcgagac gaggggtcga ctccggagca tgccctttgt	9420
gcttatgcct ccgttctgcc ttctgacgaa tttgggtactg gaagcagatg agttttggtt	9480
cactatcatt ctgaatttac acctgcgctt gctgtcagac taggcaacca agtgactttt	9540
gtgactttga tcatgttcag tgtgtttcca agtcctaata aatcaaaaag aaaaacagtt	9600
tgttaacgat tgtttgccat gtctatataa taaagttgct tttatagtag cttagaattc	9660
aatcggccaa ctttatctcg tacgctgaca gtaaaggtag atttaaaagg tgacaatgga	9720
tagtctaata cttgaactga caatagagac acattacatg tcagttgatt aagtttgtaa	9780
cagaaaaata aacaatacta cataattgca aagtttcttt gatgtctttc tttcaagaaa	9840
cacaaatata tcaatgctac agtattgctg atgaatttat ccatgttgag atgtttttct	9900
ggtttctgat ctgatcagtc tcaattggtg tgctgtttca ttttcatttg ctgatgatcg	9960
tccgagtagt taattcttac taatatttag ataatttggc atacaagcga atcacgtaga	10020
acatgatact tttgaatgaa tttatcaaag ttttatcaat tggtaggttg tttcatgggt	10080
ttctactga tgtctcttct tcagatttct cgagccctcg tctcgccgag gatgctccta	10140
atcaggctgc gcatgcactt gccgttgtag tttggcttcc tcagcgcgga catggcggcc	10200
gcaagcctgc tcggatccat gggacaagtg gctttactgt cagtcacatg cttgtaaata	10260
agtagacttt attttaataa aacataaaaa tatatatatg ttcttgaata taaaattgat	10320
aaccaaatta aaattcgaac catcacttat acataatttt actttatttt ttataaaacg	10380

tgaacgggaa ggactaccgt gaatgactat agaaccaatc atactagtat aaaatatatg	10440
atgacactac gggagagaca aactttgtct ggcgctaaat attttgccga gtgtgaattc	10500
acggggcacta ggcaaagatc ttctttgccg agtgttacgc tgggcaaagt aagacactag	10560
gtaaatcagt catttgccga gtgtccgcca ctaggcaaag caaaacactg gcaaatcaaa	10620
agtttaccta gtgccagaca ctaggcaaaa aaaaaacgct cggcaaactc gaagtttccc	10680
tagtgccaga cactagacaa agaaaaacac ttgataaact agcgtcgtca gctaacacca	10740
tccaccaacc gttaacgttg ccgagtatct gacttcgaca ctcggaag aaggtctctt	10800
tgcctagtgt cggctcggaa cactaggcaa agaggcactt tacctagtgt cgtattttga	10860
cactcagtaa aataatTTTT tttctttctg cttccaaact ttttatgatg tgttcctata	10920
gcacctagaa ctacatgtca agttttggta aaatTTTT gaagttttgct atatttactt	10980
aatTTatTTT atttaattga atttcttttg ataattcaaa tttgaactcg gcaaggtaag	11040
aagcgagggt agcctggaaa cacactttgc ctagtgttac actcggta caaggcctccc	11100
ctgcctagtgt ctgcactcga caaaagattc gcctttgcct agcgtgcac tcggcacagg	11160
agtcgccttt gcctagtgt gcactaggca aagcctccgt taccgtgcct tccatcgtcg	11220
gaccctggag aggcattgaac ccattcatgta tactcctgga cttggaagag gaaatgtcaa	11280
ccaaagtga aggttttctt ctagtgtctg ctaagagata aattgaatac tagatctctc	11340
ctaagacatc ggggcattcg tctggactac tacacatgcg aaaattgcat cttacagtgg	11400
gaagaaacta tatctcacct cttcttgccg tgtaactttg ccgaagatg ttggctcact	11460
gttggaatca ctccgccccg aacgttggat ctagcgcttg cagtgtaca tattagagca	11520
agactgacat tgcgtggag aatggaaggt attataacca tgcattgtcat ggtacatatg	11580
gaaatgtcaa aataactgga ttttcgaaaa cataccgcca acggtggcgg tctgcaagga	11640
aatgttcaag attgaaatga actacatatg caaccagggt aagcttgaga caggagataa	11700
aagtagaaac tggatacaac actttgtaac atagcgacac tctctccttc ctttctttta	11760
ccttagaact atacatagaa tctacattca ataaaaatac agtaggtacg ccgagagatt	11820
taaaatgagt aagctaacat accaactaag gcctgtttg tttcggatta taatctctcc	11880
agattatata atccagcgta aataattcag cagataaaca aacacctaaa ttatatgttc	11940
agattatata atctatagcg gagattatga taatctcgta atctcctaag agtagcttat	12000
ttgagatTTT tttggcaaaa gaccactac ccaaggttat gtaaatagaa ttacaatata	12060
tgacatcctt ctttcttcac ctcaaataaa caaacaaggg tactgtgtgc tttatgaata	12120
atctacattt atataatcta gactaacaaa caactacata tagattataa tatgtctaga	12180
ttataatcta gattatataa tttaaattat agtccatatt atataatcta taagctaaaa	12240
caaataggcc ctaattatta gctattagtt gttagctatt taaatctaag ctaaaaccaa	12300

ctaatagctt attagttgaa ttacaattag ctcaacggaa ttctctgttt ttctaaaaaa	12360
aaactgcccc tctcttacag caaattgtcc gctgcccgtc gtccagatac aatgaacgta	12420
cctagtagga actctttttac acgctcggtc gctcgccgcg gatcggagtc cccggaacac	12480
gacaccaactg tggaacacga caaagtctgc tcagaggcgg ccacaccctg gcgtgcaccg	12540
agccggagcc cggataagca cggtaaggag agtacggcgg gacgtggcga cccgtgtgtc	12600
tgctgccacg cagccttcct ccacgtagcc gcgcggccgc gccacgtacc agggcccggc	12660
gctggtataa cgcgccacct ccgctttagt tctgcataca gccaacccaa cacacaccg	12720
agcatatcac agtgacagac actacacggt gagcgccaga atcgttgtcc tctcggccac	12780
cctcctatcc gctgcccgcg cagtcgcgtc gtctgggag gacgacaacc tccaccacca	12840
cggggggccac aagtccgggc gtagcgtgcg gcggtgcgag gaccggccct ggcaccagcg	12900
cccccggtgc ctggagcagt gcagggagga ggagcgggag aagcggcaag agcggagcag	12960
gctagaggcc gacgaccgca gcggcaaggg ttcgtcggag gtagagcgcg agcgagagca	13020
ggaggaggag aagcagaagg accggcggcc gtacgtgttc gaccggcgca gcttcogtcg	13080
cgtggtcccg agcgagcagg ggtctctgag ggtgctccgg ccgttcgacg aggtgtccag	13140
gctcctccgc ggcacccggg actaccgct ggcggtccta gaggcgaacc cgcgctcgtt	13200
cgtggtgccc agccataccg tagcgcaactg catctgtac gtggtggaag gtacgttcgt	13260
ccggtccctg agaccacata gatcgagggc gacgcggcc gtgcacgtta gctgatctgt	13320
tgttgccctg tgactaggcg agggataaac catggcctcc tccgagaacg tcatcaccga	13380
gttcatgcgc ttcaagggtgc gcatggaggg caccgtgaac ggccacgagt tcgagatcga	13440
gggcgagggc gagggccgc cctacgaggg ccacaacacc gtgaagctga aggtgaccaa	13500
ggggggcccc ctgcccttcg cctgggacat cctgtcccc cagttccagt acggctccaa	13560
ggtgtacgtg aagcaccocg ccgacatccc cgactacaag aagctgtcct tccccgagg	13620
cttcaagtgg gacgcggtga tgaacttcga ggacggcggc gtggcgaccg tgaccagga	13680
ctcctccctg caggacggct gcttcatcta caagggtgaag ttcacggcg tgaacttccc	13740
ctccgacggc cccgtgatgc agaagaagac aatgggctgg gaggcctcca ccgagcgcct	13800
gtacccccgc gacggcgtgc tgaagggcga gaccacaag gccctgaagc tgaaggacgg	13860
cggccactac ctggtggagt tcaagtccat ctacatggcc aagaagcccg tgcagctgcc	13920
cggctactac tacgtggacg ccaagctgga catcacctcc cacaacgagg actacaccat	13980
cgtggagcag tacgagcgca ccgagggccg ccaccacctg ttctgagat ctcgagctga	14040
tccaaaaaag aagagaaagg tagatccaaa aaagaagaga aaggtagatc caaaaaagaa	14100
gagaaaggta ggctccaccg gatctagata agagctcggc cgccgcgctc gccaaaacga	14160
gcaagaagca acaagagggt ggcgcgcgac cgacgtgcgt acgtagctga gcctgagtgg	14220

agacgttgga cgtgtatgta tatacctctc tgcgtgttaa ctatgtacgt aagcggcagg 14280
 cagtgcata agtgtggctc tgtagtatgt acgtgcggat acgatgctgt atgctactga 14340
 ggcaagtccc ataaataaat aatgacacgt gcgtgttcta taatctcttc gcttcttcag 14400
 tgtccccctg cggagtttgg catccattga tgccgttaca ctgagaacat aacggacaca 14460
 gtagacgaac cacttgagtt cttgtatgaa attctgaccc ttttttttga aggatcagga 14520
 gggggaaacc cctactggcc tggtcctttt acaatcaata aaaagaaaag aaaaagaaag 14580
 aaaaaggaag agctatttct atatttgcca caaathtagc cagttcggga ttgaggtgtc 14640
 gaattctccc cgtgctctgt gaatgattag tcgcaattct ctactaaatt cctgagtaca 14700
 gcggagcacg gagggatctc tatttgaaaa caaccactcg tttcggcttt tccagatgct 14760
 ccaagtcata agaatagtga tctccataga gaagggaacg ttgagctggg gtttgaggtt 14820
 cacagatgcc tcttcaaggt tggaaattct ggggtgggga atgccaattg agttccagca 14880
 ggctttcgcg aagttgcatc tgaggaaaag atgataaagg gtctcctctc tttgccaag 14940
 gcagttttcg caagagtagg aatctagctc catgtttctc cttctcaaca tatttctggt 15000
 gtttagcctg ttcttaagcc acagccaata gaagactcta tgcttaggct gacacttact 15060
 tttccaaact agtttgtaga taggatgagt ttgcatctgg cccatgaaag atctgtaagc 15120
 tttttgggac gagaatcgga ccgctgcag gcccgggggc gcgccctaat tagctaacgg 15180
 ccaggatcgc cgcgtgagcc tttagcaact agctagatta attaacgcaa tctgttatta 15240
 agttgtctaa gcgtcaattt gtttacacca caatatatcc tgccaccago cagccaacag 15300
 ctccccgacc ggcagctcgg cacaaaatca ccaactcgata caggcagccc atcag 15355

<210> 33
 <211> 15441
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Casset biểu hiện của cấu trúc 22513

<220>
 <221> trình tự tăng cường
 <222> (214)..(407)
 <223> eFMV

<220>
 <221> trình tự tăng cường
 <222> (414)..(706)
 <223> e35S

<220>
 <221> vùng khởi động
 <222> (714)..(4707)
 <223> Ubi

<220>

<221> gen
 <222> (4708)..(6165)
 <223> EPSPS

<220>
 <221> trình tự kết thúc
 <222> (6179)..(7178)
 <223> Ubi

<220>
 <221> vùng khởi động
 <222> (7198)..(9285)
 <223> pPLA2

<220>
 <221> vòng gấp
 <222> (9292)..(10386)
 <223> exon4 đích của PLA2

<220>
 <221> trình tự kết thúc
 <222> (10397)..(11392)
 <223> tPLA2

<220>
 <221> vùng khởi động
 <222> (11400)..(13526)

<220>
 <221> gen
 <222> (13527)..(14204)

<220>
 <221> trình tự kết thúc
 <222> (14210)..(15222)

<400> 33
 cctgtggttg gcatgcacat acaaattggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagtttaaac tggcactagc ctaacgggtg tgactaacta ggccggttcc 180
 ctaattagct aaccggggg cgcgccggga ccagctgct tgtggggacc agacaaaaa 240
 ggaatggtgc agaattgtta ggcgcaccta caaaagcaa ctttgctttt attgcaaaga 300
 taaagcagat tcctctagta caagtgggga aaaaataac gtggaaaaga gctgtcctga 360
 cagcccactc actattgcgt ttgacgaacg cagtgaacgac cacaaaactc gagacttttc 420
 aaaaagggtt attatccgga aacctcctcg gattccattg ccagctatc tgtcacttta 480
 ttgtgaagat agtggaaaag gaaggtggct cctacaaatg ccatcattgc gataaaggaa 540
 aggctatcgt tgaagatgcc tctgccgaca gtggtcccaa agatggaccc ccaccacga 600
 ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg 660
 atatctccac tgacgtaagg gttgacgaac aatcccacta tccttcaagc tttggaaaac 720
 caaaacaacg aataagcaaa ctgcaggaaa agtatgcagt ggaaaccaac ccagattcgg 780
 acgataggaa agtatcaagt gaatgatttg ccaggaaaag gagaggggta aaaaggggag 840

aagattttaga agatctaaag cacaagaacc agagatttaga ttgaacaata gggaacttgg 900
agcatccttt ttttcttcag ggaaaaactg aaaatccaaa ccatgttgag caaaaccgag 960
tgggattgga aaccaaaaaa cccgagataa agaaactcga gaaaaagcat gaaatcgaaa 1020
ccaacttcag taaaacaaaa ggaggacaga aaagaaagtc ggaagctata aagaatacat 1080
taacattcag tgaaacagca tgctgtcttc ttcttttttt atgcacaaca gagcatacat 1140
atataccttc ccaggctgag gacttggcgg aggagagccg cggatagggt ggcggtgcag 1200
acggtctgga cgggcccga gacggagacg aacagcgggc cttcctgcc caggcaccac 1260
gcttggaacg ccaagcacgc gccaacccgc gccccgccga ggacgacgat ccccgcgaca 1320
agcgtggcgt ccatcctggg cgaccccaag ccgaggaacc tcccttcag gacgagccgt 1380
aggacggcgg tgagagaggc acccgtcgcg gaggtggcgc agcacaaggt gagcagcgcg 1440
gggaaggcgg cgagcgtggc ggcctgcagg acggtgacga gcgcgaagac ggtgacgccg 1500
gcgacgaggc agcagcagcc gaggatccag tcgtaggagg aagccggacc aaaccgggca 1560
atgcaacctg cagatgcact agacggaggt aacgaggagg aggagaaaac agagcaagag 1620
caggcggaga gaagatagag caaaacacga gtgaggcaca gcgtaagcac tcggtagaag 1680
tctccagagg cgaggtgcgc acaggagaac agatgagtaa agtcagccaa ggatccacga 1740
tccaacggct acgaattttt ggagtgcgt ggataggctc aaaggcgcca tttccatccg 1800
gctttatagt attttaaaaa aattcatttt cctccctcta gtgtgtgcgg aggcgtgagc 1860
ccgtttaacg gcgttgagaa gtctaacgga caccaaccac aaccaggaac cagcgccggc 1920
cgcgccgccg agtgaagcag actgcatacg gcacggcgcg gcctctctct ggctgcctct 1980
cgagagttcc gccccacct tcccgcggtg gcgtgggtgt ttcgctttcc gctgtcggca 2040
tccggaagtt gcgtggcaga gtggacggag acgaggccgg gtcctccagc tctctcaaa 2100
cgtcacggca ccggcatccg gcagccagcg cggtccttcc caaccactcg tttccaacct 2160
atcccccttc ctgcgccgc gtcataaata gccagcccca tcccagctt ctttcccaa 2220
cctcatcttc tctccttttg ctctgaacgc acacaccgcc cggctctccga tctccgatcc 2280
ccgatccct cgtcgatcct aggtacggcg accatcctac ccccccccc cccctctc 2340
tctctgcctt ctctagatcg gcgatccgat ccatgcttac ttggttaggg cctgctaact 2400
atgttcatgt ttgcgttaga tccgtgcatg gacgcgatct gtacacacca gacgcgttct 2460
gattgctagc taactcgcca gtacctgga atcctgggat ggctgtagcc ggccccgcac 2520
gcagacggga ccgatttcat gattctctat tttttcttt gtttcgttgc ctagggtttc 2580
gttcgatcga tccgcgttat tctttatttc catatattct ggtacgatgt tgatacggtt 2640
cgaccgtgct gcttacgttc tgtgcgcttg tttgccgggt catttttacc ttgccttttt 2700
tgtatggttt gggtgtggcg atgtggtctg gtcgggctgt cgttctagat cggagtagag 2760

tgctgtttca aactgtctag cggatctatt agatttggat ctgcatgtgt gacatatatc	2820
ttcgtagtta agatgatgca tctgtatgtg tgacatgcgg atctattaga tttggatctg	2880
tatgtgtgac atatatcttc gtagttgaga tgatgcatct gtatgtgtga catatatctt	2940
cgtagttaag attatgcatg gaaatatcaa tccttttagat aaggacgggt atacttgttg	3000
ctgtggggtt tactgggtact tcgatagatg catatacatg atctaacatg cttagataca	3060
tgaagtaaca tgctgctacg gtttaataat tcttgagttg atttttactg gtacttagat	3120
agatgtatat acatgcttag atacatgaag taacatgctc ctacagttcc tttaatcatt	3180
attgagtacc tatatatctt aataaatcag tatgttttaa attattttga ttttactggg	3240
acttagatag atgtatatat acatgctcaa acatgcttag atacatgaag taacatgctg	3300
ctacgggtta gtcattattg agtgcctata tattctaata aatcagtatg ttttaaatta	3360
ttttgatttt actgggtact agatagatgt atatatacat gctcaaacat gcttagatac	3420
atgaagtaat atgctactac ggtttaattg ttcttgagta cctatatatt ctaataaata	3480
agtatgtttt aaattatttc gatttttactg gtacttagat agatgtatat atacatgctt	3540
agatacatga agtaacatgc tactacgggt taattgttct tgaataccta tatattctaa	3600
taaatcagta tgttttaaat tatttcgatt ttactgggtac ttagatagat gtatatatac	3660
atgctcgaac atgcttagat acatgaagta acatgctaca tatatattat aataaatcag	3720
tatgtcttaa attattttga ttttactggg acttagatag atgtatatatac atgctcaaac	3780
atgcttagat acatgaagta acatgctact acggtttaat cattattgag tacctatata	3840
ttctaataaa tcagtatgtt ttcaattgtt ttgattttac tgggtacttag atatatgtat	3900
atatacatgc tcgaacatgc ttagatacgt gaagtaacat gctactatgg ttaattgttc	3960
ttgagtacct atatatctta ataaatcagt atgtttttaa ttatttcgat tttactggta	4020
cttagataga tgtatatata catgctcgaa catgcttaga tacatgaagt aacatgctac	4080
tacgggttaa tcgttcttga gtacctatat attctaataa atcagtatgt cttaaattat	4140
cttgatttta ctgggtactta gatagatgta tatacatgct tagatacatg aagtaacatg	4200
ctactatgat ttaatcggtt ttgagtaoct atatatctta ataaatcagt atgtttttta	4260
ttattttgat tttactggta cttagataga tgtatatata catgctcgaa catgcttaga	4320
tacatgaagt aacatgctac tacgggttaa tcattcttga gtacctatat attctaataa	4380
atcagtatgt ttttaattat tttgatatta ctgggtactta acatgttttag atacatcata	4440
tagcatgcac atgctgctac tgtttaatca ttcgtgaata cctatatatt ctaatatatc	4500
agtatgtctt ctaattatta tgattttgat gtacttgat ggtggcatat gctgcagcta	4560
tgtgtagatt ttgaataccc agtgtgatga gcatgcatgg cgccttcata gttcatatgc	4620
tgttttattc ctttgagact gttctttttt gttgatagtc accctgttgt ttggtgattc	4680

ttatgcagat ccagatcttc gtaaaccatg gctctggtgg ctaggccagt gctgtctgct	4740
agagtggctg cttctaggcc aaggggtggcc gctagaaagg ctgtgagggg gagcgctgcc	4800
agtcacggcg ctagctctag gccagccacc gctaggaagt ccagcggcct gagcggcacc	4860
gtgcgcaccc caggcgacaa gagcatcagt caccgcagct tcatgttcgg cggcctggcc	4920
agcggcgaga ctaggatcac cggcctgctc gagggcgagg acgtgatcaa caccggcaag	4980
gccatgcagg ctatgggctc tcgcatccgc aaggagggcg acacctggat catcgacggc	5040
gtgggtaatg gcggcctgct ggccccagag gcccactgg acttcggcaa cgccgccacc	5100
ggctgcaggc tgactatggg cctgggtggg gtgtacgact tcgacagcac ctttatcggc	5160
gacgccagcc tgaccaagag gccgatgggc aggggtgctga acccgctgcg cgagatgggc	5220
gtgcaggtca agagcgagga cggcgacagg ctgccggtga ccctgagggg cccaaagacc	5280
ccgacccccg tcacctaccg cgtgccaatg gccagcgccc aggtgaagtc agccgtgctg	5340
ctggccggcc tgaacacccc aggcacacc accgtgatcg agccgatcat gaccgcgat	5400
cacaccgaga agatgctgca gggcttcggc gccaacctga ccgtcgagac tgacgctgac	5460
ggcgtgagga ccatcaggct cgagggcagg ggcaagctga ccggccagggt gatcgacgtg	5520
ccgggcgacc caagcagcac cgccttccca ctggtggccg ccctgctggt gccaggctct	5580
gacgtgacca tcctgaacgt gctgatgaac ccgagccgca ccggcctgat cctgacctg	5640
caggagatgg gcgccgatat cgagggtgatc aacccgaggc tggctggcgg cgaggacgtc	5700
gccgacctga gggtgaggtc cagcaccctg aagggcgctga ccgtgccaga ggacagggcc	5760
ccgagtatga ttgacgagta cccgatcctg gccgtggccg ctgccttcgc cgagggcgcc	5820
accgtgatga acggcctcga ggagctgcgc gtgaaggaga gcgacaggct gagcgctgtg	5880
gccaacggcc tgaagctgaa cggcgtggac tgcgacgagg gcgagactag cctggtggtg	5940
aggggcaggc cagacggcaa gggcctgggc aacgcttcgc gcgtgccgt ggccaccac	6000
ctggatcaca ggatcgccat gtcgttcctc gtgatgggcc tcgtgagcga gaaccgggtg	6060
accgtggacg acgccaccat gatcgccacc agcttccccg agttcatgga cctgatggcc	6120
ggcctgggcg ctaagatcga gctgagcgac accaaggccg cctgatcatc tagagctcgc	6180
caaggttcaa ttaagctgct gctgtacctg ggtatctgcg tcgtctggtg ccctctggtg	6240
tacctctata tggatgtcgt cgtctaataa acatctgtgg tttgtgtgtc atgaatcgtg	6300
gttgtggctt cgttggttta atggacctgt tgtgtcctct gtgttgtagc caaaactctt	6360
ctgcagcagt atggcctgaa tccttatgaa gtttgatatt tgaacttaaa agtctgtcga	6420
ttatgttttt ttctggttat atctcctaataa taactgcctg ggatcaaatt tgattcgctg	6480
gtgtttattg gacccctccc aggttcttgc tttctaccgt ttcttgctga atgttaactt	6540
gattctgtca ggctcagttt cccactatgg cttacagctt aacgtgtttg gtttgttgaa	6600

tggttaacttg gttttgtcaa gctcagtttt ttactctggc ttacagcata acatgtttga	6660
cttttggttt tgctgctttg ttattgggtt ctgggtagtt cttgatgaat ccaaaagatc	6720
atgtgcacag ccatattatc tatttaagcg atccaggtta ttactatgaa aggatgcctt	6780
ctagctaagg agtagttagg ttttttcttc aagggttaaat tttctcgatg ctctagtgtt	6840
cctgtgacca taatcataat aattcctttg aaagctctat ggtccctgga agcagggcat	6900
acaatgcaag acagcaactt gatcacatca actgaagtat acagggttct cttaactctt	6960
ggtgacttcg gtttaatgga ccggttgtagc tcgtgttcta tccgtaaccg ttgtgatgac	7020
ttgtgtgttt ggttgcgga tagctgggac cagcagcttt ccgtctaatt ctgatggata	7080
gctatagacg gcaactgagat gggttatatta taacctctga tcctgaactc tacgagatcg	7140
tctcatccgt cattgccacc aaatacacca ttaaattact aattagctaa cggaccctat	7200
ttgtactcat tccatgtctc ataaactttg ggcaccatcc atccaacaca tccaatctaa	7260
acacacaaaa cgatggggaa tggaagagc agtattcgat tcaacaatgg caaacaata	7320
tcactgaatt agaccaagaa taaacctaat tagacaacga cctcccaacc atcattcgtc	7380
aggctgtaaa gaagataaag ctgccttggg gcatggatca agcagaacac cagagatgaa	7440
tccaaacaca cagaaaatca cgcgcgctgt ctacaatgac aacaagcccc acatttcatt	7500
gcagtacact gggctacaaa ggcacgtaca acaagagct agggaaacat tgcggagggc	7560
acgagagagc agctaacttg acaatatagc agactgagct tgcaactgta gcaggcgagg	7620
aagggaatca tggggacgga gaatggggtc catgcccgcg aaggagaagg cggacgccgc	7680
cacggtggca ccggcgacg cgacacagc gaaccgcac aggcagccaa ggatgctgcc	7740
tcgccattgc gccggtcgtc tctgccacgc tcctctctct ctcgcgctgc atcgccgtgg	7800
atggggcaag cagagagcag ggactgcgac gatctgggcg gaggactcgc cttggagagc	7860
gcggacgcag acgggattct agggagagag cgaagacggg gcgcgcgcgc cgtcgcgcgc	7920
gcgtggtggc ggcgagatta gcgggggtgg ggggagggcg gagccgtggt gaggggtgtg	7980
acgccctcct taccctctta agtagtagta gagatataat ccgttccaaa atatccatcc	8040
gttcaattta tatttcgttt gatcttttta cctaaattt gattgactca tcttattaaa	8100
aaagttcata actattatta atctttattg agatatcatt tagcatataa tatactttta	8160
gtgtgggttt agatTTTTTT taaaaaaaaa aattcgcaaa aattaaatga aacgacccaa	8220
tcaaacttga aaagtaaaac taattataaa tttgaacgga aggagtaaga ggatgtttga	8280
atgtactaga gctaatagtt ggttgcttta aaatttgcta gtagaattag ctagctaata	8340
aatatctaga taactattag ctaatttgct aaaacagcta atagttgaac tattagctag	8400
attgtttgga tgtattcggc taattttaat ggctaactat tagctatagt acaatattca	8460
aacacctcct aattaaaatg gacaaaatc tcttcttttg gtcccttgcg ttagattttt	8520

catatctcct tatttagtat aaaagaatca tcaaaaagtg gacaaccct agtggaacac	8580
cattttagta gtggttgcat gaaacctttc ggcaccagt ttctatgtgt cactctaaaa	8640
atgggacagc atgtacgtag tgcttatata tatacaagtc atctatcgtt gcctcctcag	8700
ttcatcacta atcacactta ttgtgccctc gacgagtatc tatagctagc tcattaatcg	8760
attcggggggt gtgttgctga aggcggcatt ggcgagctac tcgtcgcggc gtccaagcaa	8820
tacctgtagc acgaaggcga tcgccgggag cgtggtcggc gagcccgctg tgctggggca	8880
gagggtgacg gtgctgacgg tggacggcgg cggcgtccgg ggtctcatcc cgggaacat	8940
cctcgccttc ctggaggcca ggctgcagga gctggacgca ccggaggcga ggctggcgga	9000
ctacttcgac tacatcgccg gaaccagcac cggcggctctc atcaccgcct tgctgaccgc	9060
gccgggcaag gacaagcggc ctctctaggc tgccaaggac atcaaccact ttacatcca	9120
taactgcccg cgcacttttc ctcagaagtg agtccgatgc tgccgccatt gttcttgcat	9180
ccatccagca tcgtacgtac gtctctata catctgcgga tcatcatgtg cgcagtgttg	9240
tggcatgcat gcatgcatgt gagcaggagc aggcttgca aaaccccatg gagatctttg	9300
cgatggcgca gatcaccaa aagatgcttg ccaacaagga caaggccgag gagctgtacc	9360
cagtgaagcc gtcgaactgc cgcaggttcc tgggtgctgtc catcgggacg gggtcgacgt	9420
ccgagcaggg cctctacacg gcgcggcagt gctcccggtg gggatatctgc cgggtggctcc	9480
gcgtcgactc gggagcatgc cctttgtgct tatgcctccg ttctgccttc tgacgaattt	9540
ggtactggaa gcagatgagt tttggttcac tatcattctg aatttacacc tgcgcttgct	9600
gtcagactag gcaaccaagt gacttttgctg actttgatca tgttcagtgt gtttccaagt	9660
cctaataat caaaaagaaa aacagtttgt taacgattgt ttgccatgtc tatataataa	9720
agttgctttt atagtagctt agaattcaat cggccaactt tatctcgtac gctgacagta	9780
aaggtaacatt taaaagggtga caatggatag tctaataact gaactgacaa tagagacaca	9840
ttacatgtca gttgattaag tttgtaacag aaaaataaac aatactacat aattgcaaag	9900
tttctttgat gtctttcttt caagaaacac aaatatatca atgctacagt attgctgatg	9960
aatttatcca tgttgagatg tttttctggt ttctgatctg atcagttcca attggtgtgc	10020
tgtttcattt tcatttgctg atgacgtcc gagtagttaa ttcttactaa tatttagata	10080
atttggcata caagcgaatc acgtagaaca tgatactttt gaatgaattt atcaaagttt	10140
tatcacttgg tgagttgttt catggttttc ctactgatgt ctcttcttca gattttctga	10200
ggcggagcca ccggcagata cccacccggg agcactgccg cgcggtgtag aggcctgct	10260
cggacgtcga ccccgctccg atggacagca ccaggaacct gcggcagttc gacggcttca	10320
ctgggtacag ctctcgggc ttgtccttgc tggcaagcat ctttttggtg atctgcgtca	10380
tcgcaaggat ccatgggaca agtggcttta ctgtcagtca catgcttgta aataagtaga	10440

ctttattttta ataaaacata aaaatatata tatgttcttg aatataaaat tgataaccaa	10500
attaaaaattc gaaccatcac ttatacataa ttttacttta ttttttataa aacgtgaacg	10560
ggaaggacta ccgatgaatga ctatagaacc aatcatacta gtataaaata tatgatgaca	10620
ctacgggaga gacaaaacttt gtctggcgct aaatatcttg ccgagtgtga attcacgggc	10680
actaggcaaa gatcttcttt gccgagtgtt acgctgggca aagtaagaca ctaggtaaat	10740
cagtcatttg ccgagtgtcc gccactaggc aaagcaaaac actggcaaact caaaagttaa	10800
cctagtgcc aacactaggc aaaaaaaaaa cgctcggcaa atcggaagtt tccctagtgc	10860
cagacactag acaaagaaaa acacttgata aactagcgtc gtcagctaac accatccacc	10920
aaccgttaac gttgccgagt atctgacttc gacactcggc aaagaaggtc tctttgccta	10980
gtgtcgggtct ggaacactag gcaaagaggc actttaccta gtgtcgtatt ttgacactca	11040
gtaaaataat tttttttctt tctgcttcca aactttttat gatgtgttcc tatagcacct	11100
agaactacat gtcaagtttt ggtaaaatct ttgaagtttt tgctatatctt acttaattta	11160
ttttatttaa ttgaatttct tttgataatt caaatttgaa ctcggaagg taagaagcga	11220
gggtagcctg gaaacacact ttgcctagtg ttacactcgg tacaggagcc tccctgcct	11280
agtgtgcac tcgacaaaag attcgccttt gcctagcgtc gcactcggca caggagtcgc	11340
ctttgcctag tgctgcacta ggcaaagcct ccgttacgtt gccttccatc gtcggaccgc	11400
aaagtagcaa acaacagggt catgtgcact ataaaaagac aaaattctcg agtttcatct	11460
tttattccac ataagcctta tattttccat tttcatatga ttttttagttt aagtttgtgt	11520
cttaactttt tcgttaatac gtaattctat gcattatgga tgcgtgaagt atttttgttt	11580
aaaaaaatga aatgtcaaaa tacgttttgt gatctatttc catgttttca cctaacagggt	11640
ggtttttact atatatcttg ccataactct agccttagat gtaaactgaa aaaaaatgag	11700
agatgagctg gagatagcct tagatgaagc gtctgaaata taaaagaaag agtaatgttg	11760
aacgcagtag gtgtagcagc tgtagtcca tctctaggaa agggaaactgc aatccgggct	11820
ccgggcctcg cgcaatctgg cctgtcgtgt agatgcagcc ctgtccatga cggcccaagc	11880
aacgcccgcg gctctcgatc caccacggaa cccactccga cacacactga cacacacatg	11940
ctggatgttg atgtgctgtc caattattag tagcaattcg gtaggcacag gcacgtactg	12000
gccggtgttt tagctgtaag taccgaacca atcacggtta agaaccgatt aatccgtgcc	12060
cagccgccga gtgcgttcgt acgtgcatcg gatgcactgc atgaattgag agcatcatca	12120
tatcatacgc aggagttagta cgacgcgcgt gctgtcttgt ccggctaatag ctttgctcac	12180
agattagtcc atcgcacacg gtccgtgttg tgtggatcgc tgatgccact gctttttgtt	12240
tggtttttat tccctgata atcctccgcg tccctgaatg tatctattta ttttcatcc	12300
gaaatccctt tcacgaaaaa gaaaacgaat aaaaagagag ttacgaatac gcttccggcg	12360

gcccacatca ccttccagcg aacatcgcg cgcgctgacg tgtcgcccat cgcggccgtc 12420
catatcgcca tccgacgacc gtggaagctg gcagcggccg ctccgttccg tcgaaggggc 12480
aggtcagtcg ggtcaccac acggccacac ccgcgcgggg gatacgcggt ggaaaacccg 12540
gcgaccacat caaaacacga ggcgtctccc gcaggactgg tcaactcgga cgcaggcaga 12600
ggcagcacag cagcagccag ctccatccat cctctttccc ctctcgtctt cgcttcctcg 12660
gcggattcct cctccctcgg ccgtccccgt ccccttcttc gccgcgccag ctcgcccag 12720
ttggtacgcc catcctctgc tgtactcccc ccgttctctg tgtgcgaaac cgagcaccca 12780
aaccctagcc taagctattc gcatcgcaaa ctctattgag cgacggatcg cgaaacgcgc 12840
gccgccgcat tcggacgaga cctcgtggcc cgtaccttcc tactctaccg tcgtcgccgg 12900
agaccagctc accggttagg gtttcgggat taggggcttg gggcttgat tcccggggga 12960
ctccataccc ggtggcctta gaaggggaag ggggccagc ggtggtggtg gttcgactgt 13020
cagcgagcga gacggcgggg caccgccgat cgggcgtcgc tggacatttg attgagctgg 13080
gcaggcgaag gcgtggagct tgctcttcga ttgggattag aggagggcg aggtggtatg 13140
gtgggcggtc cggctggctc aatggagccg ctggccggtg gtggggcagg cgagatcacg 13200
ctctctcata gagcgtaatg ggttgcgtaa cacactcctg tttgtatttc gatccgatct 13260
aataattaa gtttatttac gtggtgcaca tgtagaattt tcccatgcc actgctaact 13320
tggtattgat ttttgccaaa aaaacgggtg tgcaccaat cacttatagt gaaaccgttg 13380
tgtgtttgat gtgcttctaa ccaagtacaa atccagtga aacacaccta ccattcaaca 13440
atccacattt tggttgcaat tatcagcatt ctagaagggt gcatgtgccc attgatacac 13500
ttgctattgg tgcaggcaaa cccaccatgg cctcctccga gaacgtcatc accgagttca 13560
tgcgcttcaa ggtgcgcatg gagggcacgc tgaacggcca cgagtctgag atcgagggcg 13620
agggcgaggg ccgcccctac gagggccaca acaccgtgaa gctgaagggtg accaagggcg 13680
gccccctgcc cttcgcttg gacatcctgt cccccagtt ccagtaaggc tccaagggtg 13740
acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc gagggcttca 13800
agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggc gaccgtgacc caggactcct 13860
ccctgcagga cggtgcttcc atctacaagg tgaagtctat cggcgtgaac tccccctccg 13920
acggccccgt gatgcagaag aagaccatgg gctgggaggg ctccaccgag cgctgtacc 13980
cccgcgacgg cgtgctgaag ggcgagaccc acaaggccct gaagctgaag gacggcggcc 14040
actacctggt ggagttcaag tccatctaca tggccaagaa gcccgtcag ctgcccggct 14100
actactacgt ggacgccaag ctggacatca cctcccacaa cgaggactac accatcgtgg 14160
agcagtaaga gcgcaccgag ggcgcgccacc acctgttctt gtaggagctc gcatcatgat 14220
catgcatcat ggactcggcc tactactgtg gatttgtatg ccattataga cttggtgctg 14280

tgaaagactg cttgatgatt tgcggggttg ttgctgtgta aaaaaagggtc ccttggtccc 14340
 cagaagacca tgaagggtcg gatctatcat gtaattcctt gttatctgcc aattatgtat 14400
 ggactatgga catgtgttgc gctgttcaac ttactactac aaataagtaa tcgatatgtt 14460
 ccottcccat gtctcgggtga caattgtctg gagaagctta ggggtcgttt gtttgggatt 14520
 atgtctggag aaacttattt taaactaagt gtgagttcaa gttaagttag attatataat 14580
 ctaggcagat tataattcca agcgaacagg tccttagtgt ttttggaata tcctaggtgt 14640
 tcttttggct acattgttgt gtgtgcagat cccttggttg tctgtaagcg tggggaagta 14700
 agaatcgtcc gtttctactg aagacctgct cgagttaggc accgaggatg ccggaacca 14760
 aacagagcaa tagtgtctct gtgggcacag tggagtgtga atctgtgtga tgcaaaccg 14820
 tcatttgttt agcaaaattt ccagcgttgc atgatgcagt ttctttaaca cggacttaag 14880
 ggaagggaaa aaaatgttga gccaggagat cttcaatgt gttagactga cgtgatagcc 14940
 aactaaacca cgacgcaatg ttgtcgtaa tgacaaaaaa actatttgtt cctaaatcct 15000
 tggcgacatt gcatggctgt ctcatgagat aatggtctca tctcttattt atctcttatt 15060
 tatagccgga agtggtagtg acccctgctt gattgctcgt atgcatctc aagttctcaa 15120
 ccgtgtcgag cagccatttt cccatctcaa gcgcacatc gtttcgtttg acctcatctg 15180
 ctatcctgct cctagtgtca atcacatgcg acagaaagtg tgcggaccgc ctgcaggccc 15240
 gggggcgcg cctaattagc taacggccag gatcgccgcg tgagccttta gcaactagct 15300
 agattaatta acgcaatctg ttattaagtt gtctaagcgt caatttgttt acaccacaat 15360
 atatcctgcc accagccagc caacagctcc ccgaccggca gctcggcaca aaatcaccac 15420
 tcgatacagg cagcccatca g 15441

<210> 34

<211> 10671

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 22807

<220>

<221> trình tự tăng cường

<222> (165)..(256)

<223> eNOS

<220>

<221> vùng khởi động

<222> (301)..(2097)

<223> Ubi

<220>

<221> gen

<222> (2114)..(6283)

<223> Cas9

<220>

<221> đặc điểm_hỗn tạp

<222> (6218)..(6274)

<223> NLS-TAG

<220>

<221> trình tự kết thúc

<222> (6289)..(6541)

<223> tNOS

<220>

<221> vùng khởi động

<222> (6548)..(6922)

<223> pU3

<220>

<221> đặc điểm_hỗn tạp

<222> (6924)..(6941)

<220>

<221> đặc điểm_hỗn tạp

<222> (6924)..(7026)

<223> sgRNA

<220>

<221> vùng khởi động

<222> (7037)..(9028)

<223> Ubi

<220>

<221> gen

<222> (9045)..(10223)

<223> PMI

<220>

<221> trình tự kết thúc

<222> (10246)..(10498)

<223> tNOS

<400> 34

cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa	60
atatccgatt attotaataa acgctotttt ctotttaggtt taccgcgcaa tatatcctgt	120
caaacactga tagtttaaac tgaaggcggg aaacgacaat ctgatcatga gcggagaatt	180
aaggaggtca cgttatgacc cccgccgatg acgcgggaca agccgtttta cgtttggaac	240
tgacagaacc gcaacgctgc aggaattggc cgcagcggcc atttaaaca agcttggtac	300
cattatgttg tctaggtagg ttctatatat aagaaaactt gaaatgttct aaaaaaaaaat	360
tcaagcccat gcatgattga agcaaacggt atagcaacgg tgtaaacctg atctagtgat	420
ctcttgcaat ccttaacggc cacctaccgc aggtagcaaa cggcgtcccc ctccctcgata	480
tctccgcggc gacctctggc tttttccgcg gaattgcgcg gtggggacgg attccacgag	540
accgcgacgc aaccgcctct cgccgctggg cccacaccg ctcggtgccg tagcctcacg	600
ggactctttc tccctcctcc cccgttataa attggcttca tcccctcctt gcctcatcca	660

tccaaatccc agtccccaat cccatccctt cgtaggagaa attcatcgaa gctaagcgaa	720
tcctcgcgat cctctcaagg tactgcgagt ttctgatccc cctctcgacc cctcgtatgt	780
ttgtgtttgt cgtagcgttt gattaggtat gctttccctg tttgtgttcg tcgtagcggt	840
tgattaggtg tgctttccct gttcgtgttc atcgtagtgt ttgattaggt cgtgtgaggc	900
gatggcctgc tcgcgtcctt cgatctgtag tcgatttgcg ggctcgtggtg tagatctgcg	960
ggctgtgatg aagttatttg gtgtgatctg ctgcctgat tctgcgggtt ggctcgagta	1020
gatatgatgg ttggaccggt tggttcgttt accgcgctag ggttgggctg ggatgatgtt	1080
gcatgcgccg ttgcgcgtga tcccgcagca ggacttgctg ttgattgcca gatctcgta	1140
cgattatgtg atttggtttg gacttttttag atctgtagct tctgcttatg tgccagatgc	1200
gcctactgct catatgcctg atgataatca taaatggctg tggaactaac tagttgattg	1260
cggagtcatg tatcagctac aggtgtaggg actagctaca ggtgtagggg cttgcgtcta	1320
attgtttggt cctttactca tgttgcaatt atgcaattta gtttagattg tttgttccac	1380
tcacttaggc tgtaaaaggg aactgctta gattgctgtt taatcttttt agtagattat	1440
attatattgg taacttatta cccctattac atgccatacg tgacttctgc tcatgcctga	1500
tgataatcat agatcactgt ggaattaatt agttgattgt tgaatcatgt ttcatgtaca	1560
taccacggca caattgctta gttccttaac aaatgcaaatt tttactgac catgtatgat	1620
ttgcgtgggt ctctaattgtg aaatactata gctacttggt agtaagaatc aggttcgtat	1680
gcttaatgct gtatgtgcct tctgctcatg cctgatgata atcatatata actggaatta	1740
attagttgat cgtttaataca tatatcaagt acataccatg ccacaatttt tagtcaacta	1800
acccatgcag attgaactgg tccctgcatg ttttgctaaa ttgttctatt ctgattagac	1860
catatatcat gtatTTTTTT ttggtaatgg ttctcttatt ttaaagcta tatagttctg	1920
gtacttggtta gaaagatctg ctccatagtt tagttgccta tccctcgaat taggatgctg	1980
agcagctgat cctatagctt tgtttcatgt atcaattctt ttgtgttcaa cagtcagttt	2040
ttgttagatt cattgtaact tatggctgct tactcttctg gtctcaatg cttgcaggat	2100
cgcggccgct catatggaca agaagtacag catcggcctg gacatcggca ccaacagcgt	2160
gggctgggac gtgatcaccg acgagtacaa ggtgcggagc aagaagttca aggtgctggg	2220
caacaccgac aggacagca tcaagaagaa cctgatcggc gccctgctgt tcgacagcgg	2280
cgagaccgcc gaggccacca ggctgaagag gaccgccagg aggaggtaca ccaggaggaa	2340
gaacaggatc tgctacctgc aggagatctt cagcaacgag atggccaagg tggacgacag	2400
cttcttccac aggtcggagg agagcttctt ggtggaggag gacaagaagc acgagaggca	2460
cccgatcttc ggcaacatcg tggacgaggt ggcctaccac gagaagtacc cgaccatcta	2520
ccacctgagg aagaagctgg tggacagcac cgacaaggcc gacctgaggc tgatctacct	2580

ggccctggcc cacatgatca agttcagggg ccacttcctg atcgagggcg acctgaaccc 2640
 ggacaacagc gacgtggaca agctgttcat ccagctgggt cagacctaca accagctgtt 2700
 cgaggagaac ccgatcaacg ccagcggcgt ggacgccaag gccatcctga gcgccaggct 2760
 gagcaagagc aggaggctgg agaacctgat cggccagctg ccgggcgaga agaagaacgg 2820
 cctgttcggc aacctgatcg cctgagcct gggcctgacc ccgaacttca agagcaactt 2880
 cgacctggcc gaggacgcca agctgcagct gagcaaggac acctacgacg acgacctgga 2940
 caacctgctg gccagatcg gcgaccagta cggcgacctg ttcttgccg ccaagaacct 3000
 gagcgacgcc atcctgctga gcgacatcct gaggtgaac accgagatca ccaaggcccc 3060
 gctgagcgcc agcatgatca agaggtacga cgagcaccac caggacctga cctgctgaa 3120
 ggccctgggtg aggagcagc tgccggagaa gtacaaggag atcttcttcg accagagcaa 3180
 gaacggctac gccggctaca tcgacggcgg cgccagccag gaggagtctt acaagttcat 3240
 caagccgatc ctggagaaga tggacggcac cgaggagctg ctggtgaagc tgaacagga 3300
 ggacctgctg aggaagcaga ggaccttca caacggcagc atcccgacc agatccacct 3360
 gggcgagctg cagccatcc tgaggaggca ggaggacttc taccgttcc tgaaggacaa 3420
 cagggagaag atcgagaaga tcctgacctt ccgcatcccg tactacgtgg gcccgctggc 3480
 caggggcaac agcaggttcg cctggatgac caggaagagc gaggagacca tcaccccgctg 3540
 gaacttcgag gaggtggtgg acaagggcgc cagcgcccag agcttcatcg agaggatgac 3600
 caacttcgac aagaacctgc cgaacgagaa ggtgctgccg aagcacagcc tgctgtacga 3660
 gtacttcacc gtgtacaacg agctgaccaa ggtgaagtac gtgaccgagg gcatgaggaa 3720
 gccggccttc ctgagcggcg agcagaagaa ggccatcgtg gacctgctgt tcaagaccaa 3780
 caggaaggtg accgtgaagc agctgaagga ggactacttc aagaagatcg agtgcttcga 3840
 cagcgtggag atcagcggcg tggaggacag gttcaacgcc agcctgggca cctaccacga 3900
 cctgctgaag atcatcaagg acaaggactt cctggacaac gaggagaacg aggacatcct 3960
 ggaggacatc gtgctgacct tgacctgtt cgaggacagg gagatgatcg aggagaggct 4020
 gaagacctac gccacctgt tcgacgacaa ggtgatgaag cagctgaaga ggaggaggta 4080
 caccggctgg ggcaggctga gcaggaagct gatcaacggc atcagggaca agcagagcgg 4140
 caagaccatc ctggacttcc tgaagagcga cggcttcgcc aacaggaact tcatgcagct 4200
 gatccacgac gacagcctga ccttcaagga ggacatccag aaggcccagg tgagcggcca 4260
 gggcgacagc ctgcacgagc acatcgccaa cctggccggc agcccggcca tcaagaaggg 4320
 catcctgcag accgtgaagg tggaggacga gctggtgaag gtgatgggca ggcacaagcc 4380
 ggagaacatc gtgatcgaga tggccaggga gaaccagacc acccagaagg gccagaagaa 4440
 cagcagggag aggatgaaga ggatcgagga gggcatcaag gagctgggca gccagatcct 4500

gaaggagcac ccggtggaga acaccagct gcagaacgag aagctgtacc tgtactacct 4560
 gcagaacggc agggacatgt acgtggacca ggagctggac atcaacaggc tgagcgacta 4620
 cgacgtggac cacatcgtgc cgcagagctt cctgaaggac gacagcatcg acaacaaggt 4680
 gctgaccagg agcgacaaga acaggggcaa gagcgacaac gtgccgagcg aggaggtggt 4740
 gaagaagatg aaaaactact ggaggcagct gctgaacgcc aagctgatca cccagaggaa 4800
 gttcgacaac ctgaccaagg ccgagagggg cggcctgagc gagctggaca aggccggctt 4860
 cattaanaagg cagctggtgg agaccaggca gatcaccaag cacgtggccc agatcctgga 4920
 cagcaggatg aacaccaagt acgacgagaa cgacaagctg atcagggagg tgaaggtgat 4980
 caccctgaag agcaagctgg tgagcgactt caggaaggac ttccagttct acaaggtgag 5040
 ggagatcaat aattaccacc acgcccacga cgcctacctg aacgccgtgg tgggcaccgc 5100
 cctgattaaa aagtaccoga agctggagag cgagttcgtg tacggcgact acaaggtgta 5160
 cgacgtgagg aagatgatcg ccaagagcga gcaggagatc ggcaaggcca ccgccaagta 5220
 cttcttctac agcaacatca tgaacttctt caagaccgag atcaccttg ccaacggcga 5280
 gatcaggaag aggccgctga tcgagaccaa cggcgagacc ggcgagatcg tgtgggacaa 5340
 gggcagggac ttcgccaccg tgaggaaggt gctgtccatg ccgcaggtga acatcgtgaa 5400
 gaagaccgag gtgcagaccg gcggcttcag caaggagagc atcctgccga agaggaacag 5460
 cgacaagctg atcgccagga agaaggactg ggaccogaag aagtacggcg gcttcgacag 5520
 cccgaccgtg gcctacagcg tgctggtggt ggccaaggtg gagaagggca agagcaagaa 5580
 gctgaagagc gtgaaggagc tgggtgggcat caccatcatg gagaggagca gcttcgagaa 5640
 gaaccagtg gacttcttg aggccaaggg ctacaaggag gtgaagaagg acctgatcat 5700
 taaactgccg aagtacagcc tgttcgagct ggagaacggc aggaagagga tgctggccag 5760
 cgccggcgag ctgcagaagg gcaacgagct ggccctgccg agcaagtacg tgaacttct 5820
 gtacctggcc agccactacg agaagctgaa gggcagcccg gaggacaacg agcagaagca 5880
 gctgttcgtg gagcagcaca agcactacct ggacgagatc atcgagcaga tcagcgagtt 5940
 cagcaagagg gtgatcctgg ccgacgcaa cctggacaag gtgctgagcg cctacaacaa 6000
 gcacagggac aagccgatca gggagcaggc cgagaacatc atccacctgt tcacctgac 6060
 caacctgggc gccccggcgg ccttcaagta cttcgacacc accatcgaca ggaagaggta 6120
 caccagcacc aaggaggtgc tggacgccac cctgatccac cagagcatca ccggcctgta 6180
 cgagaccagg atcgacctga gccagctggg cggcgacagc agcccgcga agaagaagag 6240
 gaaggtgagc tggaaggacg ccagcggctg gagcaggatg tgaagcttga tcgttcaaac 6300
 atttggcaat aaagtttctt aagattgaat cctgttgccg gtottgcgat gattatcata 6360
 taattttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 6420

atgagatggg tttttatgat tagagtcccg caattataca ttttaatacgc gatagaaaac 6480
 aaaatatage ggcgaaacta ggataaatta tcgcgcgcgg tgtcatctat gttactagat 6540
 cttcgaaggg atcttttaaac atacgaacag atcaacttaa gttcttctga agcaacttaa 6600
 agttatcagg catgcatgga tcttgaggga atcagatgtg cagtcaggga ccatagcaca 6660
 ggacaggcgt cttctactgg tgctaccagc aaatgctgga agccgggaac actgggtacg 6720
 ttggaaacca cgtgatgtgg agtaagataa actgtaggag aaaagcattt cgtagtgggc 6780
 catgaagcct ttcaggacat gtattgcagt atgggccggc ccattacgca attggacgac 6840
 aacaaagact agtattagta ccacctcggc tatccacata gatcaaagct ggtttaaaag 6900
 agttgtgcag atgatccgtg gcagtcaacg tggagacagg ggtttttagag ctagaatatag 6960
 caagttaaaa taaggctagt ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt 7020
 ttttttcgga ccgcgcctgc agtgcagcgt gaccggctcg tgcccctctc tagagataat 7080
 gagcattgca tgtctaagtt ataaaaaatt accacatatt ttttttgtca cacttgtttg 7140
 aagtgcagtt tatctatctt tatacatata tttaaacttt actctacgaa taatataatc 7200
 tatagtacta caataatata agtgtttttag agaatcatat aaatgaacag ttagacatgg 7260
 tctaaaggac aattgagtat tttgacaaca ggactctaca gttttatctt tttagtgtgc 7320
 atgtgttctc cttttttttt gcaaatagct tcacctatat aatacttcat ccattttatt 7380
 agtacatcca tttaggggtt aggggttaatg gtttttatag actaattttt ttagtacatc 7440
 tattttattc tatttttagcc tctaaattaa gaaaactaaa actctatttt agttttttta 7500
 ttttaataatt tagatataaa atagaataaa ataaagtgac taaaaattaa acaaatatcc 7560
 ttttaagaaat taaaaaaact aaggaaacat ttttcttggt tcgagtagat aatgccagcc 7620
 tgttaaacgc cgtcgacgag tctaacggac accaaccagc gaaccagcag cgtcgcgtcg 7680
 ggccaagcga agcagacggc acggcatctc tgctcgtgcc tctggacccc tctcgagagt 7740
 tccgctccac cgttggaact gctccgctgt cggcatccag aaattgcgtg gcggagcggc 7800
 agacgtgagc cggcacggca ggcggcctcc tctcctctc acggcaccgg cagctacggg 7860
 ggattccttt ccaccgctc cttcgtttc ccttcctcgc ccgcgtaat aaatagacac 7920
 cccctccaca cctcttttcc ccaacctcgt gttgttcgga gcgcacacac acacaaccag 7980
 atctcccca aatccaccg tcggcaactc cgttcaagg tacgcgctc gtcctcccc 8040
 cccccctct ctacctctc tagatcggg ttccgggtcca tggttagggc ccggtagttc 8100
 tacttctgtt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct gctagcgttc 8160
 gtacacggat gcgacctgta cgtcagacac gttctgattg ctaacttgcc agtgtttctc 8220
 tttggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt catgattttt 8280
 tttgtttcgt tgcatagggt ttggtttgcc cttttccttt atttcaatat atgccgtgca 8340

cttgtttgtc gggatcatctt ttcattgcttt tttttgtctt ggttgatgatg atgtgggtctg 8400
 gttgggacggt cgttctagat cggagtagaa ttctgtttca aactacctgg tggatttatt 8460
 aattttggat ctgtatgtgt gtgccataca tttcatagat tacgaattga agatgatgga 8520
 tggaaatata gatctaggat aggtatacat gttgatgcgg gttttactga tgcatataca 8580
 gagatgcttt ttgttcgctt ggttgatgatg atgtgggtgtg gttgggacggt cgttcattcg 8640
 ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa ttttggaact 8700
 gtatgtgtgt gtcatacatc ttcattagta cgagtttaag atggatggaa atatcgatct 8760
 aggataggta tacatgttga tgtgggtttt actgatgatc atacatgatg gcatatgcag 8820
 catctattca tatgctctaa ccttgagtac ctatctatta taataaaca gtatgtttta 8880
 taattatttt gatcttgata tacttggtatg atggcatatg cagcagctat atgtggattt 8940
 ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt gtcgatgctc 9000
 accctgttgt ttggtgttac ttctgcaggg atccggcagc agccatgcag aagctgatca 9060
 acagcgtgca gaactacgcc tggggcagca agaccgccct gaccgagctg tacggcatgg 9120
 agaaccccag cagccagccc atggccgagc tgtggatggg cggccacccc aagagcagca 9180
 gccgcgtgca gaacgccgcc ggccgacatc tgagcctgcg cgacgtgatc gagagcgaca 9240
 agagcaccct gctgggagag gccgtggcca agcgttcgg cgagctgccc ttctgttca 9300
 aggtgctgtg cgccgccag cccctgagca tccaggtgca cccaacaag cacaacagcg 9360
 agatcggtt cgccaaggag aacgccgccg gcatcccat ggacgccgc gagcgcaact 9420
 acaaggaccc caaccacaag cccgagctgg tgttcgctt gaccccttc ctggccatga 9480
 acgcttccg cgagttcagc gagatcgtga gcctgctgca gccctgggcc ggcgcccacc 9540
 ccgcatcgc cacttcctg cagcagccc agccgagcg cctgagcgag ctgttcgcca 9600
 gcctgctgaa catgcagggc gaggagaaga gccgcgccct ggccatcctg aagagcgccc 9660
 tggacagcca gcagggcgag cctggcaga ccacgcgct gatcagcgag ttctacccc 9720
 aggacagcg cctgttcagc cccctgctgc tgaacgtggt gaagctgaac cccggcgagg 9780
 ccatgttctt gttcgccgag acccccaag cctacctga gggcgtggcc ctggaggtga 9840
 tggccaacag cgacaacgtg ctgcgcgcg gcctgacccc caagtacatc gacatcccc 9900
 agctggtggc caacgtgaag ttcgaggcca agccgcgcaa ccagctgctg acccagcccc 9960
 tgaagcaggg cgccgagctg gacttcccca tccccgtgga cgacttcgcc ttcagcctgc 10020
 acgacctgag cgacaaggag accaccatca gccagcagag cgccgccatc ctgttctgcg 10080
 tggagggcga cgccaccctg tgggaaggga gccagcagct gcagctgaag cccggcgaga 10140
 gcgccttcac cgccgccaac gagagccccg tgaccgtgaa gggccacggc cgcttggccc 10200
 gcgtgtacaa caagctgtga taggagctcg atccgtcgac ctgcagatcg ttcaaactt 10260

tggcaataaa gtttcttaag attgaatcct gttgccggtc ttgcgatgat tatcatataa 10320
 tttctgttga attacgttaa gcatgtaata attaacatgt aatgcatgac gttatttatg 10380
 agatggggttt ttatgattag agtcccgcga ttatacatgt aatacgcgat agaaaacaaa 10440
 atatagcgcg caaactagga taaattatcg cgcgcggtgt catctatgtt actagatcgg 10500
 cgcgcgcgcaa ttgaagtttg ggcggccagc atggccgtat ccgcaatgtg ttattaagtt 10560
 gtctaagcgt caatttggtt acaccacaat atatcctgcc accagccagc caacagctcc 10620
 ccgaccggca gtcggcgaca aaatcaccac tcgatacagg cagcccatca g 10671

<210> 35
 <211> 14475
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Casset biểu hiện của cấu trúc 22808

<220>
 <221> trình tự tăng cường
 <222> (165)..(256)
 <223> eNOS

<220>
 <221> vùng khởi động
 <222> (301)..(2097)
 <223> Ubi

<220>
 <221> gen
 <222> (2116)..(5190)
 <223> TNFw

<220>
 <221> trình tự kết thúc
 <222> (2116)..(5190)
 <223> tNOS

<220>
 <221> vùng khởi động
 <222> (5483)..(7475)
 <223> Ubi

<220>
 <221> gen
 <222> (7489)..(10563)
 <223> TNRv

<220>
 <221> trình tự kết thúc
 <222> (10570)..(10822)
 <223> tNOS

<220>
 <221> vùng khởi động
 <222> (10841)..(12832)
 <223> Ubi

<220>
 <221> gen
 <222> (12849)..(14027)
 <223> PMI

<220>
 <221> trình tự kết thúc
 <222> (14050)..(14302)
 <223> tNOS

<400> 35
 cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagtttaaac tgaaggcggg aaacgacaat ctgatcatga gcggagaatt 180
 aagggagtca cgttatgacc cccgccgatg acgcgggaca agccgtttta cgtttggaac 240
 tgacagaacc gcaacgctgc aggaattggc cgcagcggcc atttaaaca agcttggtac 300
 cattatgtgg tctaggtagg ttctatatat aagaaaactt gaaatgttct aaaaaaaaaat 360
 tcaagcccat gcatgattga agcaaacggg atagcaacgg tgtaaacctg atctagtgat 420
 ctcttgcaat ccttaacggc cacctaccgc aggtagcaaa cggcgtcccc ctctcgata 480
 tctccgcggc gacctctggc tttttccgcg gaattgcgcg gtggggacgg attccacgag 540
 accgcgacgc aaccgcctct cgcgcgtggg cccacaccg ctcggtgccg tagcctcaog 600
 ggactctttc tccctcctcc cccgttataa attggcttca tcccctcctt gcctcatcca 660
 tccaaatccc agtccccaat cccatccctt cgtaggagaa attcatogaa gctaagcgaa 720
 tcctcgcat cctctcaagg tactgcgagt ttctgatccc cctctcgacc cctcgtagt 780
 ttgtgtttgt cgtagcgttt gattaggtat gctttccctg tttgtgttcg tcgtagcggt 840
 tgattaggtg tgctttccct gtctgtgttc atcgtagtgt ttgattaggt cgtgtgaggg 900
 gatggcctgc tcgcgtcctt cgatctgtag tcgatttgcg ggtcgtgggt tagatctgcg 960
 ggctgtgatg aagttatttg gtgtgatctg ctgcctgat tctgcgggtt ggctcgagta 1020
 gatatgatgg ttggaccggt tggttcgttt accgcgctag ggttgggctg ggatgatgtt 1080
 gcatgcgccg ttgcgcgtga tccgcagca ggacttgcgt ttgattgcca gatctcgta 1140
 cgattatgtg atttgggttg gacttttttag atctgtagct tctgcttatg tgccagatgc 1200
 gcctactgct catatgcctg atgataatca taaatggctg tggaactaac tagttgattg 1260
 cggagtcatg tatcagctac aggtgtaggg actagctaca ggtgtaggga cttgcgtcta 1320
 attgtttggg cctttactca tgttgcaatt atgcaattta gtttagattg tttgttccac 1380
 tcatctaggg tgtaaaaggg aactgctta gattgctgtt taatcttttt agtagattat 1440
 attatattgg taacttatta cccctattac atgccatacg tgacttctgc tcatgcctga 1500
 tgataatcat agatcactgt ggaattaatt agttgattgt tgaatcatgt ttcattgtaca 1560
 taccacggca caattgctta gttccttaac aaatgcaaat tttactgatc catgtatgat 1620

ttgcgtgggtt ctctaattgtg aaataactata gctacttggt agtaagaatc aggttcgtat	1680
gcttaatgct gtatgtgcct tctgctcatg cctgatgata atcatatatc actggaatta	1740
attagttgat cgtttaatca tatatcaagt acataccatg ccacaatttt tagtcactta	1800
acccatgcag attgaactgg tccctgcatg ttttgctaaa ttgtttctatt ctgattagac	1860
catatatcat gtattttttt ttggtaatgg ttctcttatt ttaaattgcta tatagttctg	1920
gtacttggtta gaaagatctg cttcatagtt tagttgccta tccctcgaat taggatgctg	1980
agcagctgat cctatagctt tgtttcatgt atcaattctt ttgtgttcaa cagtcagttt	2040
ttgttagatt cattgtaact tatggtcgct tactcttctg gtccctcaatg cttgcaggat	2100
cgcggccgcg ccaccatggg aaaacctatt cctaattctc tgctgggcct ggattctacc	2160
ggaggcatgg cccctaagaa aaagcgggaag gtggacggcg gagtggacct gagaacactg	2220
ggatattctc agcagcagca ggagaagatc aagcccaagg tgagatctac agtggcccg	2280
caccacgaag ccctgggtggg acacggattt acacacgccc acattgtggc cctgtctcag	2340
caccctgccg ccctgggaac agtggccgtg aaatatcagg atatgattgc cgccctgcct	2400
gaggccacac acgaagccat tgtgggagtg ggaaaacagt ggtctggagc cagagccctg	2460
gaagccctgc tgacagtggc cggagaactg agaggacctc ctctgcagct ggatacagga	2520
cagctgctga agattgccaa aaggggcgga gtgaccgcgg tggaagccgt gcacgcctgg	2580
agaaatgcc tgacaggagc ccctctgaac ctgacccccg aacaggtggg gccattgcc	2640
agccacgacg gcggcaagca ggcctggaa accgtgcaga gactgctgcc cgtgctgtgc	2700
caggcccatg gcctgacacc tgaacagggtg gtggctatcg cctctcacga cggaggaaaa	2760
caggctctgg aaacagtgca gcggctgctg cctgtgctgt gtcaggctca cggcttgact	2820
ccagaacagg tggtggctat tgcttccaat attgggggga aacaggccct ggaaactgtg	2880
cagcgccctgc tgccagtgt gtgccaggct cagggaactga ccccggaaca ggtgggtggc	2940
attgccagca acaacggcg caagcaggcc ctggaaaccg tgcagagact gctgccctg	3000
ctgtgccagg cccatggcct gacacctgaa cagggtgttg ctatcgctc taacaacgga	3060
ggaaaacagg ctctggaaac agtgcagcgg ctgctgcctg tgctgtgtca ggctcacggc	3120
ttgactccag aacaggtggg ggctattgct tccaacaacg gggggaaaca ggccctggaa	3180
actgtgcagc gcctgctgcc agtgtgtgtc caggctcacg ggctgacccc cgaacagggtg	3240
gtggccattg ccagcaacgg cggcggcaag caggccctgg aaaccgtgca gagactgctg	3300
cccgctgtgt gccaggccca tggcctgaca cctgaacagg tggtggctat cgctctcac	3360
gacggaggaa aacaggctct ggaaacagtg cagcggctgc tgccgtgtgt gtgtcaggct	3420
cacggcttga ctccagaaca ggtgggtggct attgcttcca atattggggg gaaacaggcc	3480
ctggaaaactg tgcagcgct gctgccagt ctgtgccagg ctcacggcct cactcccgaa	3540

cagggtggtgg ccattgccag caacatcggc ggcaagcagg ccctggaaac cgtgcagaga 3600
 ctgctgcccg tgetgtgccg ggcccatggc ctgacacctg aacagggtgtt ggctatcgcc 3660
 tctcacgacg gaggaaaaca ggctctggaa acagtgcagc ggctgctgcc tgtgctgtgt 3720
 caggetcacg gcttgactcc agaacagggtg gtggctattg cttccaacaa cggggggaaa 3780
 caggccctgg aaactgtgca gcgcctgctg ccagtgtgtt gccagggtca cggactgacc 3840
 cccgaacagg tgggtggccat tgccagcaac ggcggcggca agcaggccct ggaaaccgtg 3900
 cagagactgc tgcccgtgct gtgccaggcc catggcctga cacctgaaca ggtggtggct 3960
 atcgccctcta acaacggagg aaaacaagca ctcgagacag tgcagcggct gctgcctgtg 4020
 ctgtgtcagg ctcacggctt gactccagaa cagggtggtg ctattgcttc caacaacggg 4080
 gggaaacagg ccctggaaac tgtgcagcgc ctgctgccag tgetgtgccg ggctcacggg 4140
 ctgacccccg aacagggtgtt ggccattgcc agcaacatcg gcggcaagca ggccctggaa 4200
 accgtgcaga gactgctgcc cgtgctgtgc caggcccatg gcctgacacc tgaacagggtg 4260
 gtggctatcg cctctaacaa cggaggaaaa caggctctgg aaacagtgcg gcggctgctg 4320
 cctgtgctgt gtcagggtca cggcttgact ccacagcagg tctgggcaat tgctagcaat 4380
 atcggcggac ggcccgccct ggagagcatt gtggcccagc tgtctagacc tgatcctgcc 4440
 ctggccgccc tgacaaatga tcacctggtg gccctggcct gtctgggagg cagacctgcc 4500
 ctggatgccg tgaaaaaagg actgcctcac gccctgccc tgattaaaag aacaaataga 4560
 agaatccccg agcggacctc tcacagagtg gccggatccc agctggtgaa atctgagctg 4620
 gaggagaaga agtctgagct gagacacaag ctgaagtacg tgccctcacga gtacatcgag 4680
 ctgatcgaga tcgccagaaa tagcaccag gatagaatcc tggagatgaa ggtgatggag 4740
 ttcttcatga aagtgtacgg ctacagagga aagcatctgg gaggaagcag aaaacctgac 4800
 ggagccattt atacagtggg cagccctatc gattatggcg tgatcgtgga tacaaaggcc 4860
 tacagcggag gctacaatct gcctattgga caggccgatg agatgcagag atacgtggag 4920
 gagaaccaa ccaggaacaa gcatatcaac cctaacgagt ggtggaaggt gtaccttct 4980
 agcgtgaccg agttcaagtt cctgtttgtg agcggccact tcaagggcaa ttataaggcc 5040
 cagctgacca ggtgaacca catcacaat tgtaatggcg ccgtgctgtc tgtggaggaa 5100
 ctgctgattg gaggagagat gattaaggcc ggaacactga cactggaggg ggtgagaaga 5160
 aagttcaaca acggcgagat caacttctga aagcttgatc gttcaaacat ttggcaataa 5220
 agtttcttaa gattgaatcc tgttgccggg cttgcgatga ttatcatata atttctgttg 5280
 aattacgtta agcatgtaat aattaacatg taatgcatga cgttatttat gagatgggtt 5340
 tttatgatta gagtcccga attatacatt taatacgcga tagaaaacaa aatatagcgc 5400
 gcaaactagg ataaattatc gcgcgcgggtg tcatttatgt tactagatct tcgaacccta 5460

gtogaagaca accggtgcat gcctgcagtg cagcgtgacc cggtcgtgcc cctctctaga 5520
 gataatgagc attgcatgtc taagttataa aaaattacca catatTTTTT ttgtcacact 5580
 tgtttgaagt gcagtttatc tatctttata catatatTTA aactttactc tacgaataat 5640
 ataatctata gtactacaat aatatcagtg ttttagagaa tcatataaat gaacagttag 5700
 acatgggtcta aaggacaatt gagtatTTTg acaacaggac tctacagttt tatctTTTTa 5760
 gtgtgcatgt gttctccttt ttttttgcaa atagcttcac ctatataata cttcatccat 5820
 tttattagta catccattta gggtttaggg ttaatggttt ttatagacta attttttttag 5880
 tacatctatt ttattctatt ttagcctcta aattaagaaa actaaaactc tatttttagtt 5940
 tttttattta ataatttaga tataaaaatag aataaaaataa agtgactaaa aattaaacaa 6000
 atacccttta agaaattaaa aaaactaagg aaacattttt cttgtttcga gtagataatg 6060
 ccagcctgtt aaacgccgtc gacgagtcta acggacacca accagcgaac cagcagcgtc 6120
 gcgtcggggc aagcgaagca gacggcacgg catctctgtc gctgcctctg gacccctctc 6180
 gagagtccg ctccaccgtt ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg 6240
 agcggcagac gtgagccggc acggcaggcg gcctcctcct cctctcacgg caccggcagc 6300
 tacgggggat tcctttccca ccgtccttc gtttccctt cctcgcccgc cgtaataaat 6360
 agacaccccc tccacaccct ctttcccaa cctcgtgttg ttcggagcgc acacacacac 6420
 aaccagatct ccccaaactc caccgctcg cactcgcct tcaaggtagc ccgctcgtcc 6480
 tcccccccc ccctctcta cttctcttag atcggcgctt cgggtccatgg ttagggcccc 6540
 gtagttctac ttctgttcat gtttgtgtta gatccgtgtt tgtgttagat ccgtgctgct 6600
 agcgttcgta cacggatgcg acctgtacgt cagacacgtt ctgattgcta acttgccagt 6660
 gtttctcttt ggggaatcct gggatggctc tagccgttcc gcagacggga tcgatttcat 6720
 gatttttttt gtttcgttgc atagggtttg gtttgccctt ttcctttatt tcaatatatg 6780
 ccgtgcactt gtttgtcggg tcatcttttc atgctttttt ttgtcttggt tgtgatgatg 6840
 tggctctggtt gggcggctgt tctagatcgg agtagaatc tgtttcaaac tacctggtgg 6900
 atttattaat tttgatctg tatgtgtgtg ccatacatat tcatagttac gaattgaaga 6960
 tgatggatgg aaatatcgat ctaggatagg tatacatggt gatgcggggt ttactgatgc 7020
 atatacagag atgctttttg ttcgcttggt tgtgatgatg tgggtgtggtt gggcggctgt 7080
 tcattcgttc tagatcggag tagaatactg tttcaaacta cctggtgtat ttattaattt 7140
 tggaactgta tgtgtgtgtc atacatcttc atagttacga gtttaagatg gatggaaata 7200
 tcgatctagg ataggtatac atgttgatgt gggttttact gatgcatata catgatggca 7260
 tatgcagcat ctattcatat gctotaacct tgagtacct tctattataa taaacaagta 7320
 tgttttataa ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg 7380

tggatttttt tagccctgcc ttcatacgt atttatttgc ttggtactgt ttcttttgc	7440
gatgetcacc ctgttgtttg gtgttacttc tgcagcggcc gcgccaccat gggaaaacct	7500
attcctaate ctctgctggg cctggattct accggaggca tggcccctaa gaaaaagcgg	7560
aaggtggacg gcggagtgga cctgagaaca ctgggatatt ctacgcagca gcaggagaag	7620
atcaagccca aggtgagatc tacagtggcc cagcaccacg aagccctggt gggacacgga	7680
tttacacacg cccacattgt ggccctgtct cagcaccctg ccgccctggg aacagtggcc	7740
gtgaaatata aggatatgat tgccgccctg cctgaggcca cacacgaagc cattgtggga	7800
gtgggaaaac agtggctctg agccagagcc ctggaagccc tgctgacagt ggccggagaa	7860
ctgagaggac ctctctgca gctggataca ggacagctgc tgaagattgc caaaaggggc	7920
ggagtgaccg cggtggaagc cgtgcacgcc tggagaaatg ccctgacagg agcccctctg	7980
aacctgacct ccgaacaggt ggtggccatt gccagcaaca acggcggcaa gcaggccctg	8040
gaaaccgtgc agagactgct gcccgtgctg tgccaggccc atggcctgac acctgaacag	8100
gtggtggcta tcgcctctca cgacggagga aaacaggctc tggaaacagt gcagcggctg	8160
ctgcctgtgc tgtgtcaggc tcacggcttg actccagaac aggtggtggc tattgcttcc	8220
aacggcgggg ggaaacaggc cctggaaact gtgcagcgc tgctgccagt gctgtgccag	8280
gctcacggac tgacccccga acagggtggtg gccattgcca gcaacggcgg cggcaagcag	8340
gccctggaaa ccgtgcagag actgctgccc gtgctgtgcc aggcccatgg cctgacacct	8400
gaacagggtg tggctatcgc ctctcacgac ggaggaaaac aggtctctgga aacagtgcag	8460
cggctgctgc ctgtgctgtg tcaggctcac ggcttgactc cagaacaggt ggtggctatt	8520
gcttcccacg acggggggaa acaggccctg gaaactgtgc agcgctgct gccagtgtg	8580
tgccaggctc acgggctgac ccccgaacag gtggtggcca ttgccagcaa cggcggcggc	8640
aagcaggccc tggaaaccgt gcagagactg ctgccctgc tgtgccaggc ccatggcctg	8700
acacctgaac aggtggtggc tatgcctct aacggcggag gaaaacaggc tctggaaaca	8760
gtgcagcggc tgctgctgt gctgtgtcag gctcacggct tgactocaga acagggtggtg	8820
gctattgctt cccacgacgg ggggaaacag gccctggaaa ctgtgcagcg cctgtgccca	8880
gtgctgtgcc aggtcacgg cctcactccc gaacagggtg tggccattgc cagcaacaac	8940
ggcggcaagc aggccttgga aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat	9000
ggcctgacac ctgaacaggt ggtggctatc gcctctcacg acggaggaaa acaggctctg	9060
gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag	9120
gtggtggcta ttgcttccca cgacgggggg aaacaggccc tggaaactgt gcagcgcctg	9180
ctgccagtgc tgtgccaggc tcacggactg acccccgaac aggtggtggc cattgccagc	9240
aacatcggcg gcaagcaggc cctggaaacc gtgcagagac tgctgccctg gctgtgccag	9300

gcccatggcc tgacacctga acagggtggtg gctatcgccct ctaacaacgg aggaaaaacaa 9360
 gcactcgaga cagtgcagcg gctgctgcct gtgctgtgtc aggtcacagg cttgactcca 9420
 gaacagggtg tggtatttgc ttccaacggc ggggggaaac aggccctgga aactgtgcag 9480
 cgctgctgc cagtgtgtg ccaggctcac gggctgaccc ccgaacagg ggtggccatt 9540
 gccagccacg acggcggaac gcaggccctg gaaaccgtgc agagactgct gcccgctgtg 9600
 tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctaa tatcggagga 9660
 aaacaggctc tggaaacagt gcagcggctg ctgcctgtgc tgtgtcaggc tcacggcttg 9720
 actccacagc aggtcgtggc aattgctagc cacgacggcg gacggcccg cctggagagc 9780
 attgtggccc agctgtctag acctgacct gccctggccg ccctgacaaa tgatcacctg 9840
 gtggccctgg cctgtctggg aggcagacct gccctggatg ccgtgaaaaa aggactgcct 9900
 cagccccctg ccctgattaa aagaacaaat agaagaatcc ccgagcggac ctctcacaga 9960
 gtggccggat ccagctggt gaaatctgag ctggaggaga agaagtctga gctgagacac 10020
 aagctgaagt acgtgcctca cgagtacatc gagctgatcg agatcgccag aaatagcacc 10080
 caggatagaa tcttgagat gaaggatgat gagttcttca tgaaagtgtc cggctacaga 10140
 ggaaagcatc tgggaggaag cagaaaacct gacggagcca ttatacagt gggcagccct 10200
 atcgattatg gcgtgatcgt ggatacaaag gcctacagcg gaggctacaa tctgcctatt 10260
 ggacaggccg atgagatgca gagatacgtg gaggagaacc aaaccaggaa caagcatatc 10320
 aaccctaacg agtgggtgaa ggtgtaccct tctagcgtga ccgagttcaa gttcctgttt 10380
 gtgagcggcc acttcaaggg caattataag gccagctga ccaggctgaa ccacatcaca 10440
 aattgtaatg gcgccgtgct gtctgtggag gaactgctga ttggaggaga gatgattaag 10500
 gccggaacac tgacactgga ggagggtgaga agaaagttca acaacggcga gatcaacttc 10560
 tgaaagcttg atcgttcaaa catttgga taaagtttct taagattgaa tctgttgcc 10620
 ggtcttgca tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac 10680
 atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac 10740
 atttaatacg cgatagaaaa caaatatag cgcgcaaact aggataaatt atcgcgcgcg 10800
 gtgtcatcta tgttactaga tcttogaaga cggaccgcgc ctgcagtgca gcgtgacccg 10860
 gtcgtgcccc tctctagaga taatgagcat tgcattgcta agttataaaa aattaccaca 10920
 tttttttttt gtcacacttg tttgaagtgc agtttatcta tctttataca tatatttaaa 10980
 ctttactcta cgaataatat aatctatagt actacaataa tatcagtgtt ttagagaatc 11040
 atataaatga acagtttagac atggtctaaa ggacaattga gtattttgac aacaggactc 11100
 tacagtttta tcttttttagt gtgcatgtgt tctccttttt ttttgcaaat agcttcacct 11160
 atataatact tcatccattt tattagtaca tccatttagg gtttaggggtt aatggttttt 11220

atagactaat ttttttagta catctatttt attctatttt agcctctaaa ttaagaaaac	11280
taaaactcta ttttagtttt tttattttaat aatttagata taaaatagaa taaaataaag	11340
tgactaaaaa ttaaacaaat accctttaag aaattaaaaa aactaaggaa acatttttct	11400
tgtttcgagt agataatgcc agcctgttaa acgccgtcga cgagtctaac ggacaccaac	11460
cagcgaacca gcagcgtcgc gtcggggccaa gcgaagcaga cggcacggca tctctgtcgc	11520
tgcctctgga cccctctcga gagttccgct ccaccgttgg acttgctccg ctgtcggcat	11580
ccagaaaattg cgtggcggag cggcagacgt gagccggcac ggcaggcggc ctctcctcc	11640
tctcacggca ccggcagcta cgggggattc ctttcccacc gctccttcgc tttcccttcc	11700
tcgcccgcg taataaatag acacccctc cacaccctct tcccccaacc tcgtgttggt	11760
cggagcgcac acacacacaa ccagatctcc cccaaatcca cccgtcggca cctccgcttc	11820
aaggtacgcc gctcgtctc ccccccccc ctctctacct tctctagatc ggcgttccgg	11880
tccatggtta gggcccggta gttctacttc tgttcatgtt tgtgttagat ccgtgtttgt	11940
gttagatccg tgctgctagc gttcgtacac ggatgcgacc tgtacgtcag acacgttctg	12000
attgctaact tgccagtgtt tctctttggg gaatcctggg atggctctag ccgttccgca	12060
gacgggatcg atttcatgat tttttttgtt tcgttgcata gggtttggtt tgcccttttc	12120
ctttatttca atatatgccg tgcacttggt tgcgggtca tcttttcatg ctttttttg	12180
tcttggttgt gatgatgtgg tctggttggg cggtcgttct agatcggagt agaattctgt	12240
ttcaaactac ctggtggatt tattaatttt ggatctgtat gtgtgtgcc aacatattca	12300
tagttacgaa ttgaagatga tggatggaaa tatcgatcta ggataggtat acatgttgat	12360
gcggttttta ctgatgcata tacagagatg ctttttgttc gcttggttgt gatgatgtgg	12420
tgtggttggg cggtcgttca ttctgtctag atcggagtag aatactgttt caaactacct	12480
ggtgtattta ttaatttttg aactgtatgt gtgtgtcata catcttcata gttacgagtt	12540
taagatggat ggaaatatcg atctaggata ggtatacatg ttgatgtggg ttttactgat	12600
gcataacat gatggcatat gcagcatcta ttcatatgct ctaaccttga gtacctatct	12660
attataataa acaagtatgt tttataatta ttttgatctt gatatacttg gatgatggca	12720
tatgcagcag ctatatgtgg attttttttag ccctgccttc atacgtatt tatttgcttg	12780
gtactgtttc ttttgctgat gctcaccctg ttgtttggtg ttacttctgc agggatccgg	12840
cagcagccat gcagaagctg atcaacagcg tgcagaacta cgctgggggc agcaagaccg	12900
ccctgacoga gctgtacggc atggagaacc ccagcagcca gcccatggcc gagctgtgga	12960
tgggcgcca cccaagagc agcagccgcg tgcagaacgc cgccggcgac atcgtgagcc	13020
tgcgcgacgt gatcgagagc gacaagagca ccctgctggg cgaggccgtg gccaagcgct	13080
tcggcgagct gcccttcctg ttcaaggtgc tgtgcgcgc ccagcccctg agcatccagg	13140

tgcaccccaa caagcacaac agcgagatcg gcttcgccaa ggagaacgcc gccggcatcc 13200
 ccatggagcg cgccgagcg aactacaagg accccaacca caagcccagag ctggtgttcg 13260
 ccoctgacccc ctctctggcc atgaacgcct tccgcgagtt cagcgagatc gtgagcctgc 13320
 tgcagcccgt ggccggcgcc caccgcgcca tgcgccactt cctgcagcag cccgacgccg 13380
 agcgccctgag cgagctgttc gccagcctgc tgaacatgca gggcgaggag aagagccgcg 13440
 ccctggccat cctgaagagc gccctggaca gccagcaggg cgagccctgg cagaccatcc 13500
 gcctgatcag cgagttctac cccgaggaca gcggcctgtt cagccccctg ctgctgaacg 13560
 tgggtgaagct gaaccccgcc gaggccatgt tctgttcgc cgagaccccc cagccctacc 13620
 tgcagggcgt ggccctggag gtgatggcca acagcgacaa cgtgctgcgc gccggcctga 13680
 cccccaagta catcgacatc cccgagctgg tggccaacgt gaagtctgag gccaaagccc 13740
 ccaaccagct gctgacccag cccgtgaagc agggcgccga gctggacttc cccatccccg 13800
 tggacgactt cgccttcagc ctgcacgacc tgagcgacaa ggagaccacc atcagccagc 13860
 agagcgccgc catcctgttc tgcgtggagg gcgacgccac cctgtggaag ggcagccagc 13920
 agctgcagct gaagcccggc gagagcgccct tcacgcgcgc caacgagagc cccgtgaccg 13980
 tgaagggcca cggccgcctg gcccgcggtgt acaacaagct gtgataggag ctcgatccgt 14040
 cgacctgcag atcgttcaaa catttgga taaagtttct taagattgaa tctgtttgcc 14100
 ggtcttgcca tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac 14160
 atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac 14220
 atttaatacg cgatagaaaa caaaatatag cgcgcaaact aggataaatt atcgcgcgcg 14280
 gtgtcatcta tgttactaga tcggcgcgcc gcaattgaag tttggggcggc cagcatggcc 14340
 gtatccgcaa tgtgttatta agttgtctaa gcgtcaattt gtttacacca caatatatcc 14400
 tgccaccagc cagccaacag ctccccgacc ggcagctcgg cacaaaatca ccaactcgata 14460
 caggcagccc atcag 14475

<210> 36

<211> 10668

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 22873

<220>

<221> trình tự tăng cường

<222> (165)..(256)

<223> eNOS

<220>

<221> vùng khởi động

<222> (301)..(2097)
 <223> Ubi

<220>
 <221> gen
 <222> (2114)..(6280)
 <223> Cas9m

<220>
 <221> đặc điểm_hỗn tạp
 <222> (6215)..(6271)
 <223> NLS-TAG

<220>
 <221> trình tự kết thúc
 <222> (6286)..(6538)
 <223> tNOS

<220>
 <221> vùng khởi động
 <222> (6545)..(6919)
 <223> pU3

<220>
 <221> đặc điểm_hỗn tạp
 <222> (6921)..(6938)

<220>
 <221> đặc điểm_hỗn tạp
 <222> (6921)..(7023)
 <223> sgRNA

<220>
 <221> vùng khởi động
 <222> (7034)..(9025)
 <223> Ubi

<220>
 <221> gen
 <222> (9042)..(10220)
 <223> PMI

<220>
 <221> trình tự kết thúc
 <222> (10243)..(10495)
 <223> tNOS

<400> 36
 cctgtggttg gcatgcacat acaaattggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagttttaaac tgaaggcggg aaacgacaat ctgatcatga gcggagaatt 180
 aaggagagtca cgttatgacc ccgcgcgatg acgcgggaca agccgtttta cgtttggaac 240
 tgacagaacc gcaacgctgc aggaattggc cgcagcggcc atttaaacia agcttggtac 300
 cattatgttg tctaggtagg ttctatatat aagaaaactt gaaatgttct aaaaaaaaaat 360
 tcaagcccat gcatgattga agcaaacggt atagcaacgg tgtaaacctg atctagtgat 420
 ctcttgcaat cottaacggc cacctaccgc aggtagcaaa cggcgtcccc ctccctcgata 480

tctccgcggc gacctctggc tttttccgcg gaattgcgcg gtggggacgg attccacgag 540
 accgcgacgc aaccgcctct cgccgctggg cccacacccg ctcggtgccg tagcctcacg 600
 ggactctttc tccctcctcc cccgttataa attggcttca tccctcctt gcctcatcca 660
 tccaaatccc agtccccaat cccatccctt cgtaggagaa attcatcgaa gctaagcgaa 720
 tctcgcgat cctctcaagg tactgcgagt ttctgatccc cctctcgacc cctcgtatgt 780
 ttgtgtttgt cgtagcgttt gattaggtat gctttccctg tttgtgttcg tcgtagcgtt 840
 tgattaggta tgctttccct gtctgtgttc atcgtagtgt ttgattaggc cgtgtgaggg 900
 gatggcctgc tcgcgtcctt cgatctgtag tcgatttgcg ggtcgtgggt tagatctgcg 960
 ggctgtgatg aagttatttg gtgtgatctg ctgcctgat tctgcgggtt ggctcgagta 1020
 gatatgatgg ttggaccggg tggttcgttt accgcgctag ggttgggctg ggatgatgtt 1080
 gcatgcgccg ttgcgcgtga tccgcagca ggacttgctg ttgattgcca gatctcgta 1140
 cgattatgtg atttggtttg gacttttttag atctgtagct tctgcttatg tgccagatgc 1200
 gcctactgct catatgcctg atgataatca taaatggctg tggaactaac tagttgattg 1260
 cggagtcatg tatcagctac aggtgtaggg actagctaca ggtgtaggga cttgcgtcta 1320
 attgtttggg cctttactca tgttgcaatt atgcaattta gtttagattg tttgttccac 1380
 tcatctaggc tgtaaaaggg aactgctta gattgctgtt taatcttttt agtagattat 1440
 attatattgg taacttatta cccctattac atgccatacg tgacttctgc tcatgcctga 1500
 tgataatcat agatcactgt ggaattaatt agttgattgt tgaatcatgt tcatgtaca 1560
 taccacggca caattgctta gttccttaac aaatgcaaat ttactgato catgtatgat 1620
 ttgcgtgggt ctctaattgt aaatactata gctacttggt agtaagaatc aggttcgtat 1680
 gcttaatgct gtatgtgctt tctgctcatg cctgatgata atcatatata actggaatta 1740
 attagttgat cgtttaataca tatatcaagt acataccatg ccacaatttt tagtcactta 1800
 acccatgcag attgaactgg tccctgcatg ttttgctaaa ttgtttctatt ctgattagac 1860
 catatatcat gtattttttt ttggtaatgg ttctcttatt ttaaagcta tatagttctg 1920
 gtacttgcta gaaagatctg ctccatagtt tagttgcta tccctcgaat taggatgctg 1980
 agcagctgat cctatagctt tgtttcatgt atcaattctt ttgtgttcaa cagtcagttt 2040
 ttgttagatt cattgtaact tatggctgct tactcttctg gtcctcaatg cttgcaggat 2100
 cgcgccgct aaaatggata agaagtactc tattggcctc gatattggaa ccaactctgt 2160
 gggctgggac gtgatcaccg atgagtacaa ggtgccatct aagaagttca aggtgctcgg 2220
 caacaccgat aggcactcta tcaagaagaa cctcatcggg gccctcctct tcgattctgg 2280
 cgagaccgcc gaggccacca ggctcaagag gaccgccagg aggaggtaca ccaggaggaa 2340
 gaacaggatc tgctacctcc aggagatctt ctctaacgag atggccaagg tggatgattc 2400

tttcttccac aggctcgagg agtctttcct cgtggaggag gataagaagc acgagaggca	2460
cccaatcttc ggcaacatcg tggatgaggt ggctaccac gagaagtacc caaccatcta	2520
ccacctcagg aagaagctcg tggattctac cgataaggcc gatctcaggc tcctctacct	2580
cgccctcgcc cacatgatca agttcagggg ccacttcctc atcgagggcg atctcaaccc	2640
agataactct gatgtggata agctcttcat ccagctcggtg cagacctaca accagctctt	2700
cgaggagAAC ccaatcaacg cctctggcgt ggatgccaaG gccatcctct ctgccaggct	2760
ctctaagtct aggaggtcgc agaacctcat cgcccagctc ccaggcgaga agaagaacgg	2820
cctcttcggc aacctcatcg ccctctctct cggcctcacc ccaaacttca agtctaactt	2880
cgatctcgcc gaggatgcc agctccagct ctctaaggat acctacgatg atgatctcga	2940
taacctctc gcccagatcg gcgatcagta cgccgatctc ttctctgccg ccaagaacct	3000
ctctgatgcc atctctctct ctgatatctt caggggtgaac accgagatca ccaaggcccc	3060
actctctgcc tctatgatca agaggtacga tgagcaccac caggatctca ccctcctcaa	3120
ggccctcgtg aggcagcagc tcccagagaa gtacaaggag atcttcttcg atcagtctaa	3180
gaacggctac gccggctaca tcgatggcgg cgctctcag gaggagttct acaagttcat	3240
caagccaatc ctcgagaaga tggatggcac cgaggagctc ctctggaagc tcaacaggga	3300
ggatctctc aggaagcaga ggaccttcga taacggctct atcccacacc agatccacct	3360
cggcgagctc cagcccatcc tcaggaggca ggaggatttc taccattcc tcaaggataa	3420
cagggagaag atcgagaaga ttctcacctt ccgcattcca tactacgtgg gccactcgc	3480
caggggcaac tctaggttcg cctggatgac caggaagtct gaggagacca tcccccttg	3540
gaacttcgag gaggtggtgg ataagggcgc ctctgcccag tctttcatcg agaggatgac	3600
caacttcgat aagaacctcc caaacgagaa ggtgctccca aagcactctc ttctctacga	3660
gtacttcacc gtgtacaacg agctcaccaa ggtgaagtac gtgaccgagg gcatgaggaa	3720
gccagccttc ctctctggcg agcagaagaa ggccatcgtg gatctctct tcaagaccaa	3780
caggaaggtg accgtgaagc agctcaagga ggattacttc aagaagatcg agtgcttcga	3840
ttctgtggag atctctggcg tggaggatag gttcaacgcc tctctcggca cctaccacga	3900
tctctcaag atcatcaagg ataaggattt cctcgataac gaggagaacg aggatatcct	3960
cgaggatatc gtgctcacc tcacctctt cgaggatagg gagatgatcg aggagaggct	4020
caagacctac gccacctct tcgatgataa ggtgatgaag cagctcaaga ggaggaggta	4080
caccggctgg ggcaggctct ctaggaagct catcaacggc atcagggata agcagtctgg	4140
caagaccatc ctcgatttcc tcaagtctga tggcttcgcc aacaggaaact tcatgcagct	4200
catccacgat gattctctca ccttcaagga ggatatccag aaggctcagg tgtctggcca	4260
gggccactct ctccacgagc agatcgccaa cctcgccggc tctccagcca tcaagaaggg	4320

catcctccag accgtgaagg tgggtggatga gctcgtgaag gtgatgggcc acaagccaga 4380
gaacatcgtg atcgagatgg ccagggagaa ccagaccacc cagaagggcc agaagaactc 4440
tagggagagg atgaagagga tcgaggaggg catcaaggag ctcggtcttc agatcctcaa 4500
ggagcaccga gtggagaaca ccagctcca gaacgagaag ctctacctct actacctcca 4560
gaacggcagg gatatgtacg tggatcagga gctcgatata aacaggctct ctgattacga 4620
tgtggatcac atcgtgccac agtctttcct caaggatgat tctatcgata acaagggtgt 4680
caccagggtct gataagaaca ggggcaagtc tgataacgtg ccatctgagg aggtgggttaa 4740
gaagatgaag aactactgga ggcagctcct caacgccaaag ctcatcacc agaggaagtt 4800
cgataacctc accaaggccg agaggggagg cctctctgag ctcgataagg cgggttcat 4860
caagaggcag ctctggaga ccaggcagat cactaagcac gtggcccaga tctcgtatc 4920
taggatgaac accaagtacg atgagaacga taagctcatc agggagggtga aggtgatcac 4980
cctcaagtct aagctcgtgt ctgatttcag gaaggatttc cagttctaca aggtgagga 5040
gatcaacaac taccaccacg ccacgatgc ctacctcaac gccgtgggtg gcaccgcct 5100
catcaagaag taccaaaaac tcgagtctga gttcgtgtac ggcgattaca aggtgtacga 5160
tgtgaggaag atgatcgcca agtctgagca ggagatcggc aaggccaccg ccaagtactt 5220
cttctactct aacatcatga acttcttcaa gaccgagatc accctcgcca acggcgagat 5280
caggaagagg ccactcatcg agaccaacgg cgagaccggc gagatcgtgt gggataaggg 5340
cagggatttc gccaccgtga ggaagggtgt ctctatgcca caggtgaaca tcgtgaagaa 5400
gaccgagggtg cagaccggcg gcttctctaa ggagtctatc ctcccaaaga ggaactctga 5460
taagctcatc gccaggaaga aggattggga ccgaagaag tacggcgggt tcgacagccc 5520
gaccgtggcc tacagcgtgc tgggtgggtg caagggtggag aagggaaga gcaagaagct 5580
gaagagcgtg aaggagctgg tgggcatcac catcatggag aggagcagct tcgagaagaa 5640
cccagtggaac ttcttgagg ccaagggtta caaggagggt aagaaggacc tgatcattaa 5700
actgccgaag tacagcctgt tcgagctgga gaacggcagg aagaggatgc tggccagcgc 5760
cggcgagctg cagaagggca acgagctggc cctgccgagc aagtacgtga acttctgtga 5820
cctggccagc cactacgaga agctgaaggg cagcccgag gacaacgagc agaagcagct 5880
gttctgtggag cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag 5940
caagaggggtg atcctggccg acgccaacct ggacaagggt ctgagcgcct acaacaagca 6000
cagggacaag ccgatcagg agcaggccga gaacatcatc cacctgttca cctgaccaa 6060
cctgggcgc cgggcgcct tcaagtactt cgacaccacc atcgacagga agaggtacac 6120
cagcaccaag gaggtgctgg acgccaccct gatccaccag agcatcaccg gcctgtacga 6180
gaccaggatc gacctgagcc agctgggagg cgacagcagc ccgccgaaga agaagaggaa 6240

ggtgagctgg aaggacgcca gcggctggag caggatgtga agcttgatcg ttcaaacatt 6300
 tggcaataaa gtttcttaag attgaatcct gttgccggtc ttgcgatgat tatcatataa 6360
 tttctgttga attacgttaa gcatgttaata attaacatgt aatgcatgac gttattttatg 6420
 agatggggttt ttatgattag agtcccgcga ttatacathtt aatacgcgat agaaaacaaa 6480
 atatagcgcg caaactagga taaattatcg cgcgcgggtg catctatggt actagatcctt 6540
 cgaaggggatc tttaaacata cgaacagatc acttaaagtt cttctgaagc aacttaaagt 6600
 tatcaggcat gcatggatct tggaggaatc agatgtgcag tcagggacca tagcacagga 6660
 caggcgtctt ctactggtgc taccagcaaa tgctggaagc cgggaacact gggtagcttg 6720
 gaaaccacgt gatgtggagt aagataaact gtaggagaaa agcatttcgt agtgggcat 6780
 gaagcctttc aggacatgta ttgcagtatg ggccggccca ttacgcaatt ggacgacaac 6840
 aaagactagt attagtacca cctcggctat ccacatagat caaagctggt taaaagagt 6900
 tgtgcagatg atccgtggca gtcaacgtgg agacaggggt tttagagcta gaaatagcaa 6960
 gttaaaataa ggctagtccg ttatcaactt gaaaaagtgg caccgagtcg gtgctttttt 7020
 tttcggaccg cgcctgcagt gcagcgtgac ccggctcgtc ccctctctag agataatgag 7080
 cattgcatgt ctaagttata aaaaattacc acatattttt tttgtcacac ttgtttgaag 7140
 tgcagtttat ctatctttat acatatattt aaactttact ctacgaataa tataatctat 7200
 agtactacaa taatatcagt gtttttagaga atcatataaa tgaacagtta gacatggtct 7260
 aaaggacaat tgagtatttt gacaacagga ctctacagtt ttatcttttt agtgtgcatg 7320
 tgttctcctt tttttttgca aatagcttca cctatataat acttcatoca ttttattagt 7380
 acatccattt agggtttagg gttaatggtt tttatagact aattttttta gtacatctat 7440
 tttattctat tttagcctct aaattaagaa aactaaaact ctattttagt ttttttattt 7500
 aataatttag atataaaata gaataaaata aagtgactaa aaattaaaca aatacccttt 7560
 aagaaattaa aaaaactaag gaaacatttt tcttgtttcg agtagataat gccagcctgt 7620
 taaacgcgt cgacgagtct aacggacacc aaccagcgaa ccagcagcgt cgcgtcgggc 7680
 caagcgaagc agacggcacg gcatctctgt cgctgcctct ggacccctct cgagagttcc 7740
 gctccaccgt tggacttgct ccgctgtcgg catccagaaa ttgcgtggcg gagcggcaga 7800
 cgtgagccgg caccggcagg ggctcctcc tctctcacg gcaccggcag ctacggggga 7860
 ttcctttccc accgctcctt cgttttccct tctcgcctcg ccgtaataaa tagacacccc 7920
 ctccacaccc tctttcccca acctcgtggt gttcggagcg cacacacaca caaccagatc 7980
 tccccaaat ccaccgctcg gcacctccgc ttcaaggtag gccgctcgtc ccccccccc 8040
 cccctctcta ccttctctag atcggcggtt cgggtccatgg ttagggcccg gtagttctac 8100
 ttctgttcat gtttgtgtta gatccgtggt tgtgttagat ccgtgctgct agcgttcgta 8160

cacggatgcg acctgtacgt cagacacggt ctgattgcta acttgccagt gtttctcttt	8220
ggggaatcct gggatggctc tagccgttcc gcagacggga togatttcat gatttttttt	8280
gtttcggttgc ataggggttg gtttgccott ttcttttatt tcaatatatg ccgtgcactt	8340
gtttgtcggg tcatcttttc atgctttttt ttgtcttggt tgtgatgatg tggctcgggt	8400
gggcggctgt tctagatcgg agtagaattc tgtttcaaac tacctgggtg atttattaat	8460
tttggatctg tatgtgtgtg ccatacatat tcatagttaac gaattgaaga tgatggatgg	8520
aaatatcgat ctaggatagg tatacatgtt gatgcgggtt ttactgatgc atatacagag	8580
atgctttttg ttcgcttggt tgtgatgatg tgggtgtggt gggcggctgt tcattcgttc	8640
tagatcggag tagaatactg tttcaaacta cctggtgtat ttattaattt tggaactgta	8700
tgtgtgtgtc atacatcttc atagttaacga gtttaagatg gatggaaata tcgatctagg	8760
ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcagcat	8820
ctattcatal gctctaacct tgagtacctt tctattataa taaacaagta tgttttataa	8880
ttattttgat ctlgatatac ttggatgatg gcatatgcag cagctatatg tggatttttt	8940
tagccctgcc ttcatacgtt atttatttgc ttggtactgt ttcttttgtc gatgctcacc	9000
ctgttggttg gtgttacttc tgcagggatc cggcagcagc catgcagaag ctgatcaaca	9060
gcgtgcagaa ctacgcctgg ggcagcaaga ccgccctgac cgagctgtac ggcattggaga	9120
accccgagcag ccagcccatg gccgagctgt ggatgggcgc ccacccaag agcagcagcc	9180
gcgtgcagaa cgccgcccgc gacatcgtga gcctgcgcga cgtgatcgag agcgacaaga	9240
gcaccctgct gggcgaggcc gtggccaago gcttcggcga gctgcccttc ctgttcaagg	9300
tgctgtgcgc cgcccagccc ctgagcatcc aggtgcaccc caacaagcac aacagcgaga	9360
tgggcttcgc caaggagaac gccgcccga tcccatgga cgccgccgag cgcaactaca	9420
aggaccccaa ccacaagccc gagctggtgt tcgcctgac ccccttcctg gccatgaacg	9480
ccttcgcga gttcagcgag atcgtgagcc tgctgcagcc cgtggccggc gccaccccg	9540
ccatgcceca cttcctgcag cagcccgacg ccgagcgctt gagcgagctg ttccgcagcc	9600
tgctgaacat gcagggcgag gagaagagcc gcgcctggc catcctgaag agcgccctgg	9660
acagccagca gggcgagccc tggcagacca tccgcctgat cagcgagttc taccgcagg	9720
acagcggcct gttcagcccc ctgctgctga acgtggtgaa gctgaacccc ggcgaggcca	9780
tgttcctgtt cgccgagacc ccccaagcct acctgcaggg cgtggccctg gaggtgatgg	9840
ccaacagcga caacgtgctg cgcgccggcc tgaccccaa gtacatcgac atccccgagc	9900
tgggtggcaa cgtgaagtgc gaggcgaagc ccgccaacca gctgctgacc cagcccgtga	9960
agcagggcgc cgagctggac ttcccatcc ccgtggacga cttcgccttc agcctgcacg	10020
acctgagcga caaggagacc accatcagcc agcagagcgc cgccatcctg ttctgcgtgg	10080

```

agggcgacgc caccctgtgg aagggcagcc agcagctgca gctgaagccc ggcgagagcg 10140
ccttcacgc cgccaacgag agccccgtga cctgaaggg ccacggccgc ctggcccgcg 10200
tgtacaacaa gctgtgatag gagctcgatc cgtcgacctg cagatcgttc aaacatttgg 10260
caataaagtt tcttaagatt gaatcctgtt gccgggtcttg cgatgattat catataattt 10320
ctgttgaatt acgttaagca tgtaataatt aacatgtaat gcatgacgtt atttatgaga 10380
tggggttttta tgattagagt cccgcaatta tacatttaac acgcgataga aaacaaaata 10440
tagcgcgcaa actaggataa attatcgcg cgggtgtcat ctatgttact agatcggcgc 10500
gccgcaattg aagtttgggc ggccagcatg gccgtatccg caatgtgtta ttaagttgtc 10560
taagcgtcaa tttgtttaca ccacaatata tctgccacc agccagccaa cagctccccg 10620
accggcagct cggcacaaaa tcaccactcg atacaggcag cccatcag 10668

```

<210> 37
 <211> 10673
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Casset biểu hiện của cấu trúc 23123

<220>
 <221> trình tự tăng cường
 <222> (165)..(256)
 <223> eNOS

<220>
 <221> vùng khởi động
 <222> (301)..(2097)
 <223> Ubi

<220>
 <221> gen
 <222> (2114)..(6283)
 <223> Cas9

<220>
 <221> đặc điểm hỗn tạp
 <222> (6218)..(6274)
 <223> NLS-TAG

<220>
 <221> trình tự kết thúc
 <222> (6289)..(6541)
 <223> tNOS

<220>
 <221> vùng khởi động
 <222> (6548)..(6922)
 <223> pU3

<220>
 <221> đặc điểm hỗn tạp
 <222> (6924)..(6943)

<220>
 <221> đặc điểm hỗn tạp
 <222> (6924)..(7028)
 <223> sgRNA

<220>
 <221> vùng khởi động
 <222> (7039)..(9030)
 <223> Ubi

<220>
 <221> gen
 <222> (9047)..(10225)
 <223> PMI

<220>
 <221> trình tự kết thúc
 <222> (10248)..(10500)
 <223> tNOS

<400> 37
 cctgtggttg gcatgcacat acaaattggac gaacggataa accttttcac gcccttttaa 60
 atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagtttaaac tgaaggcggg aaacgacaat ctgatcatga gcggagaatt 180
 aaggagtgca cgttatgacc cccgccgatg acgcgggaca agccgtttta cgtttggaac 240
 tgacagaacc gcaacgctgc aggaattggc cgcagcggcc atttaaaca agcttggtac 300
 cattatgttg tctaggtagg ttctatatat aagaaaactt gaaatgttct aaaaaaaaaat 360
 tcaagcccat gcatgattga agcaaacggc atagcaacgg tgttaacctg atctagtgat 420
 ctcttgcaat ccttaacggc cacctaccgc aggtagcaaa cggcgctccc ctctcgata 480
 tctccgcggc gacctctggc tttttccgcg gaattgcgcg gtggggacgg attccacgag 540
 accgcgacgc aaccgcctct cgcgcgtggg cccacacccg ctcggtgccg tagcctcacg 600
 ggactctttc tccctcctcc cccgttataa attggcttca tcccctcctt gcctcatcca 660
 tccaaatccc agtcccacat cccatccctt cgtaggagaa attcatcgaa gctaagcgaa 720
 tcctcgcgat cctctcaagg tactgcgagt ttctgatccc cctctcgacc cctcgtatgt 780
 ttgtgtttgt cgtagcgttt gattaggtat gctttccctg tttgtgttcg tcgtagcgtt 840
 tgattaggta tgctttccct gtctgtgttc atcgtagtgt ttgattaggt cgtgtgaggc 900
 gatggcctgc tcgcgtcctt cgatctgtag tcgatttgcg ggtcgtggtg tagatctgcg 960
 ggctgtgatg aagttatatt gtgtgatctg ctgcctgat tctgcgggtt ggctcgagta 1020
 gatatgatgg ttggaccggc ttggtcgttt accgcgctag ggttgggctg ggatgatgtt 1080
 gcatgcgccg ttgcgcgtga tcccgcagca ggacttgctt ttgattgcca gatctcgta 1140
 cgattatgtg atttggtttg gacttttttag atctgtagct tctgcttatg tgccagatgc 1200
 gctactgct catatgcctg atgataatca taaatggctg tggaactaac tagttgattg 1260
 cggagtcatt tatcagctac aggtgtaggg actagctaca ggtgtagggg cttgcgtcta 1320

attgtttggt cctttactca tgttgcaatt atgcaattta gtttagattg tttgttccac 1380
 tcatctaggc tgtaaaaggg acaactgctta gattgctggt taatcttttt agtagattat 1440
 attatatggg taacttatta cccctattac atgccatacg tgacttctgc tcatgcctga 1500
 tgataatcat agatcactgt ggaattaatt agttgattgt tgaatcatgt ttcattgtaca 1560
 taccacggca caattgctta gttccttaac aaatgcaaat tttactgac catgtatgat 1620
 ttgcgtgggt ctctaattgt aaatactata gctacttgtt agtaagaatc aggttcgtat 1680
 gcttaatgct gtatgtgcct tctgctcatg cctgatgata atcatatatc actggaatta 1740
 attagttagt cgtttaatca tatatcaagt acataccatg ccacaatttt tagtcactta 1800
 acccatgcag attgaactgg tccctgcatg ttttgctaaa ttgttctatt ctgattagac 1860
 catatatcat gtattttttt ttggtaatgg ttctcttatt ttaaagtcta tatagttctg 1920
 gtacttggtta gaaagatctg cttcatagtt tagttgccta tccctcgaat taggatgctg 1980
 agcagctgat cctatagctt tgtttcatgt atcaattctt ttgtgttcaa cagtcagttt 2040
 ttgttagatt cattgtaact tatggctgct tactcttctg gtcctcaatg cttgcaggat 2100
 cgcgccgct catatggaca agaagtacag catcgccctg gacatcggca ccaacagcgt 2160
 gggctggggc gtgatcaccg acgagtacaa ggtgccgagc aagaagttca aggtgctggg 2220
 caacaccgac aggcacagca tcaagaagaa cctgatcggc gccctgctgt tcgacagcgg 2280
 cgagaccgcc gaggccacca ggctgaagag gaccgccagg aggaggtaca ccaggaggaa 2340
 gaacaggatc tgctacctgc aggagatctt cagcaacgag atggccaagg tggacgacag 2400
 cttcttccac aggctggagg agagcttctt ggtggaggag gacaagaagc acgagaggca 2460
 cccgatcttc ggcaacatcg tggacgaggt ggcctaccac gagaagtacc cgaccatcta 2520
 ccacctgagg aagaagctgg tggacagcac cgacaaggcc gacctgaggc tgatctacct 2580
 ggccctggcc cacatgatca agttcagggg ccacttcttg atcgagggcg acctgaacct 2640
 ggacaacagc gacgtggaca agctgttcat ccagctggtg cagacctaca accagctgtt 2700
 cgaggagaac ccgatcaacg ccagcggcgt ggacgccaaag gccatcctga gcgccaggct 2760
 gagcaagagc aggaggctgg agaacctgat cgcctcagct cggggcgaga agaagaacgg 2820
 cctgttcggc aacctgatcg ccctgagcct gggcctgacc ccgaacttca agagcaactt 2880
 cgacctggcc gaggacgcca agctgcagct gagcaaggac acctacgac acgacctgga 2940
 caacctgctg gcccgatcg ggcaccagta cgcgacctg ttcttgccg ccaagaacct 3000
 gagcgacgcc atctgctga gcgacatctt gaggggtgaac accgagatca ccaaggcccc 3060
 gctgagcgcc agcatgatca agaggtaaga cgagcaccac caggacctga ccctgctgaa 3120
 ggccctggtg aggcagcagc tgccggagaa gtacaaggag atcttcttct accagagcaa 3180
 gaacggctac gcgggtaca tcgacggcgg cgccagccag gaggagttct acaagttcat 3240

caagccgatac ctggagaaga tggacggcac cgaggagctg ctggtgaagc tgaacaggga	3300
ggacctgctg aggaagcaga ggaccttcga caacggcagc atcccgacc agatccacct	3360
gggcgagctg cacgccatcc tgaggaggca ggaggacttc taccggttcc tgaaggacaa	3420
cagggagaag atcgagaaga tcctgacctt ccgcatcccg tactacgtgg gcccgctggc	3480
caggggcaac agcaggttcg cctggatgac caggaagagc gaggagacca tcaccccgctg	3540
gaacttcgag gaggtggtgg acaagggcgc cagcgcccag agcttcacg agaggatgac	3600
caacttcgac aagaacctgc cgaacgagaa ggtgctgccg aagcacagcc tgctgtacga	3660
gtacttcacc gtgtacaacg agctgaccaa ggtgaagtac gtgaccgagg gcatgaggaa	3720
gccggccttc ctgagcggcg agcagaagaa ggccatcgtg gacctgctgt tcaagaccaa	3780
caggaagggtg accgtgaagc agctgaagga ggactacttc aagaagatcg agtgcttcga	3840
cagcgtggag atcagcggcg tggaggacag gttcaacgcc agcctgggca cctaccacga	3900
cctgctgaag atcatcaagg acaaggactt cctggacaac gaggagaacg aggacatcct	3960
ggaggacatc gtgctgacct tgacctgtt cgaggacagg gagatgatcg aggagaggct	4020
gaagacctac gccacactgt tcgacgacaa ggtgatgaag cagctgaaga ggaggaggta	4080
caccggctgg ggcaggctga gcaggaagct gatcaacggc atcagggaca agcagagcgg	4140
caagaccatc ctggacttcc tgaagagcga cggcttcgcc aacaggaact tcatgcagct	4200
gatccacgac gacagcctga ccttcaagga ggacatccag aaggcccagg tgagcggcca	4260
gggcgacagc ctgcacgagc acatcgccaa cctggccggc agcccggcca tcaagaaggg	4320
catcctgcag accgtgaagg tggaggacga gctggtgaag gtgatgggca ggcacaagcc	4380
ggagaacatc gtgatcgaga tggccaggga gaaccagacc acccagaagg gccagaagaa	4440
cagcagggag aggatgaaga ggatcgagga gggcatcaag gagctgggca gccagatcct	4500
gaaggagcac ccggtggaga acaccagct gcagaacgag aagctgtacc tgtactacct	4560
gcagaacggc agggacatgt acgtggacca ggagctggac atcaacaggc tgagcgacta	4620
cgacgtggac cacatcgtgc cgcagagctt cctgaaggac gacagcatcg acaacaaggt	4680
gctgaccagg agcgacaaga acaggggcaa gagcgacaac gtgccgagcg aggaggtggt	4740
gaagaagatg aaaaactact ggaggcagct gctgaacgcc aagctgatca cccagaggaa	4800
gttcgacaac ctgaaccaagg ccgagagggg cggcctgagc gagctggaca aggccggctt	4860
cattaaaagg cagctggtgg agaccaggca gatcaccaag cacgtggccc agatcctgga	4920
cagcaggatg aacaccaagt acgacgagaa cgacaagctg atcagggagg tgaaggtgat	4980
caccctgaag agcaagctgg tgagcgactt caggaaggac ttccagttct acaaggtgag	5040
ggagatcaat aattaccacc acgcccacga cgcctacctg aacgccgtgg tgggcaccgc	5100
cctgattaaa aagtaccgga agctggagag cgagttcgtg tacggcgact acaaggtgta	5160

cgacgtgagg aagatgatcg ccaagagcga gcaggagatc ggcaaggcca ccgccaagta 5220
 cttctttctac agcaacatca tgaacttctt caagaccgag atcacccctgg ccaacggcga 5280
 gatcaggaag aggccgctga tcgagaccaa cggcgagacc ggcgagatcg tgtgggacaa 5340
 gggcagggac ttcgccaccg tgaggaaggt gctgtccatg ccgcaggtga acatcgtgaa 5400
 gaagaccgag gtgcagaccg gcggcttcag caaggagagc atcctgccga agaggaacag 5460
 cgacaagctg atcgccagga agaaggactg ggacccgaag aagtacggcg gcttcgacag 5520
 cccgaccgtg gcctacagcg tgctggtggt ggccaaggtg gagaagggca agagcaagaa 5580
 gctgaagagc gtgaaggagc tgggtggcat caccatcatg gagaggagca gcttcgagaa 5640
 gaacccagtg gacttcctgg aggccaaagg ctacaaggag gtgaagaagg acctgatcat 5700
 taaactgccg aagtacagcc tgttcgagct ggagaacggc aggaagagga tgctggccag 5760
 cgccggcgag ctgcagaagg gcaacgagct ggccctgccg agcaagtacg tgaacttcct 5820
 gtacctggcc agccactacg agaagctgaa gggcagcccg gaggacaacg agcagaagca 5880
 gctgttcgtg gagcagcaca agcactacct ggacgagatc atcgagcaga tcagcgagtt 5940
 cagcaagagg gtgatcctgg ccgacgcaa cctggacaag gtgctgagcg cctacaacaa 6000
 gcacagggac aagccgatca gggagcaggc cgagaacatc atccacctgt tcacctgac 6060
 caacctgggc gccccggccg ccttcaagta cttcgacacc accatcgaca ggaagaggta 6120
 caccagcacc aaggaggtgc tggacgccac cctgatccac cagagcatca ccggcctgta 6180
 cgagaccagg atcgacctga gccagctggg cggcgacagc agcccgccga agaagaagag 6240
 gaagggtgagc tggaaggacg ccagcggctg gagcaggatg tgaagcttga tcgttcaaac 6300
 atttggcaat aaagtttctt aagattgaat cctgttgccg gtcttgcgat gattatcata 6360
 taatttctgt tgaattacgt taagcatgta ataattaaca tgtaatgcat gacgttattt 6420
 atgagatggg tttttatgat tagagtcccg caattataca ttttaatacgc gatagaaaac 6480
 aaaatatagc gcgcaaacta ggataaatta tcgcgcgcgg tgatcatctat gttactagat 6540
 cttcgaaggg atcttttaaac atacgaacag atcacttaaa gttcttctga agcaacttaa 6600
 agttatcagg catgcatgga tcttgaggga atcagatgtg cagtcaggga ccatagcaca 6660
 ggacaggcgt cttctactgg tgctaccagc aaatgctgga agccgggaac actgggtacg 6720
 ttggaaacca cgtgatgtgg agtaagataa actgtaggag aaaagcattt cgtagtgggc 6780
 catgaagcct ttcaggacat gtattgcagt atgggcccgc ccattacgca attggacgac 6840
 aacaaagact agtattagta ccacctcggc tatccacata gatcaaagct ggtttaaaag 6900
 agttgtgcag atgatccgtg gcagggtcaa cgtggagaca ggggttttag agctagaaat 6960
 agcaagttaa aataaggcta gtccgttatc aacttgaaaa agtggcaccc agtcggtgct 7020
 ttttttttcg gaccgcgcct gcagtgcagc gtgaccgggt cgtgcccctc tctagagata 7080

atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt cacacttggt 7140
 tgaagtgcag tttatctatc tttatacata tttttaaaact ttactctacg aataatataa 7200
 totatagtac tacaataata tcagtgtttt agagaatcat ataaatgaac agttagacat 7260
 ggtctaaagg acaattgagt attttgacaa caggactota cagtttttatc ttttttagtgt 7320
 gcatgtgttc tccttttttt ttgcaaataag cttcacctat ataatacttc atccatttta 7380
 ttagtacatc catttaggggt ttaggggttaa tggtttttat agactaattt ttttagtaca 7440
 totattttat totatttttag cctctaaaatt aagaaaaacta aaactctatt ttagtttttt 7500
 tatttaataa ttttagatata aaatagaata aaataaagtg actaaaaatt aaacaaatac 7560
 cctttaagaa attaaaaaaaa ctaaggaaac atttttcttg tttcgagtag ataatgccag 7620
 cctgttaaac gccgtcgacg agtctaacgg acaccaacca gcgaaccagc agcgtcgcg 7680
 cgggccaagc gaagcagacg gcacggcatc tctgtcgctg cctctggacc cctctcgaga 7740
 gttccgctcc accgttgac ttgctccgct gtcggcatcc agaaattgcg tggcggagcg 7800
 gcagacgtga gccggcacgg caggcggcct cctcctcctc tcacggcacc ggcagctacg 7860
 ggggattcct tccccaccgc tccttcgctt tcccttcctc gcccgccgta ataaatagac 7920
 accccctcca caccctcttt ccccaacctc gtgttggtcg gagcgcacac acacacaacc 7980
 agatctcccc caaatccacc cgtcggcacc tccgcttcaa ggtacgccgc tcgtcctccc 8040
 cccccccct ctctaccttc tctagatcgg cgttcgggtc catgggttagg gcccggtagt 8100
 tctacttctg ttcattgttg tgttagatcc gtgttttgt tagatccgtg ctgctagcgt 8160
 tcgtacacgg atgcgacctg tacgtcagac acgttctgat tgctaacttg ccagtgtttc 8220
 tctttgggga atcctgggat ggctctagcc gttccgcaga cgggatcgat ttcattgattt 8280
 tttttgtttc gttgcatagg gtttggtttg cccctttcct ttatttcaat atatgccgtg 8340
 cacttgtttg tcgggtcatc ttttcatgct tttttttgtc ttggttgtga tgatgtggtc 8400
 tggttgggcg gtcgttctag atcggagtag aattctgttt caaactacct ggtggattta 8460
 ttaatttttg atctgtatgt gtgtgccata catattcata gttacgaatt gaagatgatg 8520
 gatggaaata tcgatctagg ataggtatac atgttgatgc gggttttact gatgcatata 8580
 cagagatgct ttttgttcgc ttggttgtga tgatgtgggtg tggttgggcg gtcgttcatt 8640
 cgttctagat cggagtagaa tactgtttca aactacctgg tgtatttatt aattttggaa 8700
 ctgtatgtgt gtgtcataca ttttcatagt tacgagttta agatggatgg aaatatcgat 8760
 ctaggatagg tatacatggt gatgtgggtt ttactgatgc atatacatga tggcatatgc 8820
 agcatctatt catatgctct aaccttgagt acctatctat tataataaac aagtatgttt 8880
 tataattatt ttgatcttga tataacttga tgatggcata tgcagcagct atatgtggat 8940
 ttttttagcc ctgccttcat acgctattta tttgcttggt actgtttcct ttgtcgatgc 9000

tcaccctgtt gtttggtgtt acttctgcag ggatccggca gcagccatgc agaagctgat 9060
 caacagcgtg cagaactacg cctggggcag caagaccgcc ctgaccgagc tgtacggcat 9120
 ggagaacccc agcagccagc ccatggccga gctgtggatg ggcgcccacc ccaagagcag 9180
 cagccgctg cagaacgccg ccggcgacat cgtgagcctg cgcgacgtga tcgagagcga 9240
 caagagcacc ctgctgggcg aggccgtggc caagcgcttc ggcgagctgc ctttctgtt 9300
 caaggtgctg tgcgccgcc agcccctgag catccagggtg caccccaaca agcacaacag 9360
 cgagatcggc ttcgccaaag agaacgccgc cggcatcccc atggacgccg ccgagcgcaa 9420
 ctacaaggac cccaaccaca agcccagct ggtgttcgcc ctgacccct tctggccat 9480
 gaacgccttc cgcgagttca gcgagatcgt gagcctgctg cagcccgtag ccggcgccca 9540
 ccccgccatc gccacttcc tgcagcagc cgacgccgag cgctgagcg agctgttcgc 9600
 cagcctgctg aacatgcagg gcgaggagaa gagccgcgc ctggccatcc tgaagagcgc 9660
 cctggacagc cagcagggcg agccctggca gaccatccgc ctgatcagcg agttctaccc 9720
 cgaggacagc ggctgttca gcccctgct gctgaacgtg gtgaagctga acccggcgga 9780
 ggccatgttc ctgttcgccg agacccccca cgcctacctg cagggcgtag ccctggaggt 9840
 gatggccaac agcgacaacg tgctgcgcgc cggcctgacc cccaagtaca tcgacatccc 9900
 cgagctggtg gccaacgtga agttcgaggc caagcccgc aaccagctgc tgaccagcc 9960
 cgtgaagcag ggcgccgagc tggacttccc catcccctg gacgacttcg cttcagcct 10020
 gcacgacctg agcgacaagg agaccaccat cagccagcag agcgccgcca tctgttctg 10080
 cgtggagggc gacgccaccc tgtggaaggc cagccagcag ctgcagctga agccggcgga 10140
 gagcgcttc atcgccgcca acgagagccc cgtgaccgtg aagggccacg gccgcctggc 10200
 ccgcgtgtac aacaagctgt gataggagct cgatccgtcg aactgcagat cgttcaaaca 10260
 tttggcaata aagtttctta agattgaatc ctgttgccgg tottgcgatg attatcatat 10320
 aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttatatta 10380
 tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca 10440
 aaatatagcg cgcaaactag gataaattat cgcgcgcggg gtcactctatg ttactagatc 10500
 ggcgcgcgcg aattgaagtt tgggcggcca gcatggccgt atccgcaatg tgttattaag 10560
 ttgtctaagc gtcaatttgt ttacaccaca atatatcctg ccaccagcca gccaacagct 10620
 ccccgaccgg cagctcgga caaatcacc actcgataca ggcagcccat cag 10673

<210> 38

<211> 10693

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 23501, gARN đích 1 dẫn kép

<220>

<221> vùng khởi động

<222> (214)..(734)

<223> 35S

<220>

<221> gen

<222> (748)..(4917)

<223> Cas9

<220>

<221> đặc điểm hỗn tạp

<222> (4852)..(4908)

<223> NLS-TAG

<220>

<221> trình tự kết thúc

<222> (4923)..(5175)

<223> tNOS

<220>

<221> vùng khởi động

<222> (5196)..(5570)

<223> pU3

<220>

<221> đặc điểm hỗn tạp

<222> (5572)..(5591)

<220>

<221> đặc điểm hỗn tạp

<222> (5592)..(5636)

<223> crRNA

<220>

<221> vùng khởi động

<222> (5643)..(6158)

<223> pU6

<220>

<221> đặc điểm hỗn tạp

<222> (6159)..(6246)

<223> tracrRNA

<220>

<221> vùng khởi động

<222> (6253)..(8245)

<223> Ubi

<220>

<221> gen

<222> (8257)..(9435)

<223> PMI

<220>

<221> trình tự kết thúc

<222> (9440)..(10474)

<223> Ubi

<400> 38

cctgtggttg gcatgcacat acaaattggac gaacggataa acctttttcac gccctttttaa

60

atatccgatt attctaataa acgctctttt ctotttaggtt taccgcgcaa tatatcctgt 120
 caaacactga tagtttaaac tggcactagc ctaacgggtgt tgactaacta ggccgcttcc 180
 ctaattagct aaccgcgggg cgccgcggga ccgagtcaaa gattcaaata gaggacctaa 240
 cagaactcgc cgtaaagact ggcgaacagt tcatacagag tctcttacga ctcaatgaca 300
 agaagaaaat cttcgtcaac ttggtggagc acgacacgct agtctactcc aaaaatatca 360
 aagatacagt ctcagaagac caaagggcaa ttgagacttt tcaacaaagg gtaatatccg 420
 gaaacctcct cggattccat tgcccagcta tctgtcactt aattgtgaag atagtggaaa 480
 aggaaggtgg ctctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg 540
 cctctgccga cagtgggtccc aaagatggac cccacccac gaggagcatc gtggtaaaag 600
 aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa 660
 gggatgacgc acaatcccac tatccttcgc aagacccttc ctctatataa ggaagttcat 720
 ttcatttggga gaggataatt atccaccatg gacaagaagt acagcatcgg cctggacatc 780
 ggcaccaaca gcgtgggctg ggccgtgatc accgacgagt acaagggtgcc gagcaagaag 840
 ttcaagggtgc tgggcaacac cgacaggcac agcatcaaga agaacctgat cggcgccctg 900
 ctgttcgaca gcggcgagac cgccgaggcc accaggctga agaggaccgc caggaggagg 960
 tacaccagga ggaagaacag gatctgctac ctgcaggaga tcttcagcaa cgagatggcc 1020
 aagggtggacg acagcttctt ccacaggctg gaggagagct tcctgggtgga ggaggacaag 1080
 aagcacgaga ggcacccgat cttcggcaac atcgtggacg aggtggccta ccacgagaag 1140
 taccgacca tctaccacct gaggaagaag ctggtggaca gcaccgaaa ggccgacctg 1200
 aggtgatct acctggccct ggccacatg atcaagttca ggggccactt cctgatcgag 1260
 ggcgacctga acccggacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc 1320
 tacaaccagc tgttcgagga gaaccgatc aacgccagcg gcgtggacgc caaggccatc 1380
 ctgagcgcca ggctgagcaa gagcaggagg ctggagaacc tgatcgcca gctgccgggc 1440
 gagaagaaga acggcctgtt cggcaacctg atcgccctga gcctgggcct gaccccgaa 1500
 ttcaagagca acttcgacct ggccgaggac gccaaagtgc agctgagcaa ggacacctac 1560
 gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgttcttg 1620
 gccgccaaaga acctgagcga cgccatcctg ctgagcgaca tcctgagggg gaacaccgag 1680
 atcaccaagg ccccgctgag cgccagcatg atcaagaggt acgacgagca ccaccaggac 1740
 ctgacctgc tgaaggccct ggtgaggcag cagctgccgg agaagtacaa ggagatcttc 1800
 ttcgaccaga gcaagaacgg ctacgccggc tacatcgacg gcggcgccag ccaggaggag 1860
 ttctacaagt tcatcaagcc gatcctggag aagatggacg gcaccgagga gctgctggtg 1920
 aagctgaaca gggaggacct gctgaggaag cagaggacct tcgacaacgg cagcatcccg 1980

caccagatcc acctgggcca gctgcacgcc atcctgagga ggcaggagga cttctacccg	2040
ttcctgaagg acaacaggga gaagatcgag aagatcctga ccttccgcat cccgtactac	2100
gtggggcccg tggccaggga caacagcagg ttcgcctgga tgaccaggaa gagcgaggag	2160
accatcacc cgtggaactt cgaggagggtg gtggacaagg gcgccagcgc ccagagcttc	2220
atcgagagga tgaccaactt cgacaagaac ctgccgaacg agaagggtgct gccgaagcac	2280
agcctgctgt acgagtactt caccgtgtac aacgagctga ccaagggtgaa gtacgtgacc	2340
gagggcatga ggaagccggc cttcctgagc ggcgagcaga agaaggccat cgtggacctg	2400
ctgttcaaga ccaacaggaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag	2460
atcgagtgtc tcgacagcgt ggagatcagc ggcgtggagg acagggttaa cgccagcctg	2520
ggcacctacc acgacctgct gaagatcatc aaggacaagg acttcctgga caacgaggag	2580
aacgaggaca tcctggagga catcgtgctg accctgacct tgttcgagga caggagatg	2640
atcgaggaga ggctgaagac ctacgccac ctgttcgacg acaagggtgat gaagcagctg	2700
aagaggagga ggtacaccgg ctggggcagg ctgagcagga agctgatcaa cggcatcagg	2760
gacaagcaga gcggcaagac catcctggac ttctgaaga gcgacggctt cgccaacagg	2820
aacttcatgc agctgatcca cgacgacagc ctgaccttca aggaggacat ccagaaggcc	2880
caggtgagcg gccagggcga cagcctgcac gagcacatcg ccaacctggc cggcagcccg	2940
gcatcaaga agggcatcct gcagaccgtg aagggtggtg acgagctggt gaagggtgatg	3000
ggcaggcaca agccggagaa catcgtgatc gagatggcca gggagaacca gaccacccag	3060
aagggccaga agaacagcag ggagaggatg aagaggatcg aggagggcat caaggagctg	3120
ggcagccaga tcctgaagga gcacccggtg gagaacacct agctgcagaa cgagaagctg	3180
tacctgtact acctgcagaa cggcaggga atgtacgtgg accaggagct ggacatcaac	3240
aggctgagcg actacgacgt ggaccacatc gtgccgcaga gcttctgaa ggacgacagc	3300
atcgacaaca aggtgctgac caggagcagc aagaacaggg gcaagagcga caacgtgccg	3360
agcgaggagg tggatgaaga gatgaaaaac tactggaggc agctgctgaa cgccaagctg	3420
atcaccaga ggaagttcga caacctgacc aaggccgaga ggggcggcct gagcgagctg	3480
gacaaggccg gcttcattaa aaggcagctg gtggagacca ggcagatcac caagcacgtg	3540
gccagatcc tggacagcag gatgaacacc aagtagcagc agaacgaca gctgatcagg	3600
gaggtgaagg tgatcacctt gaagagcaag ctggtgagcg acttcaggaa ggacttccag	3660
ttctacaagg tgaggagat caataattac caccacgccc acgacgccta cctgaacgcc	3720
gtggtgggca ccgcctgat taaaaagtac ccgaagctgg agagcgagtt cgtgtacggc	3780
gactacaagg tgtacgacgt gaggaagatg atcgccaaga gcgagcagga gatcggaag	3840
gccaccgcca agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc	3900

ctggccaacg gcgagatcag gaagaggccg ctgatcgaga ccaacggcga gaccggcgag 3960
 atcgtgtggg acaagggcag ggacttcgcc accgtgagga aggtgctgtc catgccgcag 4020
 gtgaacatcg tgaagaagac cgaggtgcag accggcggtt tcagcaagga gagcatcctg 4080
 ccgaagagga acagcgacaa gctgatcgcc aggaagaagg actgggatcc gaagaagtac 4140
 ggcggcttcg acagcccgcg cgtggcctac agcgtgctgg tgggtggcaa ggtggagaag 4200
 ggcaagagca agaagctgaa gagcgtgaag gagctggtgg gcatcaccat catggagagg 4260
 agcagcttcg agaagaaccc agtggacttc ctggaggcca agggctacaa ggaggtgaag 4320
 aaggacctga tcattaaact gccgaagtac agcctgttcg agctggagaa cggcaggaag 4380
 aggatgctgg ccagcgccgg cgagctgcag aagggaacg agctggccct gccgagcaag 4440
 tacgtgaact tcctgtacct ggccagccac tacgagaagc tgaagggcag cccggaggac 4500
 aacgagcaga agcagctggt cgtggagcag cacaagcact acctggacga gatcatcgag 4560
 cagatcagcg agttcagcaa gaggggtgatc ctggccgacg ccaacctgga caaggtgctg 4620
 agcgcctaca acaagcacag ggacaagccg atcagggagc aggccgagaa catcatccac 4680
 ctgttcaccc tgaccaacct gggcgccccg gccgccttca agtacttcga caccaccatc 4740
 gacaggaaga ggtacaccag caccaaggag gtgctggacg ccacctgat ccaccagagc 4800
 atcacgggcc tgtacgagac caggatcgac ctgagccagc tgggcggcga cagcagcccg 4860
 ccgaagaaga agaggaaggt gagctggaag gacgccagcg gctggagcag gatgtgaagc 4920
 ttgatcgttc aaacatttgg caataaagtt tottaagatt gaatcctggt gccggtcttg 4980
 cgatgattat catataatct ctgttgaatt acgttaagca tgtaataatt aacatgtaat 5040
 gcatgacggt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 5100
 acgcataga aaacaaaata tagcgcgcaa actaggataa attatcgcg cgggtgtcat 5160
 ctatgttact agatcgggac cggaattcgg gaccgggat ctttaaacat acgaacagat 5220
 cacttaaagt tcttctgaag caacttaaag ttatcaggca tgcattggatc ttggaggaat 5280
 cagatgtgca gtcagggacc atagcacagg acaggcgtct totactgggtg ctaccagcaa 5340
 atgctggaag ccgggaacac tgggtacggt ggaaaccacg tgatgtggag taagataaac 5400
 tgtaggagaa aagcatttcg tagtgggcca tgaagccttt caggacatgt attgcagtat 5460
 gggccggccc attacgcaat tggacgacaa caaagactag tattagtacc acctcggcta 5520
 tccacataga tcaaagctgg tttaaaagag ttgtgcagat gatccgtggc atgcagtcga 5580
 agtaatcggc ggttttagag ctatgctggt ttgaatggtc caaaaacttt ttttttacta 5640
 gttttgtgaa agttgaatta cggcatagcc gaaggaataa cagaatcgtt tcacactttc 5700
 gtaacaaagg tcttcttatc atgtttcaga cgatggaggc aaggctgatc aaagtgatca 5760
 agcacataaa cgcatttttt taccatgttt cactccataa gcgtctgaga ttatcacaag 5820

tcacgtctag tagtttgatg gtacactagt gacaatcagt tcgtgcagac agagctcata	5880
cttgactact tgagcgatta caggcgaaaag tgtgaaacgc atgtgatgtg ggctgggagg	5940
aggagaatat atactaatgg gccgtatcct gatttgggct gcgtcggaag gtgcagccca	6000
cgcgcgcgt accgcgcggg tggcgcgtgct acccacttta gtccgttga tggggatccg	6060
atggtttgcg cgggtggcgtt gcgggggatg tttagtagca catcggaac cgaaagacga	6120
tggaaccagc ttataaacc gcgcgcgtga gtcagcttg aaccattcaa aacagcatag	6180
caagttaaaa taaggctagt ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt	6240
ttttttccta ggctgcagtg cagcgtgacc cggtcgtgcc cctctctaga gataatgagc	6300
attgcatgtc taagttataa aaaattacca catatTTTTT ttgtcacact tgtttgaagt	6360
gcagtttatc tatctttata catatatTTT aactttactc tacgaataat ataatctata	6420
gtactacaat aatatcagt ttttagagaa tcatataaat gaacagttag acatggctca	6480
aaggacaatt gagtatTTTT acaacaggac tctacagttt tatctTTTT gtgtgcatgt	6540
gttctccttt ttttttgcaa atagcttcac ctatataata ctccatccat tttattagta	6600
catccattta gggtttaggg ttaatggttt ttatagacta atTTTTTTtag tacatctatt	6660
ttattctatt ttagcctcta aattaagaaa actaaaactc tatttttagtt tttttattta	6720
ataatttaga tataaaatag aataaaataa agtgactaaa aattaaacaa atacccttta	6780
agaaattaaa aaaactaagg aaacattttt cttgtttcga gtagataatg ccagcctgtt	6840
aaacgccgtc gacgagtcta acggacacca accagcgaac cagcagcgtc gcgtcggggc	6900
aagcgaagca gacggcacgg catctctgtc gctgcctctg gacccctctc gagagtcccg	6960
ctccaccgtt ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg agcggcagac	7020
gtgagccggc acggcaggcg gcctcctcct cctctcacgg caccggcagc tacgggggat	7080
tcctttccca ccgtccttc gctttccctt cctcgccgc cgtaataaat agacaccccc	7140
tccacaccct ctttcccaa cctcgtgttg ttoggagcgc acacacacac aaccagatct	7200
ccccaaatc caccgctgg caoctcgt tcaaggtaag ccgtcgtcc tcccccccc	7260
cccctctcta ccttctctag atcggcgctt cggccatgg ttagggccc gtagttctac	7320
ttctgttcat gtttgtgtta gatccgtgtt tgtgttagat ccgtgctgct agcgttcgta	7380
cacggatgcg acctgtacgt cagacacgct ctgattgcta acttgccagt gtttctcttt	7440
ggggaatcct gggatggctc tagccgttcc gcagacggga tcgatttcat gatttttttt	7500
gtttcgttgc atagggtttg gtttgccctt ttcctttatt tcaatatatg ccgtgcactt	7560
gtttgtcggg tcatcttttc atgctttttt ttgtottggt tgtgatgatg tggctctggt	7620
gggcggctgt tctagatcgg agtagaattc tgtttcaaac tacctggtgg atttattaat	7680
tttgatctg tatgtgtgtg ccatacatat tcatagttac gaattgaaga tgatggatgg	7740

aaatatcgat ctaggatagg tatacatggt gatgcggggt ttactgatgc atatacagag	7800
atgctttttg ttcgcttggt tgtgatgatg tgggtgtggt gggcggtcgt tcattcgttc	7860
tagatcggag tagaatactg tttcaaaacta cctgggtgat ttattaattt tggaactgta	7920
tgtgtgtgtc atacatcttc atagttacga gtttaagatg gatggaaata tcgatctagg	7980
ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcagcat	8040
ctattcatat gctctaacct tgagtaccta tctattataa taaacaagta tgttttataa	8100
ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg tggatttttt	8160
tagccctgcc ttcatacgtt atttatttgc ttggtactgt ttcttttgtc gatgctcacc	8220
ctgttgtttg gtgttacttc tgcagtgact aaatagatgc agaagctgat caacagcgtg	8280
cagaactacg cctggggcag caagaccgcc ctgaccgagc tgtacggcat ggagaacccc	8340
agcagccagc ccatggccga gctgtggatg ggcgccacc ccaagagctc aagccgcgtg	8400
cagaacgccg ccggcgatat cgttagcctg cgcgacgtga tcgagagcga caagagcacc	8460
ctgctgggag aggccgtggc caagcgcttc ggcgagctgc ccttcctgtt caaggtgctg	8520
tgcgcccgtc agcccctgag catccaggtg caccctaaca agcacaacag cgagatcggc	8580
ttcgccaagg agaacgccgc cggcatcccc atggacgccg ccgagcgcaa ctacaaggac	8640
cccaaccaca agcccagct ggtgttcgcc ctgacccctt tcttgccat gaacgccttc	8700
cgcgagttca gcgagatcgt tagcctgctg cagcccgtgg ccggcgccca cccgctatc	8760
gccacttcc ttcagcagcc cgacgccgag cgcctgagcg agctgttcgc cagcctgctg	8820
aacatgcagg gtgaggagaa gtcacgcgcc ctggccatcc tgaagagcgc cctggacagc	8880
cagcagggcg agccctggca gacaatccgc ctgatcagcg agttctaccc cgaggatagc	8940
ggcctgttca gcccctgct gctgaacgtg gtgaagctga accccggcga ggccatgttc	9000
ctgttcgccg agacccccca cgcctacctg cagggcggtg ccctggaggt gatggccaac	9060
agcgacaacg tgctgcgcgc cggcctgacc cccaagtaca tcgacatccc cgagctggtg	9120
gccaacgtga agttcgaggc taagcccgcc aaccagctgc tgaccagcc cgtgaagcag	9180
ggcgccgagc tggacttccc tatccccgtt gacgacttcg ccttcagcct gcacgacctg	9240
agcgacaagg agaccactat cagccagcag agcgccgcga tctgtttctg cgtggagggc	9300
gacgccaccc tgtggaaggg cagccagcag ctgcagctga agcccggcga gagcgcttt	9360
atcgccgcca acgagagccc cgtgaccgtg aaggggccacg gccgcctggc ccgctgttac	9420
aacaagctgt gatagctacg tcatgggtcg tttaagctgc cgatgtgcct gcgtcgtctg	9480
gtgccctctc tccatatgga ggttgtcaaa gtatctgctg ttcggtgcat gagtcgtgtc	9540
agtgttggtt taataatgga ccggttgtgt tgtgtgtgcg tactaccag aactatgaca	9600
aatcatgaat aagtttgatg ttgaaaatta aagcctgtgc tcattatgtt ctgtctttca	9660

```

gttgtctcct aatatttgcc tgcagggtact ggctatctac cgttttcttac ttaggaggtg 9720
tttgaatgca ctaaaactaa tagttagtgg ctaaaattag ttaaaacatc caaacaccat 9780
agctaatagt tgaactatta gctatTTTTg gaaaattagt taatagttag gtagttatTT 9840
gtagctagc taattcaact aacaattttt agccaactaa caattagttt cagtgcattc 9900
aaacaccccc ttaatgttaa cgtggttcta tctaccgtct cctaatatat ggtagattgt 9960
tcgggtttgtt gctatgctat tgggttctga ttgctgctag ttcttgctga atccagaagt 10020
tctcgtagta tagctcagat tcatattatt tatttgagtg ataagtgatc cagggtatta 10080
ctatgtagc taggtttttt ttacaaggat aaattatctg tgatcataat tcttatgaaa 10140
gctttatgtt tcctggaggc agtggcatgc aatgcatgac agcaacttga tcacaccagc 10200
tgaggtagat acggttaaca gggtcttaaa tctgttcacc aaatcattgg agaacacaca 10260
tacacattct tgccagtctt ggtagagaa atttcatgac aaaatgccaa agctgtcttg 10320
actcttcact tttggccatg agtcgtgact tagtttggtt taatggaccg gttctcctag 10380
cttgttctac tcaaaactgt tgttgatgag aataagttgt gatggttgat ctctggattt 10440
tgttttgctc tcaatagtgg acgagattag atagcggacc gcctgcaggc cggggggcgc 10500
gccctaatta gctaacggcc aggatcgccg cgtgagcctt tagcaactag ctagattaat 10560
taacgcaatc tgttattaag ttgtctaagc gtcaatttgt ttacaccaca atatatcctg 10620
ccaccagcca gccaacagct ccccgaccgg cagctcggca caaatcacc actcgataca 10680
ggcagcccat cag 10693

```

<210> 39
 <211> 10264
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Casset biểu hiện của cấu trúc 23501, gRNA đích 1 dẫn đơn

<220>
 <221> vùng khởi động
 <222> (214)..(734)
 <223> 35S

<220>
 <221> gen
 <222> (748)..(4917)
 <223> Cas9

<220>
 <221> đặc điểm hỗn tạp
 <222> (4852)..(4908)
 <223> NLS-TAG

<220>
 <221> trình tự kết thúc

<222> (4923)..(5175)
 <223> tNOS

<220>
 <221> vùng khởi động
 <222> (5196)..(5711)
 <223> pU6

<220>
 <221> đặc điểm hỗn tạp
 <222> (5712)..(5817)
 <223> sgRNA

<220>
 <221> đặc điểm hỗn tạp
 <222> (5713)..(5732)

<220>
 <221> vùng khởi động
 <222> (5824)..(7816)
 <223> Trình tự khởi động Ubi

<220>
 <221> gen
 <222> (7828)..(9006)
 <223> PMI

<220>
 <221> trình tự kết thúc
 <222> (9011)..(10045)
 <223> Trình tự kết thúc Ubi

<400> 39
 cctgtggttg gcatgcacat acaaattggac gaacggataa acctttttcac gcccttttta 60
 atatccgatt attctaataa acgctctttt ctcttagggtt taccgcgcaa tatatcctgt 120
 caaacactga tagttttaaac tggcactagc ctaacggtgt tgactaacta ggccgcttcc 180
 ctaattagct aaccgggggg cgcgccggga ccgagtcaaa gattcaaata gaggacctaa 240
 cagaactcgc cgtaaagact ggcgaaacagt tcatacagag tctcttacga ctcaatgaca 300
 agaagaaaat ctctgtcaac ttggtggagc acgacacgct agtctactcc aaaaatatca 360
 aagatacagt ctcaagaagc caaagggcaa ttgagacttt tcaacaaagg gtaatatccg 420
 gaaacctcct cggattccat tgcccagcta tctgtcactt aattgtgaag atagtggaaa 480
 aggaagggtg ctcctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg 540
 cctctgccga cagtgggtccc aaagatggac cccacccac gaggagcatc gtggtaaaaag 600
 aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa 660
 gggatgacgc acaatcccac tctccttcgc aagacccttc ctctatataa ggaagttcat 720
 ttcatattgga gaggataatt atccaccatg gacaagaagt acagcatcgg cctggacatc 780
 ggcaccaaca gcgtgggctg ggccgtgatc accgacgagt acaagggtgc gagcaagaag 840
 ttcaagggtgc tgggcaacac cgacaggcac agcatcaaga agaacctgat cggcgccctg 900
 ctgttcgaca gcggcgagac cgccgaggcc accaggctga agaggaccgc caggaggagg 960

tacaccagga ggaagaacag gatctgctac ctgcaggaga tcttcagcaa cgagatggcc	1020
aaggtggacg acagcttctt ccacaggctg gaggagagct tcctgggtgga ggaggacaag	1080
aagcacgaga ggcacccgat cttcggcaac atcgtggacg aggtggccta ccacgagaag	1140
tacccgacca tctaccacct gaggaagaag ctggtggaca gcaccgacaa ggccgacctg	1200
aggctgatct acctggccct ggcccacatg atcaagttca ggggccactt cctgatcgag	1260
ggcgacctga acccggacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc	1320
tacaaccagc tgttcgagga gaacccgatc aacgccagcg gcgtggacgc caaggccatc	1380
ctgagcgcca ggctgagcaa gagcaggagg ctggagaacc tgatcgccca gctgccgggc	1440
gagaagaaga acggcctggt cggcaacctg atcgccctga gcctgggcct gacccgaac	1500
ttcaagagca acttcgacct ggccgaggac gccaaactgc agctgagcaa ggacacctac	1560
gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgttcctg	1620
gccgccaaga acctgagcga cgccatcctg ctgagcgaca tcctgagggt gaacaccgag	1680
atcaccaagg ccccgctgag cgccagcatg atcaagaggt acgacgagca ccaccaggac	1740
ctgacctgac tgaaggccct ggtgaggcag cagctgccgg agaagtacaa ggagatcttc	1800
ttcgaccaga gcaagaacgg ctacgccggc tacatcgacg gcggcgccag ccaggaggag	1860
ttctacaagt tcatcaagcc gatcctggag aagatggacg gcaccgagga gctgctggtg	1920
aagctgaaca gggaggacct gctgaggaag cagaggacct tcgacaacgg cagcatcccc	1980
caccagatcc acctgggcga gctgcacgcc atcctgagga ggcaggagga cttctacccc	2040
ttcctgaagg acaacaggga gaagatcgag aagatcctga ccttcgcgat cccgtactac	2100
gtgggcccgc tggccagggg caacagcagg ttgcctgga tgaccaggaa gacgaggag	2160
accatcacc cgtggaactt cgaggagggt gtggacaagg gcgccagcgc ccagagcttc	2220
atcgagagga tgaccaactt cgacaagaac ctgccgaacg agaaggtgct gccgaagcac	2280
agcctgctgt acgagtactt caccgtgtac aacgagctga ccaaggtgaa gtacgtgacc	2340
gagggcatga ggaagccggc cttcctgagc ggcgagcaga agaaggccat cgtggacctg	2400
ctgttcaaga ccaacaggaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag	2460
atcgagtgct tcgacagcgt ggagatcagc ggctggagg acaggttcaa cgccagcctg	2520
ggcacctacc acgacctgct gaagatcacc aaggacaagg acttcctgga caacgaggag	2580
aacgaggaca tcctggagga catcgtgctg acctgaccc tgttcgagga caggagatg	2640
atcgaggaga ggctgaagac ctacgccac ctgttcgacg acaaggtgat gaagcagctg	2700
aagaggagga ggtacaccgg ctggggcagg ctgagcagga agctgatcaa cgccatcagg	2760
gacaagcaga gcggcaagac catcctggac ttctgaaga gcgacggctt cgccaacagg	2820
aacttcatgc agctgatcca cgacgacagc ctgaccttca aggaggacat ccagaaggcc	2880

caggtgagcg gccagggcga cagcctgcac gagcacatcg ccaacctggc cggcagcccg	2940
gccatcaaga agggcatcct gcagaccgtg aaggtggtgg acgagctggt gaaggtgatg	3000
ggcaggcaca agccggagaa catcgtgatc gagatggcca gggagaacca gaccacccag	3060
aagggccaga agaacagcag ggagaggatg aagaggatcg aggagggcat caaggagctg	3120
ggcagccaga tcctgaagga gcacccggtg gagaacaccc agctgcagaa cgagaagctg	3180
tacctgtact acctgcagaa cggcagggac atgtacgtgg accaggagct ggacatcaac	3240
aggctgagcg actacgacgt ggaccacatc gtgccgcaga gcttcctgaa ggacgacagc	3300
atcgacaaca aggtgctgac caggagcgac aagaacaggg gcaagagcga caacgtgccg	3360
agcgaggagg tggatgaagaa gatgaaaaac tactggaggc agctgctgaa cgccaagctg	3420
atcaccacaga ggaagttcga caacctgacc aaggccgaga ggggcggcct gagcgagctg	3480
gacaaggccg gcttcattaa aaggcagctg gtggagacca ggcagatcac caagcacgtg	3540
gcccagatcc tggacagcag gatgaacacc aagtacgacg agaacgacaa gctgatcagg	3600
gaggtgaagg tgatcacccct gaagagcaag ctggtgagcg acttcaggaa ggacttcag	3660
ttctacaagg tgagggagat caataattac caccacgccc acgacgccta cctgaacgcc	3720
gtggtgggca ccgccctgat taaaaagtac ccgaagctgg agagcgagtt cgtgtacggc	3780
gactacaagg tgtacgacgt gaggaagatg atcgccaaga gcgagcagga gatcggaag	3840
gccaccgcca agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc	3900
ctggccaacg gcgagatcag gaagaggccg ctgatcgaga ccaacggcga gaccggcgag	3960
atcgtgtggg acaagggcag ggacttcgcc accgtgagga aggtgctgtc catgccgcag	4020
gtgaacatcg tgaagaagac cgaggtgcag accggcggct tcagcaagga gagcatcctg	4080
ccgaagagga acagcgacaa gctgatcgcc aggaagaagg actgggatcc gaagaagtac	4140
ggcggcttcg acagcccgcac cgtggcctac agcgtgctgg tggtgccaa ggtggagaag	4200
ggcaagagca agaagctgaa gagcgtgaag gagctggtgg gcatcaccat catggagagg	4260
agcagcttcg agaagaaccc agtggacttc ctggaggcca agggctacaa ggaggtgaag	4320
aaggacctga tcattaaact gccgaagtac agcctgttcg agctggagaa cggcaggaag	4380
aggatgctgg ccagcgccgg cgagctgcag aagggaacg agctggccct gccgagcaag	4440
tacgtgaact tcctgtacct ggccagccac tacgagaagc tgaagggcag cccggaggac	4500
aacgagcaga agcagctgtt cgtggagcag cacaagcact acctggacga gatcatcgag	4560
cagatcagcg agttcagcaa gaggtgatc ctggccgacg ccaacctgga caaggtgctg	4620
agcgctaca acaagcacag ggacaagccg atcagggagc aggccgagaa catcatccac	4680
ctgttcaccc tgaccaacct gggcgccccg gccgccttca agtacttcga caccaccatc	4740
gacaggaaga ggtacaccag caccaaggag gtgctggacg ccacctgat ccaccagagc	4800

atcaccggcc	tgtacgagac	caggatcgac	ctgagccagc	tgggcggcga	cagcagcccg	4860	
ccgaagaaga	agaggaaggt	gagctggaag	gacgccagcg	gctggagcag	gatgtgaagc	4920	
ttgatcgttc	aaacatttgg	caataaagtt	tcttaagatt	gaatcctggt	gccggtcttg	4980	
cgatgattat	catataat	ctgttgaatt	acgttaagca	tgtataaatt	aacatgtaat	5040	
gcatgacggt	at	ttatgaga	tgggttttta	tgattagagt	cccgcaatta	tacatttaat	5100
acgcgataga	aaacaaaata	tagcgcgcaa	actaggataa	attatcgcg	gcggtgtcat	5160	
ctatgttact	agatcgggac	cggaattcgg	gaccctttgt	gaaagttgaa	ttacggcata	5220	
gccgaaggaa	taacagaatc	gtttcacact	ttcgtaacaa	aggctcttct	atcatgtttc	5280	
agacgatgga	ggcaaggctg	atcaaagtga	tcaagcacat	aaacgcattt	ttttaccatg	5340	
tttcaactcca	taagcgtctg	agattatcac	aagtcacgtc	tagtagtttg	atggtacact	5400	
agtgacaatc	agttcgtgca	gacagagctc	atacttgact	acttgagcga	ttacaggcga	5460	
aagtgtgaaa	cgcatgtgat	gtgggctggg	aggaggagaa	tatatactaa	tgggccgtat	5520	
cctgatttgg	gctgcgtcgg	aagggtgcagc	ccacgcgcgc	cgtaccgcgc	gggtggcgct	5580	
gctacccact	ttagtccggt	ggatggggat	ccgatgggtt	gcgcggtggc	gttgcggggg	5640	
atgttttagta	ccacatcgga	aaccgaaaga	cgatggaacc	agcttataaa	cccgcgcgct	5700	
gtagtcagct	tgtgcagtcg	aagtaatcgg	cggtttttaga	gctagaaata	gcaagttaaa	5760	
ataaggctag	tccgttatca	acttgaaaaa	gtggcaccga	gtcggtgctt	tttttttcct	5820	
aggctgcagt	gcagcgtgac	ccggtcgtgc	ccctctctag	agataatgag	cattgcatgt	5880	
ctaagttata	aaaaattacc	acatat	tttgtcacac	ttgtttgaag	tgcagtttat	5940	
ctatctttat	acatatat	aaactttact	ctacgaataa	tataatctat	agtactacaa	6000	
taatatcagt	gttttagaga	atcatataaa	tgaacagtta	gacatggtct	aaaggacaat	6060	
tgagtatttt	gacaacagga	ctctacagtt	ttatcttttt	agtgtgcatg	tgttctcctt	6120	
tttttttgca	aatagcttca	cctatataat	acttcatcca	ttttattagt	acatccattt	6180	
agggttttagg	gttaatgggt	tttatagact	aattttttta	gtacatctat	tttattctat	6240	
tttagcctct	aaattaagaa	aactaaaact	ctat	tttttagt	ttttttat	aataatttag	6300
atataaaaata	gaataaaaata	aagtgactaa	aaattaacaa	aatacccttt	aagaaattaa	6360	
aaaaactaag	gaaacatttt	tcttggttcg	agtagataat	gccagcctgt	taaacgccgt	6420	
cgacgagtct	aacggacacc	aaccagcgaa	ccagcagcgt	cgcgtcgggc	caagcgaagc	6480	
agacggcacg	gcatctctgt	cgctgcctct	ggaccctct	cgagagttcc	gctccaccgt	6540	
tggacttgct	ccgctgtcgg	catccagaaa	ttgcgtggcg	gagcggcaga	cgtgagccgg	6600	
cacggcaggc	ggcctcctcc	tcctctcacg	gcaccggcag	ctacggggga	ttcctttccc	6660	
accgctcctt	cgctttccct	tcctcgcccg	ccgtaataaa	tagacacccc	ctccacaccc	6720	

tctttcccca acctcgtggt gttcggagcg cacacacaca caaccagatc tccccaaat	6780
ccacccgctcg gcacctccgc ttcaaggtac gccgctcgtc ctcccccccc cccctctct	6840
accttctcta gatcggcggt ccggtccatg gttagggccc ggtagttcta cttctgttca	6900
tgtttgtggt agatccgtgt ttgtgttaga tccgtgctgc tagcgttcgt acacggatgc	6960
gacctgtacg tcagacacgt tctgattgct aacttgccag tgtttctctt tggggaatcc	7020
tgggatggct ctagccgttc cgcagacggg atcgatttca tgattttttt tgtttcgttg	7080
catagggttt ggtttgccct tttcctttat ttcaatatat gccgtgcact tgtttgtcgg	7140
gtcatctttt catgcttttt tttgtcttgg ttgtgatgat gtggtctggt tgggcggtcg	7200
ttctagatcg gagtagaatt ctgtttcaaa ctacctggtg gatttattaa ttttggatct	7260
gtatgtgtgt gccatacata ttcatagtta cgaattgaag atgatggatg gaaatatcga	7320
tctaggatag gtatacatgt tgatgcgggt tttactgatg catatacaga gatgcttttt	7380
gttcgcttgg ttgtgatgat gtggtgtggt tgggcggtcg ttcattcgtt ctagatcgga	7440
gtagaatact gtttcaaact acctggtgta tttattaatt ttggaactgt atgtgtgtgt	7500
catacatctt catagttacg agtttaagat ggatggaaat atcgatctag gataggtata	7560
catgttgatg tgggttttac tgatgcatat acatgatggc atatgcagca tctattcata	7620
tgctctaacc ttgagtacct atctattata ataaacaagt atgttttata attattttga	7680
tcttgatata cttggatgat ggcatatgca gcagctatat gtggattttt ttagccctgc	7740
cttcatacgc tatttatttg cttggtactg tttcttttgt cgatgctcac cctgttggtt	7800
ggtgttactt ctgcagtgc taaatagatg cagaagctga tcaacagcgt gcagaactac	7860
gcctggggca gcaagaccgc cctgaccgag ctgtacggca tggagaacct cagcagccag	7920
cccatggccg agctgtggat gggcgccac cccaagagct caagccgct gcagaacgcc	7980
gccggcgata tcgttagcct gcgcgacgtg atcgagagcg acaagagcac cctgctgggc	8040
gaggccgtgg ccaagcgctt cggcgagctg cccttcctgt tcaaggtgct gtgcgcgct	8100
cagcccctga gcatccaggt gcaccctaac aagcacaaca gcgagatcgg cttcgccaag	8160
gagaacgccg ccggcatccc catggacgcc gccgagcgca actacaagga cccaaccac	8220
aagcccgagc tgggtgttcgc cctgaccccc ttcttgcca tgaacgcctt ccgcgagttc	8280
agcgagatcg ttagcctgct gcagcccgtg gccggcgccc acccgctat cgcocacttc	8340
cttcagcagc ccgacgccga gcgcctgagc gagctgttcg ccagcctgct gaacatgcag	8400
ggtgaggaga agtcacgcgc cctggccatc ctgaagagcg ccctggacag ccagcagggc	8460
gagccctggc agacaatccg cctgatcagc gagttctacc ccgaggatag cggcctgttc	8520
agccccctgc tgctgaacgt ggtgaagctg aacccggcg aggccatgtt cctgttcgcc	8580
gagaccccc acgcctacct gcagggcggt gccctggagg tgatggccaa cagcgacaac	8640

gtgctgcgcg cgggcctgac cccaagtac atcgacatcc ccgagctggt ggccaacgtg 8700
 aagttcgagg ctaagcccgc caaccagctg ctgaccagc ccgtgaagca gggcgccgag 8760
 ctggacttcc ctatccccgt tgacgacttc gccttcagcc tgcacgacct gagcgacaag 8820
 gagaccacta tcagccagca gagcgccgcg atcctgttct gcgtggaggg cgacgccacc 8880
 ctgtggaagg gcagccagca gctgcagctg aagcccggcg agagcgctt tatcgccgcc 8940
 aacgagagcc ccgtgaccgt gaagggccac ggccgcctgg cccgcgtgta caacaagctg 9000
 tgatagctac gtcattgggtc gtttaagctg ccgatgtgcc tgcgtcgtct ggtgccctct 9060
 ctccatatgg aggttgtcaa agtatctgct gttcgtgtca tgagtcgtgt cagtgttggt 9120
 ttaataatgg accggttggt ttgtgtgtgc gtactacca gaactatgac aaatcatgaa 9180
 taagtttgat gtttgaaatt aaagcctgtg ctcatatgt tctgtcttcc agttgtctcc 9240
 taatatattgc ctgcaggtac tggctatcta ccgtttctta cttaggaggt gtttgaatgc 9300
 actaaaacta atagttagtg gctaaaatta gttaaaacat ccaaaccacca tagctaatag 9360
 ttgaactatt agctattttt ggaaaattag ttaatagtga ggtagttatt tgtagctag 9420
 ctaattcaac taacaatttt tagccaacta acaattagtt tcagtgcatt caaacacccc 9480
 cttaatgta acgtggttct atctaccgtc tcctaataata tggttgattg ttcggtttgt 9540
 tgctatgcta ttgggttctg attgctgcta gttcttgctg aatccagaag ttctcgtagt 9600
 atagctcaga ttcatattat ttatttgagt gataagtgt ccaggttatt actatgttag 9660
 ctaggttttt tttaacaagga taaattatct gtgatcataa ttcttatgaa agctttatgt 9720
 ttcttgagg cagtggcatg caatgcatga cagcaacttg atcacaccag ctgaggtaga 9780
 tacggtaaca aggttcttaa atctgttcac caaatcattg gagaacacac atacacattc 9840
 ttgccagtct tggtagaga aatttcatga caaatgcc aagctgtctt gactcttcac 9900
 ttttggccat gagtcgtgac ttagtttggt ttaatggacc ggttctccta gcttgttcta 9960
 ctcaaaactg ttgttgatgc gaataagttg tgatggttga tctctggatt ttgttttgc 10020
 ctcaatagtg gacgagatta gatagcggac cgctgcagg ccggggggcg cgccctaatt 10080
 agctaacggc caggatcgcc gcgtgagcct ttagcaacta gctagattaa ttaacgcaat 10140
 ctgttattaa gttgtctaag cgtcaatttg tttaaccac aatatatcct gccaccagcc 10200
 agccaacagc tccccgaccg gcagctggc acaaatcac cactcgatac aggcagccca 10260
 tcag 10264

<210> 40
 <211> 10692
 <212> ADN
 <213> Trình tự nhân tạo
 <220>

<223> Casset biểu hiện của cấu trúc 23501, gARN đích 2 dẫn kép

<220>

<221> vùng khởi động

<222> (214)..(734)

<223> 35S

<220>

<221> gen

<222> (748)..(4917)

<223> Cas9

<220>

<221> đặc điểm_hỗn tạp

<222> (4852)..(4908)

<223> NLS-TAG

<220>

<221> trình tự kết thúc

<222> (4923)..(5175)

<223> tNOS

<220>

<221> vùng khởi động

<222> (5196)..(5570)

<223> pU3

<220>

<221> đặc điểm_hỗn tạp

<222> (5572)..(5590)

<220>

<221> đặc điểm_hỗn tạp

<222> (5591)..(5635)

<223> crRNA

<220>

<221> vùng khởi động

<222> (5642)..(6157)

<223> pU6

<220>

<221> đặc điểm_hỗn tạp

<222> (6158)..(6245)

<223> tracrRNA

<220>

<221> vùng khởi động

<222> (6252)..(8244)

<223> Trình tự khởi động Ubi

<220>

<221> gen

<222> (8256)..(9434)

<223> PMI

<220>

<221> trình tự kết thúc

<222> (9439)..(10473)

<223> Trình tự kết thúc Ubi

<400> 40

cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa

60

atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt	120
caaacactga tagttttaaac tggcactagc ctaacggtgt tgactaacta ggccgcttcc	180
ctaattagct aaccgcgggg cgcgccggga ccgagtcaaa gattcaaata gaggacctaa	240
cagaactcgc cgtaaagact ggcgaaacagt tcatacagag tctcttacga ctcaatgaca	300
agaagaaaat cttcgtcaac ttggtggagc acgacacgct agtctactcc aaaaatatca	360
aagatacagt ctcagaagac caaagggcaa ttgagacttt tcaacaaagg gtaatatccg	420
gaaacctcct cggattccat tgcccagcta tctgtcactt aattgtgaag atagtggaaa	480
aggaaggtgg ctctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg	540
cctctgccga cagtgggtccc aaagatggac cccacccac gaggagcatc gtggtaaaag	600
aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa	660
gggatgacgc acaatcccac tatccttcgc aagacccttc ctctatataa ggaagttcat	720
ttcatttgga gaggataatt atccaccatg gacaagaagt acagcatcgg cctggacatc	780
ggcaccaaça gcgtgggctg ggccgtgatc accgacgagt acaaggtgcc gagcaagaag	840
ttcaaggtgc tgggcaacac cgacaggcac agcatcaaga agaacctgat cggcgccctg	900
ctgttcgaca gcggcgagac cgccgaggcc accaggctga agaggaccgc caggaggagg	960
tacaccagga ggaagaacag gatctgctac ctgcaggaga tcttcagcaa cgagatggcc	1020
aaggtggacg acagcttctt ccacaggctg gaggagagct tcctgggtgga ggaggacaag	1080
aagcacgaga ggcaccgat cttcggaac atcgtggacg aggtggccta ccacgagaag	1140
taccgcacca tctaccacct gaggaagaag ctggtggaca gcaccgaca ggccgacctg	1200
aggtgatct acctggccct ggcccacatg atcaagttca ggggccactt cctgatcgag	1260
ggcgacctga acccggacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc	1320
taçaaccagc tgttcgagga gaacctgatc aacgccagcg gcgtggacgc caaggccatc	1380
ctgagcgcca ggctgagcaa gagcaggagg ctggagaacc tgatcgcca gctgccgggc	1440
gagaagaaga acggcctgtt cggcaacctg atcgccctga gcctgggcct gacccgaac	1500
ttcaagagca acttcgacct ggccgaggac gccaagctgc agctgagcaa ggacacctac	1560
gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgttctctg	1620
gccgccaaga acctgagcga cgccatcctg ctgagcgaca tcctgagggt gaacaccgag	1680
atcaccaagg ccccgctgag cgccagcatg atcaagaggt acgacgagca ccaccaggac	1740
ctgacctgc tgaaggccct ggtgaggcag cagctgccgg agaagtacaa ggagatcttc	1800
ttcgaccaga gcaagaacgg ctacgccggc tacatcgacg gcggcgccag ccaggaggag	1860
ttctacaagt tcatcaagcc gatcctggag aagatggacg gcaccgagga gctgctggtg	1920
aagctgaaca gggaggacct gctgaggaag cagaggacct tcgacaacgg cagcatcccc	1980

caccagatcc acctgggcca gctgcacgcc atcctgagga ggcaggagga cttctacccg	2040
ttcctgaagg acaacaggga gaagatcgag aagatcctga ccttccgcat cccgtactac	2100
gtgggcccgc tggccagggg caacagcagg ttcgcctgga tgaccaggaa gagcgaggag	2160
accatcaccc cgtggaactt cgaggagggtg gtggacaagg gcgccagcgc ccagagcttc	2220
atcgagagga tgaccaactt cgacaagaac ctgccgaacg agaagggtgct gccgaagcac	2280
agcctgctgt acgagtactt caccgtgtac aacgagctga ccaagggtgaa gtacgtgacc	2340
gagggcatga ggaagccggc cttcctgagc ggcgagcaga agaaggccat cgtggacctg	2400
ctgttcaaga ccaacaggaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag	2460
atcgagtgtc tcgacagcgt ggagatcagc ggcgtggagg acaggttcaa cgccagcctg	2520
ggcacctacc acgacctgtc gaagatcatc aaggacaagg acttcctgga caacgaggag	2580
aacgaggaca tcctggagga catcgtgctg accctgacct tgttcgagga cagggagatg	2640
atcgaggaga ggctgaagac ctacgcccac ctgttcgacg acaagggtgat gaagcagctg	2700
aagaggagga ggtacaccgg ctggggcagg ctgagcagga agctgatcaa cggcatcagg	2760
gacaagcaga gcggcaagac catcctggac ttcctgaaga gcgacggctt cgccaacagg	2820
aacttcatgc agctgatcca cgacgacagc ctgaccttca aggaggacat ccagaaggcc	2880
caggtgagcg gccaggggca cagcctgcac gagcacatcg ccaacctggc cggcagcccg	2940
gccatcaaga agggcatcct gcagaccgtg aagggtggtg acgagctggt gaagggtgatg	3000
ggcaggcaca agccggagaa catcgtgatc gagatggcca gggagaacca gaccaccag	3060
aagggccaga agaacagcag ggagaggatg aagaggatcg aggagggcat caaggagctg	3120
ggcagccaga tcctgaagga gcacccggtg gagaacaccc agctgcagaa cgagaagctg	3180
tacctgtact acctgcagaa cggcaggga atgtacgtg accaggagct ggacatcaac	3240
aggtgagcg actacgacgt ggaccacatc gtgccgcaga gcttcctgaa ggacgacagc	3300
atcgacaaca aggtgctgac caggagcgac aagaacaggg gcaagagcga caacgtgccg	3360
agcgaggagg tggatgaagaa gatgaaaaac tactggaggc agctgctgaa cgccaagctg	3420
atcaccacga ggaagttcga caacctgacc aaggccgaga ggggcggcct gagcgagctg	3480
gacaaggccg gcttcattaa aaggcagctg gtggagacca ggcagatcac caagcacgtg	3540
gccagatcc tggacagcag gatgaacacc aagtacgacg agaacgacaa gctgatcagg	3600
gaggtgaagg tgatcacctt gaagagcaag ctggtgagcg acttcaggaa ggacttcag	3660
ttctacaagg tgaggagat caataattac caccacgccc acgacgccta cctgaacgcc	3720
gtggtgggca ccgccctgat taaaaagtac ccgaagctgg agagcgagtt cgtgtacggc	3780
gactacaagg tgtacgacgt gaggaagatg atcgccaaga gcgagcagga gatcggaag	3840
gccaccgcca agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc	3900

ctggccaacg gcgagatcag gaagaggccg ctgatcgaga ccaacggcga gaccggcgag	3960
atcgtgtggg acaagggcag ggacttcgcc accgtgagga aggtgctgtc catgccgcag	4020
gtgaacatcg tgaagaagac cgaggtgcag accggcggtt tcagcaagga gagcatcctg	4080
ccgaagagga acagcgacaa gctgatcgcc aggaagaagg actgggatcc gaagaagtac	4140
ggcggcttcg acagcccgac cgtggcctac agcgtgctgg tggaggccaa ggtggagaag	4200
ggcaagagca agaagctgaa gagcgtgaag gagctggtgg gcatcaccat catggagagg	4260
agcagcttcg agaagaacct agtggacttc ctggaggcca agggctacaa ggaggtgaag	4320
aaggacctga tcattaaact gccgaagtac agcctgttcg agctggagaa cggcaggaag	4380
aggatgctgg ccagcgccgg cgagctgcag aagggaacg agctggccct gccgagcaag	4440
tacgtgaact tctgtacct ggccagccac tacgagaagc tgaagggcag cccggaggac	4500
aacgagcaga agcagctgtt cgtggagcag cacaagcact acctggacga gatcatcgag	4560
cagatcagcg agttcagcaa gagggtgatc ctggccgacg ccaacctgga caaggtgctg	4620
agcgcctaca acaagcacag ggacaagccg atcagggagc aggccgagaa catcatccac	4680
ctgttcaccc tgaccaacct gggcgccccg gccgccttca agtacttcga caccaccatc	4740
gacaggaaga ggtacaccag caccaaggag gtgctggacg ccacctgat ccaccagagc	4800
atcaccggcc tgtacgagac caggatcgac ctgagccagc tgggcggcga cagcagcccg	4860
ccgaagaaga agaggaaggt gagctggaag gacgccagcg gctggagcag gatgtgaagc	4920
ttgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg	4980
cgatgattat catataatct ctgttgaatt acgttaagca tgtaataatt aacatgtaat	5040
gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat	5100
acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat	5160
ctatgttact agatcgggac cggaattcgg gaccgggat ctttaaacad acgaacagat	5220
cacttaaagt tcttctgaag caacttaaag ttatcaggca tgcattggatc ttggagggaat	5280
cagatgtgca gtcagggacc atagcacagg acaggcgtct tctactggtg ctaccagcaa	5340
atgctggaag ccgggaacac tgggtacgtt ggaaaccacg tgatgtggag taagataaac	5400
tgtaggagaa aagcatttcg tagtgggcca tgaagccttt caggacatgt attgcagtat	5460
gggccggccc attacgcaat tggacgacaa caaagactag tattagtacc acctcggcta	5520
tccacataga tcaaagctgg tttaaaagag ttgtgcagat gatccgtggc agagaccggc	5580
aggtagctcg gtttttagagc tatgctgttt tgaatggtcc caaaactttt tttttactag	5640
ttttgtgaaa gttgaattac ggcatagccg aaggaataac agaatcgttt cacactttcg	5700
taacaaaggt cttcttatca tgtttcagac gatggaggca aggctgatca aagtgatcaa	5760
gcacataaac gcattttttt accatgtttc actccataag cgtctgagat tatcacaagt	5820

cacgtctagt agtttgatgg tacactagtg acaatcagtt cgtgcagaca gagctcatac	5880
ttgactactt gagcgattac aggcgaaagt gtgaaacgca tgtgatgtgg gctgggagga	5940
ggagaatata tactaatggg ccgtatcctg atttgggctg cgtcggaagg tgcagcccac	6000
gcgcgccgta ccgcgcgggt ggcgctgcta cccactttag tccgttgat ggggatccga	6060
tggtttgcgc ggtggcgttg cgggggatgt ttagtaccac atcggaacc gaaagacgat	6120
ggaaccagct tataaaccg cgcgctgtag tcagcttga accattcaaa acagcatagc	6180
aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt	6240
tttttcctag gctgcagtgc agcgtgacct ggctgtgccc ctctctagag ataatgagca	6300
ttgcatgtct aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagtg	6360
cagtttatct atctttatac atatatttaa actttactct acgaataata taatctatag	6420
tactacaata atatcagtgt tttagagaat catataaatg aacagttaga catggtctaa	6480
aggacaattg agtattttga caacaggact ctacagtttt atcttttttag tgtgcatgtg	6540
ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcattccatt ttattagtag	6600
atccatttag ggttttaggt taatggtttt tatagactaa ttttttttagt acatctattt	6660
tattctattt tagcctctaa attaagaaaa ctaaaactct atttttagttt ttttatttaa	6720
taatttagat ataaaaataga ataaaaataa gtgactaaaa attaaacaaa taccctttaa	6780
gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc cagcctgtta	6840
aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg cgtcgggcca	6900
agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg agagttccgc	6960
tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga gcggcagacg	7020
tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct acgggggatt	7080
cctttccac cgctccttcg ctttcccttc ctgcgccgcc gtaataaata gacacccct	7140
ccacaccctc tttccccaac ctctgtgtgt tcggagcgca cacacacaca accagatctc	7200
ccccaaatcc acccgctcggc acctccgctt caaggtagcg cgctcgtcct ccccccccc	7260
ccctctctac cttctctaga tcggcgttcc ggtccatggt tagggcccgg tagttctact	7320
tctgttcatg tttgtgttag atccgtgttt gtgttagatc cgtgctgcta gcgttcgtac	7380
acggatgcga cctgtacgtc agacacgttc tgattgctaa cttgccagtg tttctctttg	7440
gggaatcctg ggatggctct agccgttccg cagacgggat cgatttcatg attttttttg	7500
tttcgttgca tagggtttg tttgcccttt tcctttattt caatatatgc cgtgcacttg	7560
tttgtcgggt catcttttca tgcttttttt tgtcttggtt gtgatgatgt ggtctggttg	7620
ggcggtcgtt ctagatcgga gtagaattct gtttcaaact acctggtgga tttattaatt	7680
ttggatctgt atgtgtgtgc catacatatt catagttacg aattgaagat gatggatgga	7740

aatatcgatc taggataggt atacatgttg atgcggggtt tactgatgca tatacagaga	7800
tgctttttgt tgccttggtt gtgatgatgt ggtgtgggtg ggcggtcgtt cattcgttct	7860
agatcggagt agaatactgt ttcaaactac ctgggtgtatt tattaatttt ggaactgtat	7920
gtgtgtgtca tacatcttca tagttacgag tttaagatgg atggaaatat cgatctagga	7980
taggtataca tgttgatgtg gggttttactg atgcatatac atgatggcat atgcagcatc	8040
tattcatatg ctctaacctt gagtacctat ctattataat aaacaagtat gttttataat	8100
tattttgatc ttgatatact tggatgatgg catatgcagc agctatatgt ggattttttt	8160
agccctgcct tcatacgcta tttatttgct tgggtactgtt tcttttgctg atgctcacc	8220
tgttgtttggtg tgttacttct gcagtgacta aatagatgca gaagctgatc aacagcgtgc	8280
agaactacgc ctggggcagc aagaccgccc tgaccgagct gtacggcatg gagaacccca	8340
gcagccagcc catggccgag ctgtggatgg gcgcccaccc caagagctca agccgcgtgc	8400
agaacgccgc cggcgatatc gttagcctgc gcgacgtgat cgagagcgac aagagcacc	8460
tgctgggcga ggccgtggcc aagcgcttcg gcgagctgcc ctctctgttc aaggtgctgt	8520
gcgccgctca gcccctgagc atccaggtgc accctaaca gcacaacagc gagatcggt	8580
tcgccaagga gaacgcgcgc ggcattccca tggacgcgc cgagcgcaac tacaaggacc	8640
ccaaccacaa gcccgagctg gtgttcgccc tgacccccctt cctggccatg aacgccttc	8700
gcgagttcag cgagatcggt agcctgctgc agcccgtggc cggcgccac cccgctatcg	8760
cccacttct tcagcagccc gacgccgagc gcctgagcga gctgttcgcc agcctgctga	8820
acatgcaggg tgaggagaag tcacgcgccc tggccatcct gaagagcgcc ctggacagcc	8880
agcagggcga gccctggcag acaatccgcc tgatcagcga gttctacccc gaggatagcg	8940
gcctgttcag ccccctgctg ctgaacgtgg tgaagctgaa ccccgcgag gccatgttcc	9000
tgttcgccga gacccccac gcctacctgc agggcgtggc cctggaggtg atggccaaca	9060
gcgacaacgt gctgcgcgcc ggcctgacct ccaagtacat cgacatcccc gagctgggtg	9120
ccaacgtgaa gttcgaggct aagcccgcca accagctgct gaccagccc gtgaagcagg	9180
gcgcgagct ggacttccct atccccgttg acgacttcgc cttcagcctg cacgacctga	9240
gcgacaagga gaccactatc agccagcaga gcgcccgat cctgttctgc gtggagggcg	9300
acgccaccct gtggaaggc agccagcagc tgcagctgaa gcccgcgag agcgccttta	9360
tcgccgccaa cgagagcccc gtgaccgtga agggccacgg ccgcctggcc cgcgtgtaca	9420
acaagctgtg atagctacgt catgggtcgt ttaagctgcc gatgtgcctg cgtcgtctgg	9480
tgccctctct ccatatggag gttgtcaaag tatctgctgt tcgtgtcatg agtcgtgtca	9540
gtgttggttt aataatggac cggttggtgt gtgtgtgcgt actaccaga actatgacaa	9600
atcatgaata agtttgatgt ttgaaattaa agcctgtgct cattatgttc tgtctttcag	9660


```

ttgtctccta atatttgcct gcaggtactg gctatctacc gtttcttact taggaggtgt      9720
ttgaatgcac taaaactaat agttagtggc taaaattagt taaaacatcc aaacaccata      9780
gctaatagtt gaactattag ctatTTTTTgG aaaattagtt aatagtgagg tagttatttg      9840
ttagctagct aattcaacta acaatttttta gccaaactaac aattagtttc agtgcattca      9900
aacaccccct taatgttaac gtggttctat ctaccgtctc ctaatatatg gttgattgtt      9960
cggtttggtg ctatgctatt gggttctgat tgctgctagt tcttgctgaa tccagaagtt     10020
ctcgtagtat agctcagatt catattatTTT atttgagtga taagtgatcc aggttattac     10080
tatgttagct aggtTTTTTTTT tacaaggata aattatctgt gatcataatt cttatgaaag     10140
ctttatgttt cctggaggca gtggcatgca atgcatgaca gcaacttgat cacaccagct     10200
gaggtagata cggtaacaag gttctttaaT ctgttcacca aatcattgga gaacacacat     10260
acacattctt gccagtcttg gtttagagaaa tttcatgaca aaatgccaaa gctgtcttga     10320
ctcttcactt ttggccatga gtcgtgactt agtttggttt aatggaccgg ttctcctagc     10380
ttgttctact caaaactgtt gttgatgcga ataagttgtg atgggtgatc tctggatttt     10440
gttttgctct caatagtgga cgagattaga tagcggaccg cctgcaggcc cggggg'gcgcg     10500
ccctaattag ctaacggcca ggatcgccgc gtgagccttt agcaactagc tagattaatt     10560
aacgcaatct gttattaagt tgtctaagcg tcaatttggt tacaccacaa tatatcctgc     10620
caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca ctogatacag     10680
gcagcccatc ag                                                                10692

```

<210> 41

<211> 10262

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Casset biểu hiện của cấu trúc 23501, gRNA đích 2 dẫn đơn

<220>

<221> vùng khởi động

<222> (214)..(734)

<223> Trình tự khởi động 35S

<220>

<221> gen

<222> (748)..(4917)

<223> Cas9

<220>

<221> trình tự kết thúc

<222> (4923)..(5175)

<223> tNOS

<220>

<221> vùng khởi động

<222> (5196)..(5711)

<220>

<221> đặc điểm hỗn tạp

<222> (5712)..(5730)

<220>

<221> đặc điểm hỗn tạp

<222> (5712)..(5815)

<223> sgRNA

<220>

<221> vùng khởi động

<222> (5822)..(7814)

<223> Trình tự khởi động Ubi

<220>

<221> gen

<222> (7826)..(9004)

<223> PMI

<220>

<221> trình tự kết thúc

<222> (9009)..(10043)

<223> Trình tự kết thúc Ubi

<400> 41

cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac gcccttttaa	60
atatccgatt attctaataa acgctctttt ctcttaggtt taccgcgcaa tatatcctgt	120
caaacactga tagttttaaac tggcactagc ctaacggtgt tgactaacta ggccgcttcc	180
ctaattagct aaccggggg cgcgccggga ccgagtcaaa gattcaaata gaggacctaa	240
cagaactcgc cgtaaagact ggccaacagt tcatacagag tctcttacga ctcaatgaca	300
agaagaaaat ctctgtcaac ttggtggagc acgacacgct agtctactcc aaaaatatca	360
aagatacagt ctcaagaagc caaagggcaa ttgagacttt tcaacaaagg gtaatatccg	420
gaaacctcct cggattccat tgcccagcta tctgtcactt aattgtgaag atagtggaaa	480
aggaaggttg ctctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg	540
cctctgccga cagtgggtccc aaagatggac cccacccac gaggagcatc gtggtaaaag	600
aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa	660
gggatgacgc acaatcccac tacccttcgc aagacccttc ctctatataa ggaagttcat	720
ttcatttgga gaggataatt atccaccatg gacaagaagt acagcatcgg cctggacatc	780
ggcaccaaca gcgtgggctg ggccgtgatc accgacgagt acaagggtgcc gagcaagaag	840
ttcaaggtgc tgggcaacac cgacaggcac agcatcaaga agaactgat cggcgccttg	900
ctgttcgaca gcggcgagac cgccgaggcc accaggctga agaggaccgc caggaggagg	960
tacaccagga ggaagaacag gatctgctac ctgcaggaga tcttcagcaa cgagatggcc	1020
aaggtggacg acagcttctt ccacaggctg gaggagagct tcctgggtgga ggaggacaag	1080
aagcacgaga ggcacccgat cttcggcaac atcgtggacg aggtggccta ccacgagaag	1140

tacccgacca	tctaccacct	gaggaagaag	ctggtggaca	gcaccgacaa	ggccgacctg	1200
aggctgatct	acctggccct	ggcccacatg	atcaagttca	ggggccactt	cctgatcgag	1260
ggcgacctga	acccggacaa	cagcgacgtg	gacaagctgt	tcattccagct	ggtgcagacc	1320
tacaaccagc	tgttcgagga	gaacccgatc	aacgccagcg	gcgtggacgc	caaggccatc	1380
ctgagcgcca	ggctgagcaa	gagcaggagg	ctggagaacc	tgatcgccca	gctgccgggc	1440
gagaagaaga	acggcctgtt	cggcaacctg	atcgccctga	gcctgggcct	gaccccgaac	1500
ttcaagagca	acttcgacct	ggccgaggac	gccaaagctgc	agctgagcaa	ggacacctac	1560
gacgacgacc	tggacaacct	gctggcccag	atcggcgacc	agtacgccga	cctgttcctg	1620
gccgccaaga	acctgagcga	cgccatcctg	ctgagcgaca	tcctgagggt	gaacaccgag	1680
atcaccaagg	ccccgctgag	cgccagcatg	atcaagaggt	acgacgagca	ccaccaggac	1740
ctgacctgtc	tgaaggccct	ggtgaggcag	cagctgccgg	agaagtacaa	ggagatcttc	1800
ttcgaccaga	gcaagaacgg	ctacgccggc	tacatcgacg	gcggcgccag	ccaggaggag	1860
ttctacaagt	tcattcaagcc	gatcctggag	aagatggacg	gcaccgagga	gctgctggtg	1920
aagctgaaca	gggaggacct	gctgaggaag	cagaggacct	tcgacaacgg	cagcatcccc	1980
caccagatcc	acctgggcga	gctgcacgcc	atcctgagga	ggcaggagga	cttctacccg	2040
ttcctgaagg	acaacaggga	gaagatcgag	aagatcctga	ccttcgcgat	ccgtacttac	2100
gtgggcccgc	tggccagggg	caacagcagg	ttgcctgga	tgaccaggaa	gagcgaggag	2160
accatcacc	cgtggaactt	cgaggagggtg	gtggacaagg	gcgccagcgc	ccagagcttc	2220
atcgagagga	tgaccaactt	cgacaagaac	ctgccgaacg	agaagggtgct	gccgaagcac	2280
agcctgctgt	acgagtactt	caccgtgtac	aacgagctga	ccaagggtgaa	gtacgtgacc	2340
gagggcatga	ggaagccggc	cttcctgagc	ggcgagcaga	agaaggccat	cgtggacctg	2400
ctgttcaaga	ccaacaggaa	ggtgaccgtg	aagcagctga	aggaggacta	cttcaagaag	2460
atcgagtgt	tcgacagcgt	ggagatcagc	ggcgtggagg	acaggttcaa	cgccagcctg	2520
ggcacctacc	acgacctgt	gaagatcatc	aaggacaagg	acttcctgga	caacgaggag	2580
aacgaggaca	tcctggagga	catcgtgctg	acctgaccc	tgttcgagga	caggagatg	2640
atcgaggaga	ggctgaagac	ctacgcccac	ctgttcgacg	acaagggtgat	gaagcagctg	2700
aagaggagga	ggtacaccgg	ctggggcagg	ctgagcagga	agctgatcaa	cggcacgag	2760
gacaagcaga	gcggcaagac	catcctggac	ttcctgaaga	gcgacggctt	cgccaacagg	2820
aacttcatgc	agctgatcca	cgacgacagc	ctgaccttca	aggaggacat	ccagaaggcc	2880
caggtgagcg	gccagggcga	cagcctgcac	gagcacatcg	ccaacctggc	cggcagcccc	2940
gccatcaaga	agggcatcct	gcagaccgtg	aagggtggtg	acgagctggt	gaagggtgatg	3000
ggcaggcaca	agccggagaa	catcgtgatc	gagatggcca	gggagaacca	gaccacccag	3060

aagggccaga agaacagcag ggagaggatg aagaggatcg aggagggcat caaggagctg 3120
ggcagccaga tcctgaagga gcacccggtg gagaacaccc agctgcagaa cgagaagctg 3180
tacctgtact acctgcagaa cggcagggac atgtacgtgg accaggagct ggacatcaac 3240
aggctgagcg actacgacgt ggaccacatc gtgccgcaga gcttcctgaa ggacgacagc 3300
atcgacaaca aggtgctgac caggagcgac aagaacaggg gcaagagcga caacgtgccg 3360
agcgaggagg tggatgaagaa gatgaaaaac tactggaggc agctgctgaa cgccaagctg 3420
atcaccacaga ggaagtgcga caacctgacc aaggccgaga ggggcggcct gagcgagctg 3480
gacaaggccg gcttcattaa aaggcagctg gtggagacca ggcagatcac caagcacgtg 3540
gccagatcc tggacagcag gatgaacacc aagtacgacg agaacgacaa gctgatcagg 3600
gaggtgaagg tgatcacccct gaagagcaag ctggtgagcg acttcaggaa ggacttccag 3660
ttctacaagg tgaggggagat caataattac caccacgccc acgacgccta cctgaacgcc 3720
gtggtgggca ccgccctgat taaaaagtac ccgaagctgg agagcgagtt cgtgtacggc 3780
gactacaagg tgtacgacgt gaggaagatg atcgccaaga gcgagcagga gatcggcaag 3840
gccaccgcca agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc 3900
ctggccaacg gcgagatcag gaagaggccg ctgatcgaga ccaacggcga gaccggcgag 3960
atcgtgtggg acaagggcag ggacttcgcc accgtgagga aggtgctgtc catgccgcag 4020
gtgaacatcg tgaagaagac cgaggtgcag accggcggct tcagcaagga gagcatcctg 4080
ccgaagagga acagcgacaa gctgatcgcc aggaagaagg actgggatcc gaagaagtac 4140
ggcggcttcg acagcccgac cgtggcctac agcgtgctgg tgggtggcaa ggtggagaag 4200
ggcaagagca agaagctgaa gagcgtgaag gagctggtgg gcatcaccat catggagagg 4260
agcagcttcg agaagaaccc agtggacttc ctggaggcca agggctacaa ggaggtgaag 4320
aaggacctga tcattaaact gccgaagtac agcctgttcg agctggagaa cggcaggaag 4380
aggatgctgg ccagcgccgg cgagctgcag aagggcaacg agctggccct gccgagcaag 4440
tacgtgaact tcctgtacct ggccagccac tacgagaagc tgaagggcag cccggaggac 4500
aacgagcaga agcagctggt cgtggagcag cacaagcact acctggacga gatcatcgag 4560
cagatcagcg agttcagcaa gaggggtgatc ctggccgacg ccaacctgga caaggtgctg 4620
agcgctaca acaagcacag ggacaagccg atcagggagc aggccgagaa catcatccac 4680
ctgttcaccc tgaccaacct gggcgccccg gccgccttca agtacttcga caccaccatc 4740
gacaggaaga ggtacaccag caccaaggag gtgctggacg ccacctgat ccaccagagc 4800
atcaccggcc tgtacgagac caggatcgac ctgagccagc tgggcggcga cagcagcccc 4860
ccgaagaaga agaggaaggt gagctggaag gacgccagcg gctggagcag gatgtgaagc 4920
ttgatcggtc aaacatttgg caataaagtt tcttaagatt gaatcctggt gccgggtctg 4980

cgatgattat catataatTT ctgttgaatt acgttaagca tgtaataatt aacatgtaat 5040
 gcatgacggt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 5100
 acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat 5160
 ctatgttact agatcgggac cggaattcgg gaccctttgt gaaagttgaa ttacggcata 5220
 gccgaaggaa taacagaatc gtttcacact ttcgtaacaa aggtcttctt atcatgtttc 5280
 agacgatgga ggcaaggctg atcaaagtga tcaagcacat aaacgcattt ttttaccatg 5340
 tttcactcca taagcgtctg agattatcac aagtcacgtc tagtagtttg atggtacact 5400
 agtgacaatc agttcgtgca gacagagctc atacttgact acttgagcga ttacaggcga 5460
 aagtgtgaaa cgcatgtgat gtgggctggg aggaggagaa tatatactaa tgggccgtat 5520
 cctgatttgg gctgcgtcgg aaggtgcagc ccacgcgcgc cgtaccgcgc ggggtggcgt 5580
 gctaccact ttagtccgtt ggatggggat ccgatggttt gcgcggtggc gttgcggggg 5640
 atgttttagta ccacatcgga aaccgaaaga cgatggaacc agcttataaa cccgcgcgct 5700
 gtagtcagct tgagaccggc aggtacgtcg gttttagagc tagaaatagc aagttaaaat 5760
 aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt tttttcctag 5820
 gctgcagtg cgcgtgacct ggtcgtgccc ctctctagag ataagagca ttgcatgtct 5880
 aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagtg cagtttatct 5940
 atctttatac atatatTTaa actttactct acgaataata taatctatag tactacaata 6000
 atatcagtgt tttagagaat catataaatg aacagttaga catgggtctaa aggacaattg 6060
 agtatTTtga caacaggact ctacagtttt atctTTTTtag tgtgcatgtg ttctcctttt 6120
 tttttgcaaa tagcttcacc tatataatac ttcattccatt ttatttagtac atccatttag 6180
 ggttttagggT taatggTTTT tatagactaa ttttttttagt acatctatTT tattctatTT 6240
 tagcctctaa attaagaaaa ctaaaactct atTTtagttt ttttatTTaa taatttagat 6300
 ataaaaataga ataaaaataa gtgactaaaa attaaacaaa taccctTTaa gaaattaaaa 6360
 aaactaagga aacattTTTT ttgtttcgag tagataatgc cagcctgtta aacgccgtcg 6420
 acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg cgtcgggcca agcgaagcag 6480
 acggcacggc atctctgtcg ctgcctctgg acccctctcg agagttccgc tccaccgttg 6540
 gacttgctcc gctgtcggca tccagaaatt gcgtggcgga gcggcagacg tgagccggca 6600
 cggcaggcgg cctcctctc ctctcacggc accggcagct acgggggatt cctttccac 6660
 cgctccttg ctttcccttc ctgcccgc gtaataaata gacacccct ccacaccctc 6720
 tttccccaac ctggtgttgt tcggagcgca cacacacaca accagatctc ccccaaTcc 6780
 acccgtcggc acctccgctt caaggtacgc cgctcgtcct cccccccc cctctctac 6840
 cttctctaga tcggcgTtcc ggtccatggt tagggcccgg tagttctact tctgttcag 6900

tttgtgtag atccgtgttt gtgtagatc cgtgctgcta gcgttcgtac acggatgcga 6960
 cctgtacgtc agacacgttc tgattgctaa cttgccagt tttctctttg gggaatcctg 7020
 ggatggctct agccgttccg cagacgggat cgatttcattg attttttttg tttcgttgca 7080
 tagggtttg tttgcccttt tcctttatct caatatatgc cgtgcacttg tttgtcgggt 7140
 catcttttca tgcttttttt tgtcttggtt gtgatgatgt ggtctgggtg ggcggtcgtt 7200
 ctagatcggg gtagaattct gtttcaaact acctgggtga tttattaatt ttggatctgt 7260
 atgtgtgtgc catacatatt catagttacg aattgaagat gatggatgga aatatcgatc 7320
 taggataggt atacatgttg atgcgggttt tactgatgca tatacagaga tgctttttgt 7380
 tcgcttggtt gtgatgatgt ggtgtgggtg ggcggtcgtt cattcgctct agatcggagt 7440
 agaatactgt ttcaaactac ctggtgtatt tattaatttt ggaactgtat gtgtgtgtca 7500
 tacatcttca tagttacgag tttaagatgg atggaaatat cgatctagga taggtataca 7560
 tggtgatgtg ggttttactg atgcatatac atgatggcat atgcagcatc tattcatatg 7620
 ctctaacctt gagtacctat ctattataat aaacaagtat gttttataat tattttgatc 7680
 ttgatatact tggatgatgg catatgcagc agctatatgt ggattttttt agccctgcct 7740
 tcatacgcta tttatttget tggtaactgt tcttttgtcg atgctcacc tgttgttttg 7800
 tgttacttct gcagtgacta aatagatgca gaagctgatc aacagcgtgc agaactacgc 7860
 ctggggcagc aagaccgccc tgaccgagct gtacggcatg gagaacccca gcagccagcc 7920
 catggccgag ctgtggatgg gcgcccacc caagagctca agccgcgtgc agaacgccgc 7980
 cggcgatatc gttagcctgc gcgacgtgat cgagagcgac aagagcacc tgctgggcga 8040
 ggccgtggcc aagcgcttcg gcgagctgcc ctctctgttc aagggtgctgt gcgccgtca 8100
 gccctgagc atccaggtgc accctaaca gcacaacagc gagatcggct tcgccaagga 8160
 gaacgccgcc ggcattccca tggacgccgc cgagcgcaac tacaaggacc ccaaccacaa 8220
 gcccgagctg gtgttcgccc tgacccccct cctggccatg aacgccttcc gcgagttcag 8280
 cgagatcgtt agcctgctgc agcccggtgg cggcgcccac cccgctatcg cccacttcct 8340
 tcagcagccc gacgccgagc gcctgagcga gctgttcgcc agcctgctga acatgcaggg 8400
 tgaggagaag tcacgcgccc tggccatcct gaagagcgcc ctggacagcc agcagggcga 8460
 gccctggcag acaatccgcc tgatcagcga gttctacccc gaggatagcg gcctgttcag 8520
 cccctgctg ctgaacgtgg tgaagctgaa ccccgccgag gccatgttcc tgctcgcga 8580
 gacccccac gcctacctgc agggcggtgg cctggaggtg atggccaaca gcgacaacgt 8640
 gctgcgcgcc ggctgaccc ccaagtacat cgacatcccc gagctggtgg ccaacgtgaa 8700
 gttcgagget aagcccgcca accagctgct gaccagccc gtgaagcagg gcgccgagct 8760
 ggacttccct atccccgttg acgacttcgc cttcagcctg cagcacctga gcgacaagga 8820

gaccactatc agccagcaga gcgccgcgat cctgttctgc gtggagggcg acgccaccct 8880
 gtggaagggc agccagcagc tgcagctgaa gcccggcgag agcgccctta tcgccgccaa 8940
 cgagagcccc gtgaccgtga agggccacgg ccgcctggcc cgcgtgtaca acaagctgtg 9000
 atagctacgt catgggtcgt ttaagctgcc gatgtgcctg cgtcgtctgg tgccctctct 9060
 ccatatggag gttgtcaaag tatctgctgt tcgtgtcatg agtcgtgtca gtgttggttt 9120
 aataatggac cggttgtgtt gtgtgtgctg actaccaga actatgacaa atcatgaata 9180
 agtttgatgt ttgaaattaa agcctgtgct cattatgttc tgtctttcag ttgtctccta 9240
 atatttgctt gcaggtactg gctatctacc gtttcttact taggaggtgt ttgaatgcac 9300
 taaaactaat agttagtggc taaaattagt taaaacatcc aaacaccata gctaatagtt 9360
 gaactattag ctatTTTTTgg aaaattagtt aatagtggag tagttatttg ttagctagct 9420
 aattcaacta acaatTTTTta gccaaactaac aattagtttc agtgcattca aacacccctt 9480
 taatgttaac gtggttctat ctaccgtctc ctaatatatg gttgattgtt cggtttggtg 9540
 ctatgctatt gggttctgat tgctgctagt tcttgctgaa tccagaagtt ctcgtagtat 9600
 agctcagatt catattatTT atttgagtga taagtgatcc aggttattac tatgttagct 9660
 aggtTTTTTTT tacaaggata aattatctgt gatcataatt cttatgaaag ctttatgttt 9720
 cctggaggca gtggcatgca atgcatgaca gcaacttgat cacaccagct gaggtagata 9780
 cggtaacaag gttctttaaT ctgttcacca aatcattgga gaacacacat acacattctt 9840
 gccagtcttg gttagagaaa tttcatgaca aaatgcaaaa gctgtcttga ctcttcactt 9900
 ttggccatga gtctgactt agtttggttt aatggaccgg ttctcctagc ttgttctact 9960
 caaaactgtt gttgatgcga ataagttgtg atggttgatc tctggatttt gttttgctct 10020
 caatagtgga cgagattaga tagcggaccg cctgcaggcc cggggggcgcg ccctaattag 10080
 ctaacggcca ggatcgccgc gtgagccttt agcaactagc tagattaatt aacgcaatct 10140
 gttattaagt tgtctaagcg tcaatttggt tacaccacaa tatatcctgc caccagccag 10200
 ccaacagctc cccgaccggc agctcggcac aaaatcacca ctcgatacag gcagcccatc 10260
 ag 10262

<210> 42

<211> 1366

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến do TALEN ở Sự kiện 38A ID 22808-4108 alen 2

<400> 42

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60

atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120

aatacctgta gcacgaaggc gatggccggg agcgtgggtcg gcgagcccggt cgtgctgggg 180
cagaggggtga cgggtgctgac ggtggacggc ggccggcgctcc ggggtctcat cccgggaacc 240
atcctcgccct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
gactacttcg actacatcg cccaaccagc accggcggtc tcataccgc catgctcacc 360
gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
aggaagccaa agtacaacgg caagtgcag cgcagcctga ttaggagcat cctcggcgag 540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
tccgcccgcg cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
gagcgcgaat acaacctcat cgacggcggt gtggcgccca acaaccgac gatggttgcg 720
atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc 840
gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgcg gtggctccgc 900
aacaacggca tggcccccat catcgacatc ttcatggcg ccagctcgga cctggtggac 960
atccacgtcg ccgcatggt ccagtcgctc cacagcgacg gcgactacct ggcacatccag 1020
gacaactcgc tccgtggcg cgcggccacc gtggacgcgg cgacgcgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccaggg caacgtggag 1140
acaggacgaa ccggtgactg gcgaaggaa caatgccgat gccctcggtg ggctcgtag 1200
gcagctctcc gaggagagga gaacaaggct cgcgcgccgc gtctctgcca tcaacccaag 1260
aggctctaga tgtcgtcgt acgatata agacaagtgg ctttactgtc agtcacatgc 1320
ttgtaaataa gtagacttta ttttaataaa acataaaaat atatat 1366

<210> 43

<211> 1366

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do CRISPR ở Sự kiện 27A ID 22807-4073 alen 2

<400> 43

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
aatacctgta gcacgaaggc gatggccggg agcgtgggtcg gcgagcccggt cgtgctgggg 180
cagaggggtga cgggtgctgac ggtggacggc ggccggcgctcc ggggtctcat cccgggaacc 240
atcctcgccct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
gactacttcg actacatcg cccaaccagc accggcggtc tcataccgc catgctcacc 360

gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattctt ccttcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcattg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggt gtatctgccg gtggctccgc 900
 aacaacggca tggcccccatt catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcattccg 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggagaa ccggtgactg gcgaaggaag caatgccgat gccctcgggt ggctcgttag 1200
 gcagctctcc gaggagagga gaacaaggct cgcgcgccgc gtctctgcca tcaacccaag 1260
 aggctctaga tgtgcgtcgt acgatattta agacaagtgg ctttactgtc agtcacatgc 1320
 ttgtaaataa gtagacttta ttttaataaa acataaaaat atatat 1366

<210> 44

<211> 1363

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đột biến MTL do CRISPR ở Sự kiện 27A ID 22807-4081 alen 2

<400> 44

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcgggc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggccggcgtcc ggggtctcat cccgggaacc 240
 atcctcgctt tctcagaggc caggctgcag gagctggacg gaccggaggc gaggtggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcattctt ccttcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcattg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600

tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaaat acaacctcat cgacggcggt gtggcggccca acaacccgac gatgggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggt gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcggg cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acaggaaccg gtgactggcg aaggaagcaa tgccgatgcc ctcggtgggc tcgctaggca 1200
 gctctccgag gagaggagaa caaggctcgc gcgccgcgtc tctgccatca acccaagagg 1260
 ctctagatgt gcgtcgtacg atatctaaga caagtggctt tactgtcagt cacatgcttg 1320
 taaataagta gactttattt taataaaaca taaaaatata tat 1363

<210> 45
 <211> 1371
 <212> ADN
 <213> Zea mays

<400> 45
 agttcatcac taatcacact tattgttccc tcgacgagta totagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcgcca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctata gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgctc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggttggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaaag acatcaacca cttttacatg 420
 gagaactgcc cgcgcattct cctcagaag agcaggcttg cggccgcat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc totgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaaat acaacctcat cgacggcggt gtggcggccca acaacccgac gatgggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggt gtatctgccg gtggctccgc 900

aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
atccacgtcg ccgcatgtt ccagtcgtc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggg caacgtggag 1140
acagggaggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 46

<211> 1371

<212> ADN

<213> Zea mays

<400> 46

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagctcgt cgtgctgggg 180
cagaggggta cgggtgctgac ggtggacggc ggccggtcc ggggtctcat cccgggaacc 240
atcctcgctt tctcagaggc caggctgcag gagctggacg gaccggaggc gaggetggcg 300
gactacttcg actacatcgc cgggaaccagc accggcggtc tcatcacgc catgctcacc 360
gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
aggaagccaa agtacaacgg caagtgcatt cgcagcctga ttaggagcat cctcggcgag 540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
tccgccgcgc cgacctacct ccggcgccac tacttccaga ctgaagacgc caacggcaag 660
gagcgcgaat acaacctcat cgacggcggg gtggcggcca acaaccgac gatggttgcg 720
atgacgcaga tcacaaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
gagcagggcc tctacacggc gcggcagtgc tcccgggtgg gtatctgccg gtggctccgc 900
aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
atccacgtcg ccgcatgtt ccagtcgtc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggg caacgtggag 1140
acagggaggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260

ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 47
 <211> 1371
 <212> ADN
 <213> Zea mays

<400> 47
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcg cgaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc tgcgcatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcag cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccgggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcggg gtggcgggcca acaaccgcgac gatggttgcg 720
 atgacgcaga tcacaaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctgggtggac 960
 atccaogtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggaggt acgaaccggg gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
 ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 48
 <211> 1371
 <212> ADN
 <213> Zea mays

<400> 48
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccg cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgcat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcattg cgcagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcggaat acaacctcat cgacggcggg gtggcgccca acaaccgac gatggttacg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttctct gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccatt catcgacatc ttcattggcg ccagctcggg cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct ggcattccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgcggga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggaggt acgaaccggg gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggtcgcgc gcgcgtgtc tgccatcaac 1260
 ccaagaggct ctagatgtgc gtcgtaogat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatctta ataaaacata aaaatatata t 1371

<210> 49
 <211> 1371
 <212> ADN
 <213> Zea mays

<400> 49
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccg cgtgctgggg 180
 cagaggggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300

gactacttgc actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgccccgca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt cctcagaag agcaggcttg cgcccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcag cgagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcggaat acaacctcat cgacggcggg gtggcgcca acaaccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtacca 780
 gtgaagccgt cgaactgccg caggttcttg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtg tccgggtggg gtatctgcg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgac gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgccgga gaacatacgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggagg acgaaccggg gactggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
 ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 50

<211> 1371

<212> ADN

<213> Zea mays

<400> 50

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcgggc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctg gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgtgac ggtggacggc ggcggcgctc ggggtctcat cccgggaacc 240
 atctcgcct tctcagaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300
 gactacttgc actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360
 gcgccccgca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt cctcagaag agcaggcttg cgcccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcag cgagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660

gagcgcgaat acaacctcat cgacggcgggt gtggcggcca acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccgggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960
 atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
 gacaactcgc tccgtggcgc cgcgccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
 acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
 acagggaggt acgaaccggt gattggcgaa ggaagcaatg ccgatgccct cgggtgggctc 1200
 gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
 ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
 catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 51
 <211> 1371
 <212> ADN
 <213> Zea mays

<400> 51
 agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60
 atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120
 aatacctgta gcacgaaggc gatggccggg agcgtggctc gcgagcccgt cgtgctgggg 180
 cagaggggtga cgggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240
 atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggtcggcg 300
 gactacttcg actacatcgc cggaaccagc acgggcggtc tcatcacgc catgctcacc 360
 gcgcccggca aggacaagcg gcctctctac gctgccaaag acatcaacca cttttacatg 420
 gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cgcccgccat gtccgcgctg 480
 aggaagccaa agtacaacgg caagtgcatt gcgagcctga ttaggagcat cctcggcgag 540
 acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
 tccgcgcgcg cgacctacct ccgggcgcac tacttccaga ctgaagacgc caacggcaag 660
 gagcgcgaat acaacctcat cgacggcgggt gtggcggcca acaacccgac gatggttgcg 720
 atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
 gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
 gagcagggcc tctacacggc gcggcagtgc tcccgggtggg gtatctgccg gtggctccgc 900
 aacaacggca tggcccccat catcgacatc ttcattggcg ccagctcgga cctggtggac 960

atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
acagggagggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cggtagggctc 1200
gctaggcagc tctccgagga gaggaaaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
ccaagagggt ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 52

<211> 401

<212> PRT

<213> Zea mays

<400> 52

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Tyr Ser Thr Lys
1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 53

<211> 401

<212> PRT

<213> Zea mays

<400> 53

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Leu Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 54
 <211> 401
 <212> PRT
 <213> Zea mays

<400> 54

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
 100 105 110

Cys Leu Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 55
 <211> 401
 <212> PRT
 <213> Zea mays

<400> 55

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser

115		120		125
Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile				
130		135		140
Arg Ser Ile Leu Gly Glu Thr Arg Ala Lys Ser Thr Pro Leu Lys Asn				
145		150		155
				160
Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr				
		165		170
				175
Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg				
		180		185
				190
Glu Tyr Asn Leu Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met				
		195		200
				205
Val Thr Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys				
		210		215
				220
Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu				
225		230		235
				240
Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr				
		245		250
				255
Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn				
		260		265
				270
Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu				
		275		280
				285
Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly				
		290		295
				300
Asp Tyr Leu Arg Ile Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr				
305		310		315
				320
Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly				
		325		330
				335
Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly				
		340		345
				350
Arg Tyr Glu Pro Val Thr Gly Glu Gly Ser Asn Ala Asp Ala Leu Gly				
		355		360
				365
Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg				

370 375 380
 Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

 Ile

 <210> 56
 <211> 401
 <212> PRT
 <213> Zea mays

 <400> 56

 Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

 Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

 Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

 Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

 Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

 Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

 Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
 100 105 110

 Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

 Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

 Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

 Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

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

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 57
 <211> 401
 <212> PRT.
 <213> Zea mays

<400> 57

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

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

Arg Tyr Glu Pro Val Ile Gly Glu Gly Ser Asn Ala Asp Ala Leu Gly
 355 360 365

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 58

<211> 401

<212> PRT

<213> Zea mays

<400> 58

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Glu Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

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

Ala Leu Leu Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr
 165 170 175

Leu Pro Ala His Tyr Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg
 180 185 190

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

Val Ala Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys
 210 215 220

Ala Glu Glu Leu Tyr Pro Val Lys Pro Ser Asn Cys Arg Arg Phe Leu
 225 230 235 240

Val Leu Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr
 245 250 255

Ala Arg Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn
 260 265 270

Gly Met Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu
 275 280 285

Val Asp Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly
 290 295 300

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

Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly
 325 330 335

Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly
 340 345 350

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

Gly Leu Ala Arg Gln Leu Ser Glu Glu Arg Lys Thr Arg Leu Ala Arg
 370 375 380

Arg Val Ser Ala Ile Asn Pro Arg Gly Ser Arg Cys Ala Ser Tyr Asp
 385 390 395 400

Ile

<210> 59
 <211> 379
 <212> PRT
 <213> Zea mays

<400> 59

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile

130 135 140
 Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160
 Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175
 Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190
 Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205
 Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220
 Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240
 Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255
 Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270
 Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285
 Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300
 Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320
 Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335
 Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350
 Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365
 Gln Arg Val Ser Arg Val Asn Val Glu Asn Arg
 370 375

<210> 60

<211> 396

<212> PRT

<213> Zea mays

<400> 60

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Thr Gly Asp Trp
 370 375 380

Arg Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390 395

<210> 61
 <211> 379
 <212> PRT
 <213> Zea mays

<400> 61

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Asn Arg
 370 375

<210> 62

<211> 394

<212> PRT

<213> Zea mays

<400> 62

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Thr Gly Asp Trp Arg Arg
 370 375 380

Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390

<210> 63
 <211> 395
 <212> PRT
 <213> Zea mays

<400> 63

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr

152

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Met Ser
115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Thr Gly Asp Trp Arg Arg
 370 375 380

Lys Gln Cys Arg Cys Pro Arg Trp Ala
 385 390

<210> 65
 <211> 395
 <212> PRT
 <213> Zea mays

<400> 65

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175
 Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190
 Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205
 Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220
 Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240
 Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255
 Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270
 Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285
 Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300
 Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320
 Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335
 Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350
 Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365
 Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Thr Gly Asp Trp Arg
 370 375 380
 Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390 395

<210> 66
 <211> 398
 <212> PRT

<213> Zea mays

<400> 66

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Arg Glu Val Arg Thr Gly
 370 375 380

Asp Trp Arg Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390 395

<210> 67

<211> 398

<212> PRT

<213> Zea mays

<400> 67

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser

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

159

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Lys Gly Gly Thr Asn Arg
 370 375 380

<210> 69
 <211> 383
 <212> PRT
 <213> Zea mays

<400> 69

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

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Gly Thr Asn Arg
370 375 380

<210> 70

<211> 396

<212> PRT

<213> Zea mays

<400> 70

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
 180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
 195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
 210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
 225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
 245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
 260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Thr Gly Asp Trp
 370 375 380

Arg Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390 395

<210> 71
 <211> 396
 <212> PRT
 <213> Zea mays

<400> 71

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
 1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
 85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
 100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
 115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
 130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
 145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
 165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp


```

180              185              190
Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
195              200              205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
210              215              220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
225              230              235              240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
245              250              255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
260              265              270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
275              280              285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
290              295              300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
305              310              315              320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
325              330              335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
340              345              350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
355              360              365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Thr Gly Asp Trp
370              375              380

Arg Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
385              390              395

```

```

<210> 72
<211> 395
<212> PRT
<213> Zea mays

```

```

<400> 72

```

```

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Thr Cys Ser Thr Lys
1              5              10              15

```

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln Arg
20 25 30

Val Thr Val Leu Thr Val Asp Gly Gly Gly Val Arg Gly Leu Ile Pro
35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp Lys
85 90 95

Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn His Phe Tyr Met Gln Asn
100 105 110

Cys Pro Arg Ile Phe Pro Gln Lys Ser Arg Leu Ala Ala Ala Met Ser
115 120 125

Ala Leu Arg Lys Pro Lys Tyr Asn Gly Lys Cys Met Arg Ser Leu Ile
130 135 140

Arg Ser Ile Leu Gly Glu Thr Arg Val Ser Glu Thr Leu Thr Asn Val
145 150 155 160

Ile Ile Pro Ala Phe Asp Ile Arg Leu Leu Gln Pro Ile Ile Phe Ser
165 170 175

Thr Tyr Asp Ala Lys Ser Thr Pro Leu Lys Asn Ala Leu Leu Ser Asp
180 185 190

Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His Tyr
195 200 205

Phe Gln Thr Glu Asp Ala Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile
210 215 220

Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln
225 230 235 240

Ile Thr Lys Lys Met Leu Ala Ser Lys Asp Lys Ala Glu Glu Leu Tyr
245 250 255

Pro Val Asn Pro Ser Asn Cys Arg Arg Phe Leu Val Leu Ser Ile Gly
260 265 270

Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg Gln Cys Ser
 275 280 285

Arg Trp Gly Ile Cys Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile
 290 295 300

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Val
 305 310 315 320

Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Ala Ala Thr
 340 345 350

Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 355 360 365

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Thr Gly Asp Trp Arg
 370 375 380

Arg Lys Gln Cys Arg Cys Pro Arg Trp Ala Arg
 385 390 395

<210> 73
 <211> 437
 <212> PRT
 <213> Sorghum bicolor

<400> 73

Met Ala Thr Tyr Tyr Ser Ser Arg Arg Pro Cys Asn Ala Cys Ser Thr
 1 5 10 15

Lys Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Leu Gly Gln
 20 25 30

Arg Val Thr Val Leu Thr Val Asp Gly Gly Gly Ile Arg Gly Leu Ile
 35 40 45

Pro Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp
 50 55 60

Gly Pro Glu Val Arg Leu Ala Asp Tyr Phe Asp Tyr Ile Ala Gly Thr
 65 70 75 80

Ser Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Lys Asp
 85 90 95

Arg Arg Pro Leu Tyr Ala Ala Lys Asp Ile Asn Gln Phe Tyr Met Glu
 100 105 110

Asn Cys Pro Arg Ile Phe Pro Gln Lys Ser Ser Arg Leu Ala Ala Ala
 115 120 125

Met Ser Ala Leu Arg Lys Pro Arg Tyr Asn Gly Lys Cys Leu Arg Asn
 130 135 140

Leu Ile Met Ser Met Leu Gly Glu Thr Arg Val Ser Asp Thr Leu Thr
 145 150 155 160

Asn Val Ile Ile Pro Thr Phe Asp Val Arg Leu Leu Gln Pro Ile Ile
 165 170 175

Phe Ser Thr Tyr Asp Ala Lys Ser Met Pro Leu Lys Asn Ala Leu Leu
 180 185 190

Ser Asp Val Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala
 195 200 205

His Tyr Phe Gln Thr Lys Asp Ala Gly Ser Gly Lys Glu Arg Glu Tyr
 210 215 220

Asn Leu Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala
 225 230 235 240

Met Thr Gln Ile Thr Lys Lys Met Leu Ala Ser Lys Glu Lys Ala Glu
 245 250 255

Glu Leu Tyr Pro Val Lys Pro Trp Asn Cys Arg Lys Phe Leu Val Leu
 260 265 270

Ser Ile Gly Thr Gly Ser Thr Ser Glu Gln Gly Leu Tyr Thr Ala Arg
 275 280 285

Gln Cys Ser Arg Trp Gly Ile Cys Arg Trp Ile Arg Asn Asn Gly Met
 290 295 300

Ala Pro Ile Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp
 305 310 315 320

Ile His Val Ala Ala Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr
 325 330 335

Leu Arg Ile Gln Asp Asn Ser Leu His Gly Ala Ala Ala Thr Val Asp
 340 345 350

Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val Gly Ile Gly Glu Arg
355 360 365

Met Leu Ala Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Tyr
370 375 380

Glu Pro Val Pro Gly Glu Gly Ser Asn Ala Asp Ala Leu Ala Gly Ile
385 390 395 400

Ala Arg Gln Leu Ser Glu Glu Arg Arg Thr Arg Leu Ala Arg Arg Thr
405 410 415

Ser Ala Ile Val Ser Ser Gly Gly Ala Ser Arg Arg Thr Cys Ala Ser
420 425 430

Lys Val Ser Asn Val
435

<210> 74
<211> 421
<212> PRT
<213> *Setaria italica*

<400> 74

Met Ala Ser Tyr Ser Ser Arg Arg Pro Cys Asn Ala Cys Arg Thr Lys
1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Glu Pro Val Val Pro Gly Gln Arg
20 25 30

Val Thr Val Leu Thr Ile Asp Gly Gly Gly Ile Arg Gly Leu Ile Pro
35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
50 55 60

Pro Glu Ala Arg Leu Ala Asp Tyr Phe Asp Cys Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Ile Thr Thr Pro Gly Glu Asp Lys
85 90 95

Arg Pro Leu Phe Ala Ala Arg Asp Ile Asn Arg Phe Tyr Phe Asp Asn
100 105 110

Cys Pro Arg Ile Phe Pro Gln Ser Arg Ser Ser Leu Ala Ala Ala Met
115 120 125

Ser Ala Leu Arg Lys Pro Arg Tyr Asn Gly Lys Tyr Leu Arg Ser Thr
 130 135 140
 Ile Arg Ser Met Leu Gly Glu Thr Arg Val Ser Asp Ala Leu Thr Asn
 145 150 155 160
 Val Val Ile Pro Thr Phe Asp Ile Lys Leu Ile Gln Pro Ile Ile Phe
 165 170 175
 Ser Thr Tyr Asp Val Lys Asn Met Pro Leu Lys Asn Ala Leu Leu Ser
 180 185 190
 Asp Val Cys Ile Ser Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His
 195 200 205
 Tyr Phe Gln Ile Gln Asp Ala Gly Gly Lys Thr Arg Glu Tyr Asn Leu
 210 215 220
 Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr
 225 230 235 240
 Gln Ile Thr Lys Met Met Leu Ala Lys Asp Lys Glu Glu Leu Tyr Pro
 245 250 255
 Val Lys Pro Glu Asp Cys Arg Lys Phe Leu Val Leu Ser Ile Gly Thr
 260 265 270
 Gly Ser Thr Ser Asp Glu Gly Leu Phe Thr Ala Arg Gln Cys Ser Arg
 275 280 285
 Trp Gly Val Val Arg Trp Leu Arg Asn Asn Gly Met Ala Pro Ile Ile
 290 295 300
 Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Ala Ala
 305 310 315 320
 Val Leu Phe Gln Ser Leu His Ser Asp Gly His Ser Leu Arg Gly Ala
 325 330 335
 Ala Ala Thr Val Asp Ala Ala Thr Pro Glu Asn Met Arg Thr Leu Val
 340 345 350
 Gly Ile Gly Glu Arg Met Leu Ala Gln Arg Val Ser Arg Val Asn Val
 355 360 365
 Glu Thr Gly Arg Tyr Glu Pro Val Pro Gly Glu Gly Ser Asn Ala Asp
 370 375 380

Ala Leu Val Ala Leu Ala Arg Gln Leu Ser Asp Glu Arg Arg Ala Arg
385 390 395 400

Ile Ala Arg Arg Ala Ala Ala Ala Cys Ala Gly Gly Ser Arg Cys Cys
405 410 415

Ser Pro Val Lys Thr
420

<210> 75
<211> 422
<212> PRT
<213> Hordeum vulgare

<400> 75

Met Ala Ser Tyr Trp Cys Arg Arg Pro Cys Glu Ser Cys Ser Thr Arg
1 5 10 15

Ala Met Ala Gly Ser Val Val Gly Gln Pro Val Ala Pro Gly Gln Arg
20 25 30

Val Thr Val Leu Thr Ile Asp Gly Gly Gly Ile Arg Gly Leu Ile Pro
35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
50 55 60

Pro Asp Ala Arg Leu Ala Asp Tyr Phe Asp Cys Ile Ala Gly Thr Ser
65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Leu Thr Ala Pro Gly Gln Asp Gly
85 90 95

Arg Pro Leu Phe Ala Ala Lys Asp Val Asn Arg Phe Tyr Leu Asp Asn
100 105 110

Gly Pro Tyr Ile Phe Pro Gln Arg Arg Cys Ala Leu Ala Ala Val Thr
115 120 125

Ala Ser Leu Arg Arg Pro Arg Tyr Ser Gly Lys Tyr Leu His Gly Lys
130 135 140

Ile Arg Ser Met Leu Gly Glu Thr Arg Leu Cys Asp Ala Leu Thr Asp
145 150 155 160

Val Val Ile Pro Thr Phe Asp Val Lys Leu Leu Gln Pro Ile Ile Phe
165 170 175

Ser Thr Tyr Asp Ala Arg Asn Met Pro Leu Lys Asn Ala Arg Leu Ala

180	185	190
Asp Ile Cys Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His 195 200 205		
His Phe His Thr Gln Asp Asp Asn Gly Lys Glu Arg Glu Tyr Asn Leu 210 215 220		
Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Thr Met Thr 225 230 235 240		
Gln Ile Thr Lys Lys Met Met Val Lys Asp Arg Glu Glu Leu Tyr Pro 245 250 255		
Val Lys Pro Ser Asp Cys Gly Lys Phe Leu Val Leu Ser Ile Gly Thr 260 265 270		
Gly Ser Thr Ser Asp Gln Gly Leu Tyr Thr Ala Lys Gln Cys Ser Gln 275 280 285		
Trp Gly Ile Ile Arg Trp Leu Arg Asn Lys Gly Met Ala Pro Ile Ile 290 295 300		
Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Ala Ala 305 310 315 320		
Val Leu Phe Gln Ser Leu His Ser Asp Gly Asn Tyr Leu Arg Ile Gln 325 330 335		
Asp Asn Ser Leu His Gly Pro Ala Ala Thr Val Asp Ala Ala Thr Pro 340 345 350		
Glu Asn Met Ala Glu Leu Leu Arg Ile Gly Glu Arg Met Leu Ala Gln 355 360 365		
Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Tyr Glu Glu Ile Arg 370 375 380		
Gly Ala Gly Ser Asn Ala Asp Ala Leu Ala Gly Phe Ala Lys Gln Leu 385 390 395 400		
Ser Asp Glu Arg Arg Thr Arg Leu Gly Arg Arg Arg Val Gly Ala Gly 405 410 415		
Arg Leu Lys Ser Arg Arg 420		

<210> 76

<211> 423
 <212> PRT
 <213> *Brachypodium distachyon*

<400> 76

Met Ala Ser Tyr Ala Cys Arg Arg Pro Cys Glu Ser Cys Arg Thr Arg
 1 5 10 15

Ala Met Ala Gly Gly Val Val Gly Glu Pro Thr Thr Pro Gly Gln Arg
 20 25 30

Val Thr Val Leu Thr Ile Asp Gly Gly Gly Ile Arg Gly Leu Ile Pro
 35 40 45

Gly Thr Ile Leu Ala Phe Leu Glu Asp Arg Leu Gln Glu Leu Asp Gly
 50 55 60

Pro Asp Ala Arg Leu Ala Asp Tyr Phe Asp Cys Ile Ala Gly Thr Ser
 65 70 75 80

Thr Gly Gly Leu Ile Thr Ala Met Ile Thr Ala Pro Gly Glu Glu Gly
 85 90 95

Arg Pro Leu Phe Ala Ala Glu Asp Ile Asn Arg Phe Tyr Leu Asp Asn
 100 105 110

Gly Pro Gln Ile Phe Pro Gln Lys Arg Ser Ser Leu Met Ser Val Leu
 115 120 125

Ala Ser Leu Thr Arg Pro Arg Tyr Asn Gly Lys Phe Leu His Gly Lys
 130 135 140

Ile Arg Ser Met Leu Gly Glu Thr Arg Val Cys Asp Thr Leu Thr Asp
 145 150 155 160

Val Val Ile Pro Thr Phe Asp Val Arg Leu Leu Gln Pro Ile Ile Phe
 165 170 175

Ser Thr Tyr Asp Ala Lys Ser Met Pro Leu Lys Asn Ala Leu Leu Ser
 180 185 190

Asp Val Cys Ile Ser Thr Ser Ala Ala Pro Thr Phe Leu Pro Ala His
 195 200 205

Tyr Phe Gln Thr Glu Asp Asp Asn Gly Lys Val Arg Glu Tyr Asn Leu
 210 215 220

Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr
 225 230 235 240

Gln Ile Thr Lys Lys Ile Met Ala Lys Asp Lys Glu Glu Leu Tyr Pro
245 250 255

Val Lys Pro Ser Asp Cys Gly Lys Phe Leu Val Leu Ser Ile Gly Thr
260 265 270

Gly Ser Thr Ser Asp Gln Gly Leu Tyr Thr Ala Lys Gln Cys Ser Arg
275 280 285

Trp Gly Ile Ile Arg Trp Leu Arg Asn Lys Gly Met Ala Pro Ile Ile
290 295 300

Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Ala Ala
305 310 315 320

Val Leu Phe Gln Ser Leu His Ser Asp Gly Asp Cys Tyr Leu Arg Ile
325 330 335

Gln Asp Asn Ser Leu Arg Gly Ala Ala Ala Thr Val Asp Thr Ala Thr
340 345 350

Pro Asp Asn Met Arg Glu Leu Val Arg Ile Gly Glu Arg Met Leu Ala
355 360 365

Gln Arg Val Ser Lys Val Asn Val Glu Thr Gly Arg Tyr Glu Glu Met
370 375 380

Gln Gly Ala Gly Thr Asn Ala Asp Ala Leu Ala Gly Phe Ala Arg Gln
385 390 395 400

Leu Ser Asp Glu Arg Arg Ala Arg Phe Gly Pro Arg Asp Gly Ala Pro
405 410 415

Ala Asn Gly Lys Ser Arg Cys
420

<210> 77
<211> 414
<212> PRT
<213> Oryza sativa

<400> 77

Met Ala Gly Cys Val Val Gly Glu Pro Ala Ser Ala Pro Gly Gln Arg
1 5 10 15

Val Thr Leu Leu Ala Ile Asp Gly Gly Gly Ile Arg Gly Leu Ile Pro
20 25 30

Gly Thr Ile Leu Ala Phe Leu Glu Ala Arg Leu Gln Glu Leu Asp Gly
 35 40 45

Pro Asp Ala Arg Leu Ala Asp Tyr Phe Asp Cys Ile Ala Gly Thr Ser
 50 55 60

Thr Gly Gly Leu Ile Thr Ala Met Leu Ala Ala Pro Gly Asp His Gly
 65 70 75 80

Arg Pro Leu Phe Ala Ala Ser Asp Ile Asn Arg Phe Tyr Leu Asp Asn
 85 90 95

Gly Pro Arg Ile Phe Pro Gln Lys Arg Cys Gly Met Ala Ala Ala Met
 100 105 110

Ala Ala Leu Thr Arg Pro Arg Tyr Asn Gly Lys Tyr Leu Gln Gly Lys
 115 120 125

Ile Arg Lys Met Leu Gly Glu Thr Arg Val Arg Asp Thr Leu Thr Asn
 130 135 140

Val Val Ile Pro Thr Phe Asp Val Arg Leu Leu Gln Pro Thr Ile Phe
 145 150 155 160

Ser Thr Tyr Asp Ala Lys Ser Met Pro Leu Lys Asn Ala Leu Leu Ser
 165 170 175

Asp Ile Cys Ile Ser Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His
 180 185 190

Cys Phe Gln Thr Thr Asp Asp Ala Thr Gly Lys Val Arg Glu Phe Asp
 195 200 205

Leu Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met
 210 215 220

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

Pro Val Lys Pro Ser Asp Cys Gly Lys Phe Leu Val Leu Ser Leu Gly
 245 250 255

Thr Gly Ser Thr Ser Asp Gln Gly Met Tyr Thr Ala Arg Gln Cys Ser
 260 265 270

Arg Trp Gly Ile Val Arg Trp Leu Arg Asn Lys Gly Met Ala Pro Ile
 275 280 285

Ile Asp Ile Phe Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Ala
 290 295 300

Ala Val Met Phe Gln Ser Leu His Ser Asp Gly Asp Tyr Leu Arg Ile
 305 310 315 320

Gln Asp Asn Thr Leu His Gly Asp Ala Ala Thr Val Asp Ala Ala Thr
 325 330 335

Arg Asp Asn Met Arg Ala Leu Val Gly Ile Gly Glu Arg Met Leu Ala
 340 345 350

Gln Arg Val Ser Arg Val Asn Val Glu Thr Gly Arg Tyr Val Glu Val
 355 360 365

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

Leu Ser Glu Glu Arg Arg Ala Arg Leu Gly Arg Arg Asn Ala Cys Gly
 385 390 395 400

Gly Gly Gly Glu Gly Glu Pro Ser Gly Val Ala Cys Lys Arg
 405 410

<210> 78

<211> 275

<212> PRT

<213> Triticum aestivum

<400> 78

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

Pro Thr Phe Asp Val Lys Leu Leu Gln Pro Ile Ile Phe Ser Thr Tyr
 20 25 30

Asp Ala Lys Ser Met Pro Leu Lys Asn Ala Arg Leu Ala Asp Val Cys
 35 40 45

Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Ala His His Phe His
 50 55 60

Thr His Asp Gly Asn Gly Lys Glu Arg Glu Tyr Asn Leu Ile Asp Gly
 65 70 75 80

Gly Val Ala Ala Asn Asn Pro Thr Met Val Ala Met Thr Gln Ile Thr
 85 90 95

Lys Lys Met Met Gly Lys Asp Arg Glu Glu Leu Tyr Pro Val Glu Pro
 100 105 110

Ser Asp Cys Gly Lys Phe Leu Val Leu Ser Val Gly Thr Gly Ser Thr
 115 120 125

Ser Asp Gln Gly Leu Tyr Thr Ala Lys Gln Cys Ser Gln Trp Gly Ile
 130 135 140

Ile Ser Trp Leu Arg Asn Lys Gly Met Ala Pro Ile Ile Asp Ile Phe
 145 150 155 160

Met Ala Ala Ser Ser Asp Leu Val Asp Ile His Ala Ala Val Leu Phe
 165 170 175

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

Leu His Gly Pro Ala Ala Thr Val Asp Ala Ala Thr Pro Glu Asn Met
 195 200 205

Ala Glu Leu Leu Arg Ile Gly Glu Arg Met Leu Ala Gln Arg Val Ser
 210 215 220

Arg Val Asn Val Glu Thr Gly Arg Tyr Glu Glu Val Lys Gly Ala Gly
 225 230 235 240

Asn Asn Ala Asp Ala Leu Ala Gly Phe Ala Arg Gln Leu Ser Asp Glu
 245 250 255

Arg Arg Thr Arg Leu Gly Ser Arg Arg Gly Gly Ala Gly Arg Leu Lys
 260 265 270

Ser Ser Arg
 275

<210> 79
 <211> 416
 <212> PRT
 <213> Musa acuminata

<400> 79

Met Ala Ser Asp Gln Thr Pro Pro Ala Thr Ala Asp Ala Pro Ile Ser
 1 5 10 15

Thr Pro Pro Pro Ser Phe Gly Lys Arg Val Thr Val Leu Cys Ile Asp
 20 25 30

Gly Gly Gly Val Arg Gly Leu Ile Pro Ala Thr Ile Ile Ala Phe Leu

35

40

45

Glu Ala Glu Leu Gln Lys Leu Asp Gly Pro Glu Ala Arg Ile Ala Asp
 50 55 60

Tyr Phe Asp Val Ile Ala Gly Thr Ser Thr Gly Gly Leu Val Thr Ala
 65 70 75 80

Met Leu Thr Ala Pro Asn Lys Glu Lys Arg Pro Leu Phe Ala Ala Lys
 85 90 95

Asp Ile Val Gln Phe Tyr Leu Asp Asn Ser Pro Lys Ile Phe Pro Gln
 100 105 110

Lys Asn Ala Gly Leu Phe Asn Ser Ala Leu Asn Leu Val Gly Ala Val
 115 120 125

Ser Gly Pro Lys Tyr Asp Gly Lys Tyr Leu His Ala Ile Ile Arg Gln
 130 135 140

Leu Leu Gly Asp Thr Arg Leu Ser Gln Thr Leu Thr Asn Val Val Ile
 145 150 155 160

Pro Thr Phe Asp Ile Lys Phe Leu Gln Pro Thr Ile Phe Ser Thr Tyr
 165 170 175

Gln Thr Lys Ser Thr Pro Leu Lys Asp Ala Leu Leu Ser Asp Ile Cys
 180 185 190

Ile Gly Thr Ser Ala Ala Pro Thr Tyr Leu Pro Gly His Tyr Phe Glu
 195 200 205

Thr Lys Asp Asp Asn Gly Asn Lys Arg Ser Phe Asn Leu Val Asp Gly
 210 215 220

Gly Val Ala Ala Asn Asn Pro Thr Leu Thr Ala Met Thr Glu Val Ser
 225 230 235 240

Lys Glu Ile Leu Leu Ser Asn Pro Asp Phe Phe Ser Tyr Gln Pro Val
 245 250 255

Glu Tyr Asp Arg Phe Leu Val Ile Ser Leu Gly Thr Gly Ala Pro Lys
 260 265 270

Gln Glu Glu Lys Phe Thr Ala Gln Glu Ser Ala Lys Trp Gly Val Leu
 275 280 285

Gly Trp Leu Leu Asn Lys Gly Thr Thr Pro Leu Ile Asp Ile Phe Thr

290 295 300
 Gln Ala Ser Ala Asp Met Val Asp Ile His Ala Ser Val Leu Phe Gln
 305 310 315 320
 Ala Leu Asn Lys Gly Lys Asn Tyr Leu Arg Ile Glu Asp Asp Thr Leu
 325 330 335
 Thr Gly Gln Thr Ser Ser Val Asp Val Ser Thr Lys Lys Asn Leu Gln
 340 345 350
 Asp Leu Val Asp Ile Gly Asn Ser Leu Leu Lys Lys Pro Val Ser Arg
 355 360 365
 Val Asn Ile Glu Thr Gly His Ser Glu Ala Val Asp Gly Glu Gly Thr
 370 375 380
 Asn Glu Ala Ala Leu Thr Gly Phe Ala Lys Lys Leu Ser Asp Glu Lys
 385 390 395 400
 Arg Arg Arg Gln Ser Lys Gln Leu Thr Ser Ser Asp Ala Thr Gln His
 405 410 415

 <210> 80
 <211> 402
 <212> PRT
 <213> Elaeis guineensis

 <400> 80
 Met Gly Ser Asn Glu Ser Pro Val Ala Asn Pro Pro Pro Cys Lys Gly
 1 5 10 15
 Lys Val Val Thr Ile Leu Ser Ile Asp Gly Gly Gly Val Arg Gly Ile
 20 25 30
 Ile Pro Gly Thr Ile Leu Glu Phe Leu Glu Ala Lys Leu Gln Glu Leu
 35 40 45
 Asp Gly Pro Asp Ala Arg Ile Ala Asp Tyr Phe Asp Ile Ile Ala Gly
 50 55 60
 Thr Ser Thr Gly Gly Leu Val Thr Thr Met Leu Thr Ala Pro Asn Lys
 65 70 75 80
 Asp Asn Arg Pro Leu Phe Ser Ala Lys Asp Ile Ile Gln Phe Tyr Leu
 85 90 95
 Glu Asn Cys Pro Lys Ile Phe Pro Gln Arg Thr Gly Leu Leu Ala Gly
 100 105 110

Ala Leu Asn Leu Phe Gly Ala Val Ser Gly Pro Lys Tyr Asp Gly Lys
 115 120 125

Phe Leu His Ser Lys Val Lys Glu Leu Leu Gly Asp Thr Lys Leu His
 130 135 140

Gln Thr Leu Thr Asn Ile Val Ile Pro Ala Phe Asp Ile Lys Leu Leu
 145 150 155 160

Gln Pro Val Ile Phe Ser Thr Phe Glu Thr Lys Thr Asp Pro Ser Lys
 165 170 175

Asp Ala Leu Leu Ser Asp Ile Cys Ile Ser Thr Ser Ala Ala Pro Thr
 180 185 190

Tyr Leu Pro Ala His Tyr Phe Glu Thr Lys Asp Ser Gln Gly Lys Thr
 195 200 205

Arg Ser Phe Asn Leu Val Asp Gly Gly Val Ala Ala Asn Asn Pro Met
 210 215 220

Leu Ile Ala Thr Ser Gln Ile Thr Lys Gln Ile Phe Trp Asn Asn Glu
 225 230 235 240

Asp Phe Ser Lys Phe Lys Pro Thr Asp Phe Ala Lys Phe Leu Val Val
 245 250 255

Ser Leu Gly Thr Gly Ser Pro Lys Gln Glu Gln Lys Phe Thr Ala Pro
 260 265 270

Glu Ser Ala Lys Trp Gly Leu Leu Gly Trp Leu Gln Asn Lys Gly Ser
 275 280 285

Thr Pro Ile Ile Asp Ile Phe Ser Gln Ser Ser Ala Asp Met Val Asp
 290 295 300

Ile His Ala Ser Ile Leu Phe Gln Ala Leu Arg Ser Glu Lys Asn Tyr
 305 310 315 320

Leu Arg Ile Gln Asp Asp Thr Leu Glu Gly Asp Thr Ala Ser Val Asp
 325 330 335

Val Ser Thr Ser Glu Asn Leu Arg Lys Leu Val Gln Val Gly Gln Asp
 340 345 350

Leu Leu Lys Lys Pro Val Ser Arg Val Asn Leu Glu Thr Gly Val Ser
 355 360 365

Glu Ala Cys Asp Val Glu Gly Thr Asn Glu Asp Ala Leu Ile Arg Phe
 370 375 380

Ala Lys Met Leu Ser Asn Glu Arg Lys Ser Arg Asn Ala Lys Met Ser
 385 390 395 400

Ala Ala

<210> 81
 <211> 407
 <212> PRT
 <213> Arabidopsis thaliana

<400> 81

Met Gln Met Asp Ser Pro Lys Ser Pro Leu Gln Pro Pro Thr Tyr Gly
 1 5 10 15

Asn Leu Val Thr Ile Leu Ser Ile Asp Gly Gly Gly Ile Arg Gly Leu
 20 25 30

Ile Pro Ala Val Ile Leu Gly Phe Leu Glu Ser Glu Leu Gln Lys Leu
 35 40 45

Asp Gly Glu Glu Ala Arg Leu Ala Asp Tyr Phe Asp Val Ile Ala Gly
 50 55 60

Thr Ser Thr Gly Gly Leu Val Thr Ala Met Leu Thr Ala Pro Asn Lys
 65 70 75 80

Glu Gly Arg Pro Leu Phe Ala Ala Ser Glu Ile Lys Asp Phe Tyr Leu
 85 90 95

Glu Gln Cys Pro Lys Ile Phe Pro Gln Asp His Phe Pro Phe Ser Ala
 100 105 110

Ala Lys Lys Leu Val Lys Ser Leu Thr Gly Pro Lys Tyr Asp Gly Lys
 115 120 125

Tyr Leu His Gln Leu Ile His Ala Lys Leu Gly Asp Thr Lys Leu Ser
 130 135 140

Gln Thr Leu Thr Asn Val Val Ile Pro Thr Phe Asp Ile Lys His Leu
 145 150 155 160

Gln Pro Thr Ile Phe Ser Ser Tyr Glu Val Lys Asn His Pro Leu Lys
 165 170 175

Asp Ala Ala Leu Ala Asp Ile Ala Ile Ser Thr Ser Ala Ala Pro Thr
 180 185 190

Tyr Leu Pro Ala His Phe Phe Lys Val Glu Asp Leu Asn Gly Asn Ala
 195 200 205

Lys Glu Tyr Asn Leu Ile Asp Gly Gly Val Ala Ala Asn Asn Pro Ala
 210 215 220

Leu Leu Ala Ile Gly Glu Val Thr Asn Glu Ile Ser Gly Gly Ser Ser
 225 230 235 240

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

Ser Leu Gly Thr Gly Asn His Lys Ala Glu Glu Lys Phe Asn Ala Lys
 260 265 270

Glu Val Ala Gly Trp Gly Leu Leu Asn Trp Leu Thr His Asp Asn Ser
 275 280 285

Thr Pro Ile Ile Asp Ala Phe Ser Gln Ala Ser Ser Asp Met Val Asp
 290 295 300

Phe His Leu Ser Ala Val Phe Arg Ala Leu His Ser Glu Ala Asn Tyr
 305 310 315 320

Ile Arg Ile Gln Asp Asp Thr Leu Thr Gly Asp Ala Ala Ser Val Asp
 325 330 335

Ile Ala Thr Val Glu Asn Leu Asp Ile Leu Ala Lys Thr Gly Asp Glu
 340 345 350

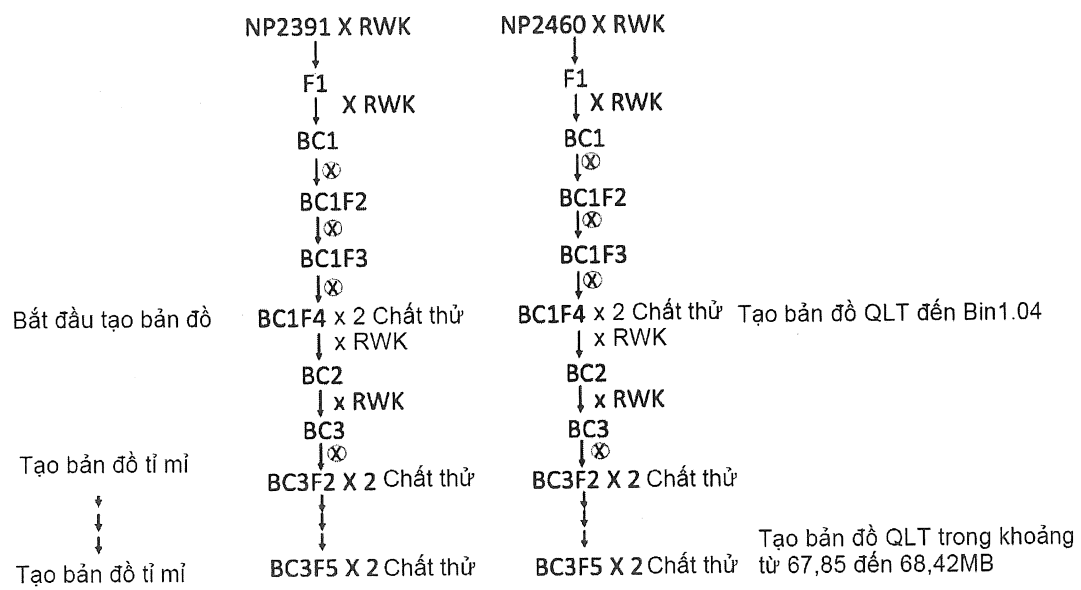
Leu Leu Lys Lys Pro Val Ala Arg Val Asn Leu Asp Ser Gly Cys Asn
 355 360 365

Glu Asn Ala Tyr Glu Thr Thr Asn Glu His Ala Leu Ile Lys Leu Ala
 370 375 380

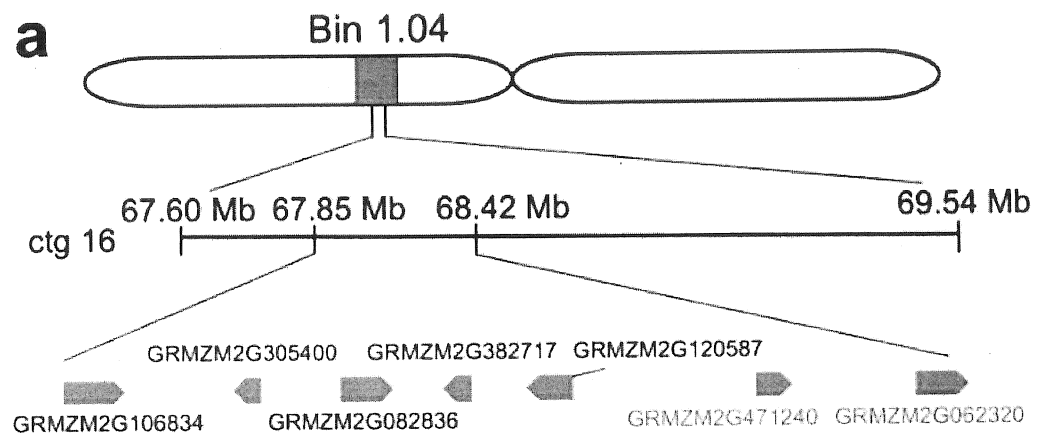
Gly Ile Leu Ser Lys Glu Lys Lys Ile Arg Asp Ile Arg Ser Pro His
 385 390 395 400

Ala Lys Ala Pro Ile Arg Ile
 405

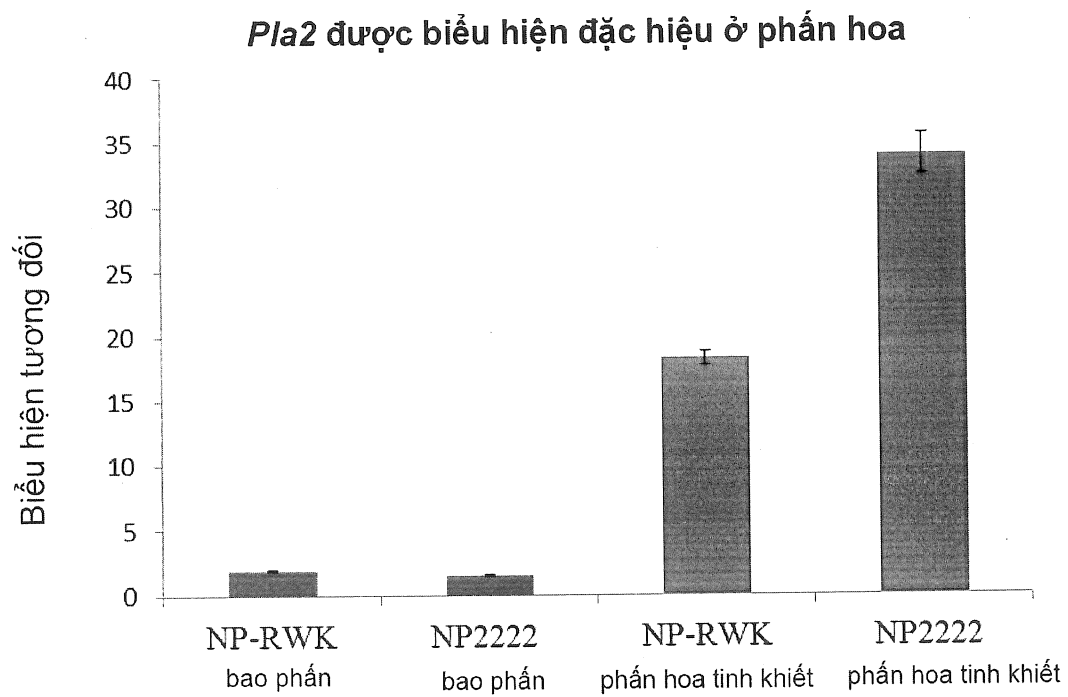
HÌNH 1.



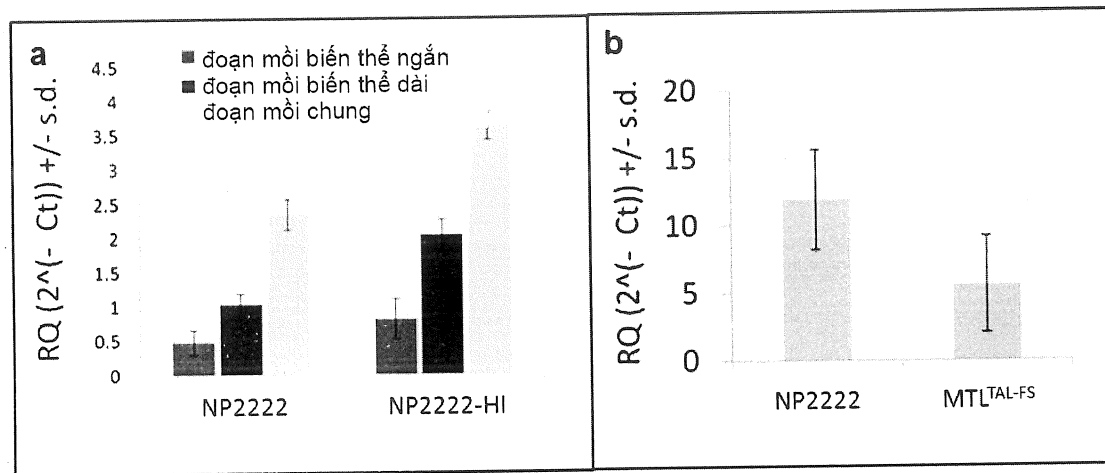
HÌNH 2.



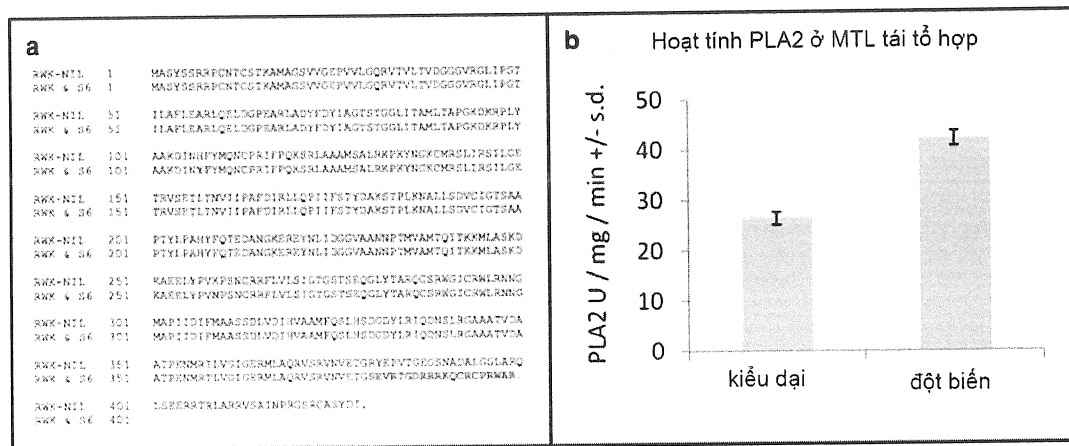
HÌNH 3.



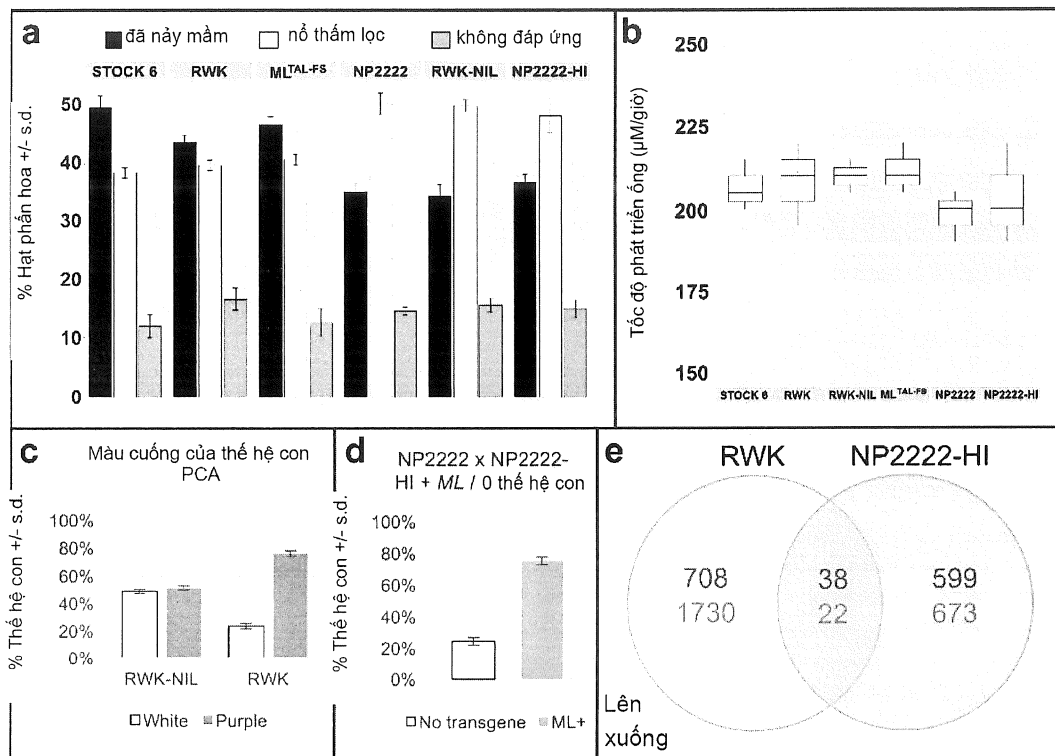
HÌNH 4.



HÌNH 5.



HÌNH 6.

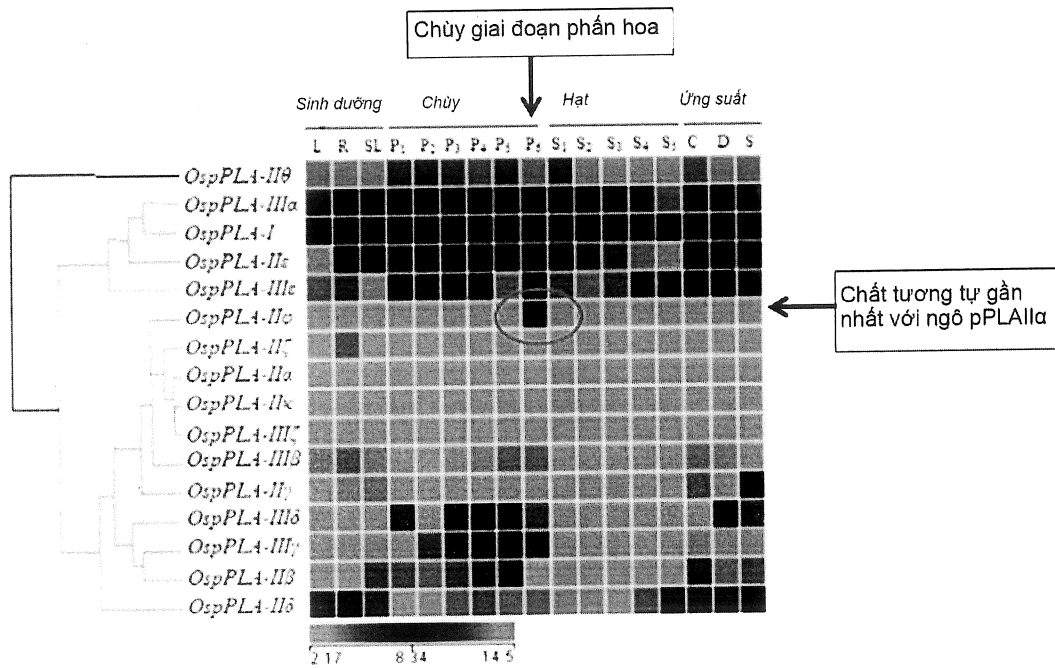


7/31

HÌNH 7.

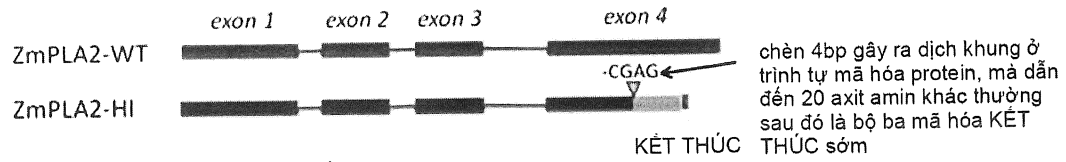
<i>Zea mays</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Sorghum bicolor</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Setaria italica</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Hordeum vulgare</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Brachypodium distachyon</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Oryza sativa v. japonica</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Oryza sativa v. indica</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Triticum aestivum</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Musa acuminata</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Eleais guineensis</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Arabidopsis thaliana</i>	MAA-YSSRRPCNTCSTWAMAGSVVGEFFV-VLQGRVTVLTVDDGGGRLIPSTILAFLEARLQELDGEARLADYFDYAGTST
<i>Zea mays</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Sorghum bicolor</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Setaria italica</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Hordeum vulgare</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Brachypodium distachyon</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Oryza sativa v. indica</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Oryza sativa v. japonica</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Triticum aestivum</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Musa acuminata</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Eleais guineensis</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Arabidopsis thaliana</i>	GGITAMLTAFQNDKSPYAAADINRFYMNCRPIFFQK----SSLAAMMSALKKFFYNGKRLNIMSMLGSTRVSDILNVIIFTFD
<i>Zea mays</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Sorghum bicolor</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Setaria italica</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Hordeum vulgare</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Brachypodium distachyon</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Oryza sativa v. indica</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Oryza sativa v. japonica</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Triticum aestivum</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Musa acuminata</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Eleais guineensis</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Arabidopsis thaliana</i>	IRLLQPIESTYDANMELNALLSDVCTGSAAPTYLPAHYPTEDA-NHKEFSYNLIDGVAANNFTWVANTQITKMLAKGMAEEL
<i>Zea mays</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Sorghum bicolor</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Setaria italica</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Hordeum vulgare</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Brachypodium distachyon</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Oryza sativa v. indica</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Oryza sativa v. japonica</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Triticum aestivum</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Musa acuminata</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Eleais guineensis</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Arabidopsis thaliana</i>	YVNFEDDGRFLVLSIGTSTEDQWYTAQCSQNGIHWLNKMGMAPIIDIHMAASSDLVDIHAAYVPSLHSDGN-YLRIQNSLRGA
<i>Zea mays</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Sorghum bicolor</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Setaria italica</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Hordeum vulgare</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Brachypodium distachyon</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Oryza sativa v. indica</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Oryza sativa v. japonica</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Triticum aestivum</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Musa acuminata</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Eleais guineensis</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI
<i>Arabidopsis thaliana</i>	NATVDAATFENMRLVIGGERMLAQRVSRVNVETGRYEVPTGEGSNA DALAGARQLSEERPARLARNVSAINTFOS--ACAS--YDI

HÌNH 8.

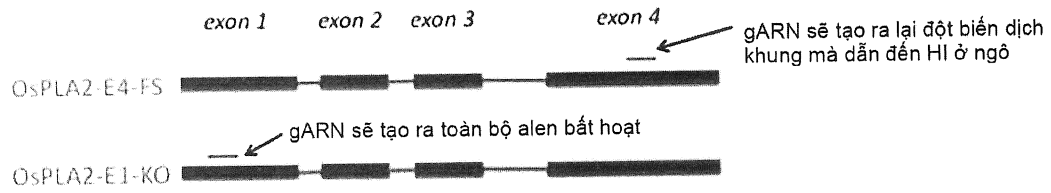


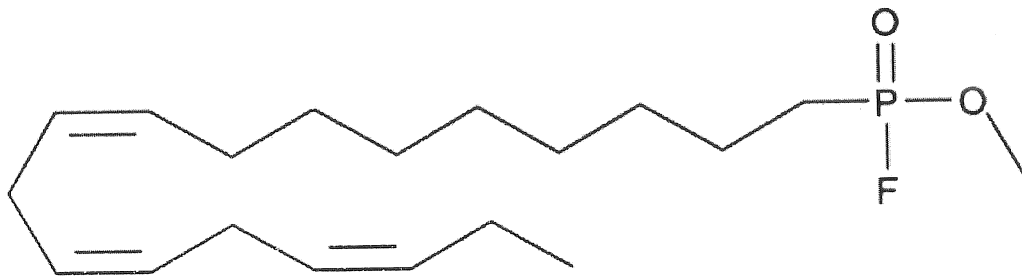
HÌNH 9.

ĐỘT BIẾN BẨM SINH Ở NGÔ: Kích tạo đơn bội do dịch khung ở exon thứ 4 của PLA2

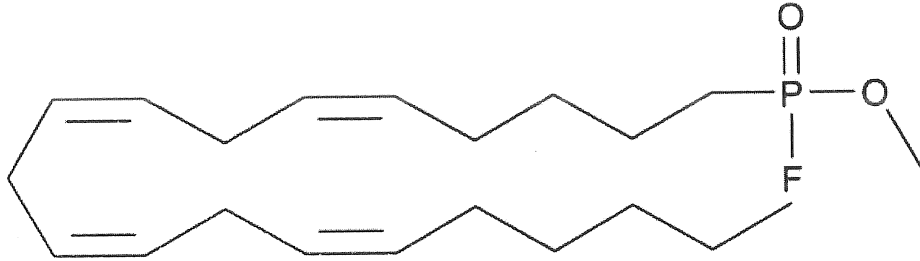


Chỉnh sửa OsPLA2 ở exon thứ nhất và thứ tư

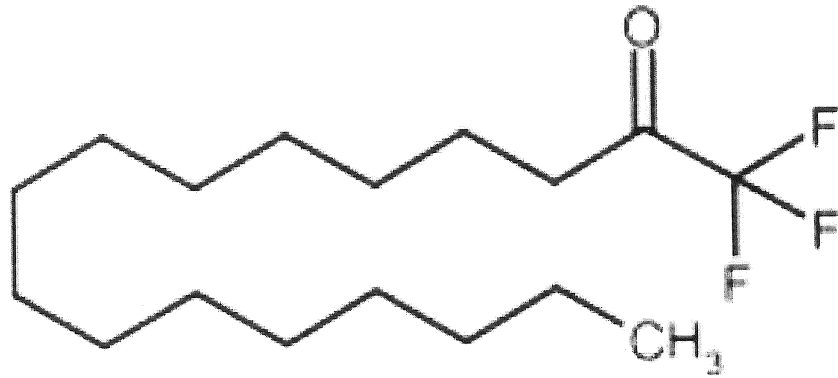




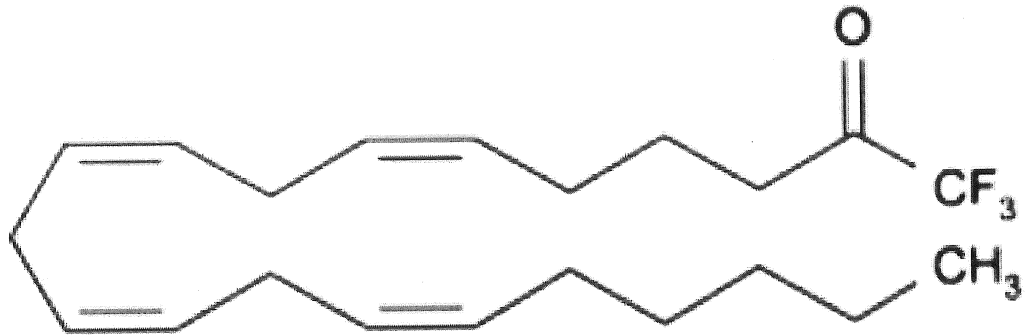
HÌNH 11.



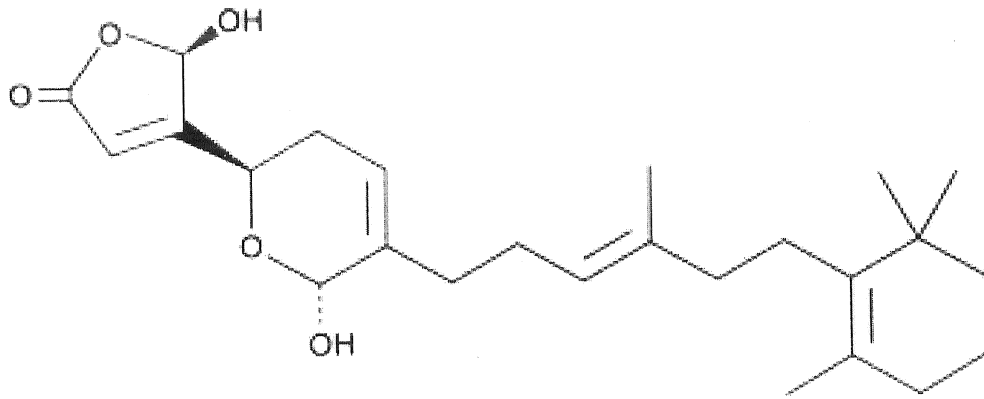
HÌNH 12



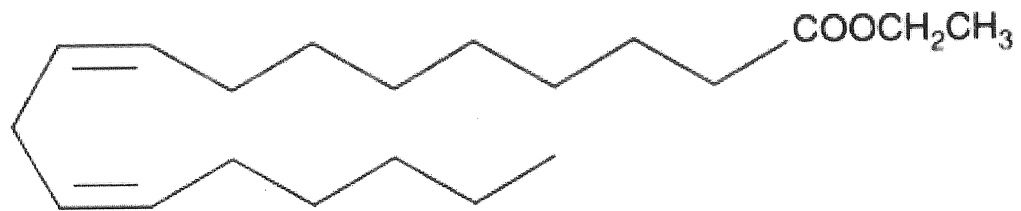
HÌNH 13.



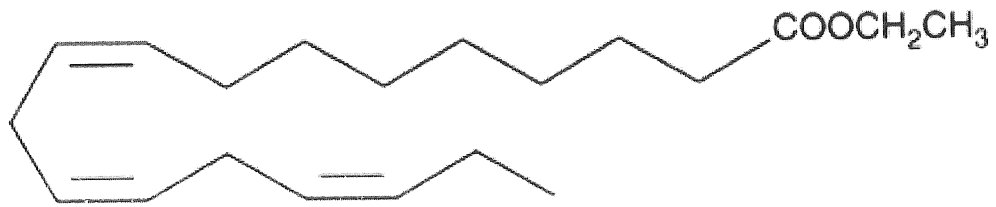
HÌNH 14.



HÌNH 15.

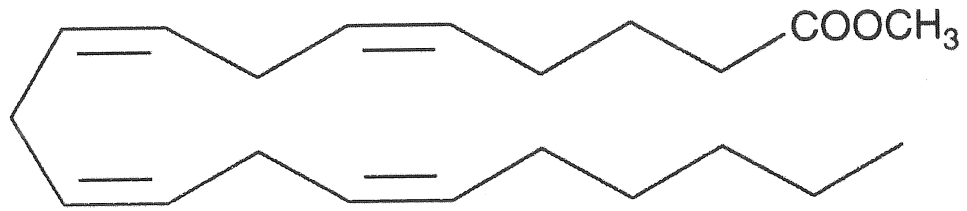


HÌNH 16.

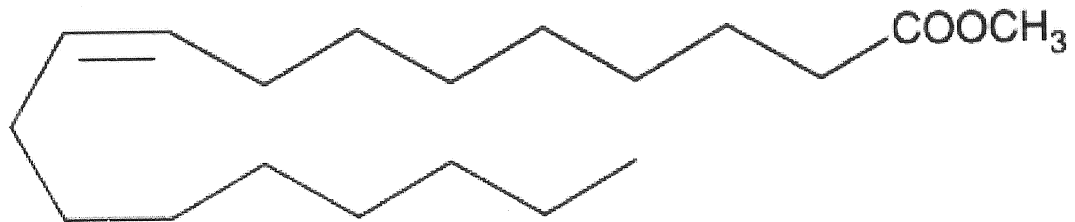


17/31

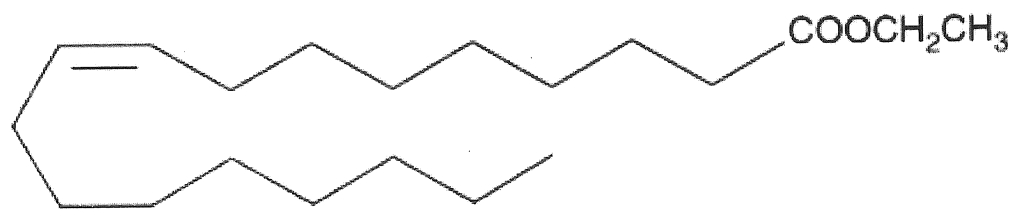
HÌNH 17.



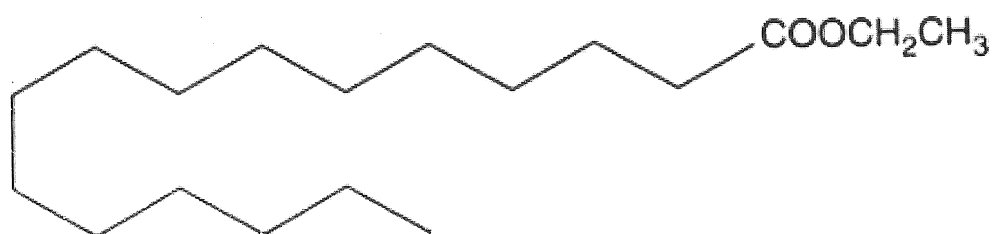
HÌNH 18.



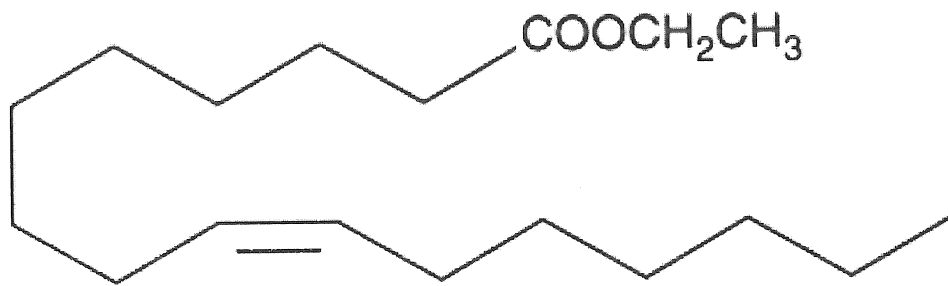
HÌNH 19.



HÌNH 20.

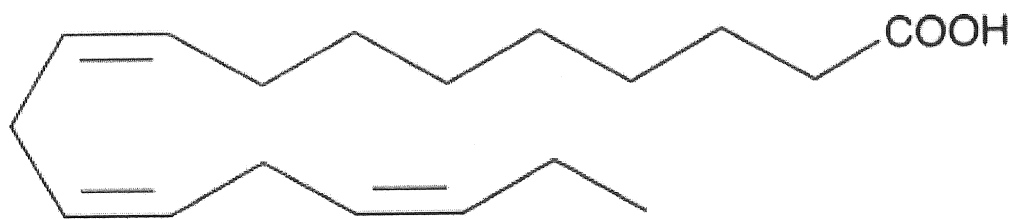


HÌNH 21.

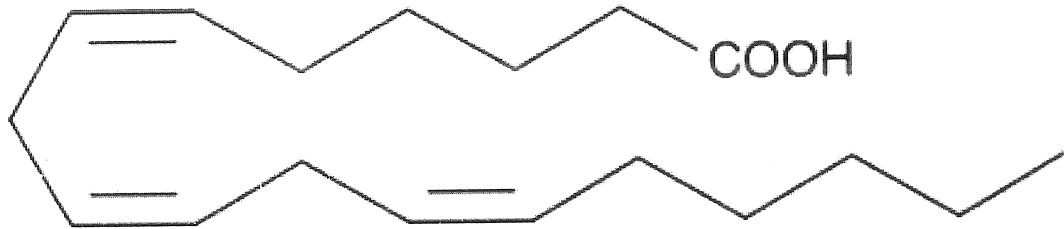


22/31

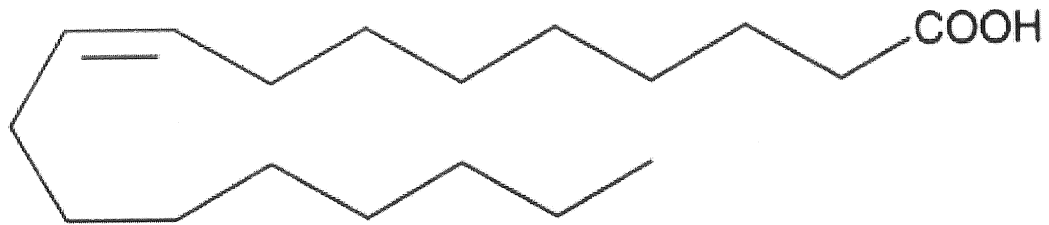
HÌNH 22.



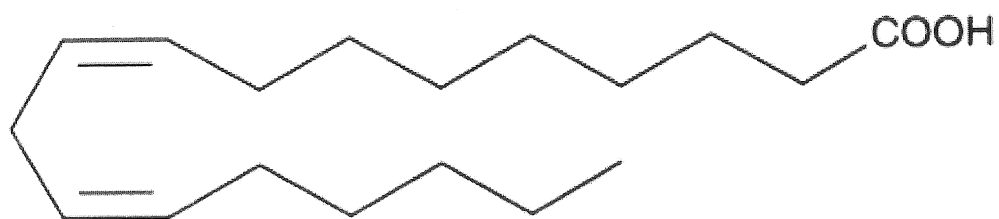
HÌNH 23.



HÌNH 24.

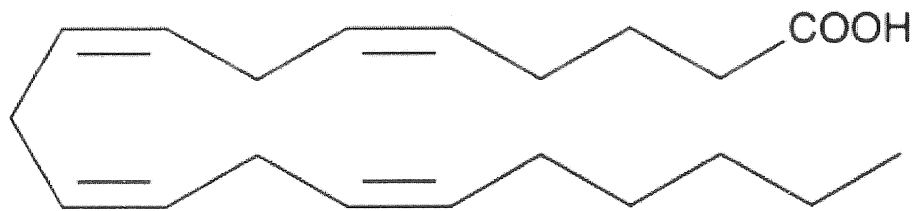


HÌNH 25.



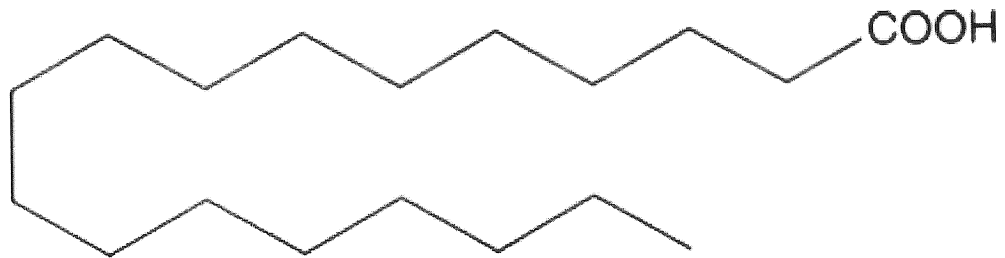
26/31

HÌNH 26.

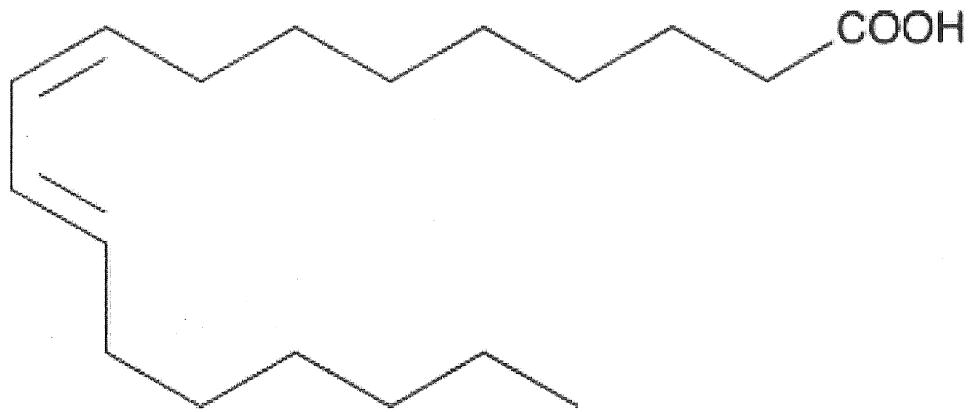


27/31

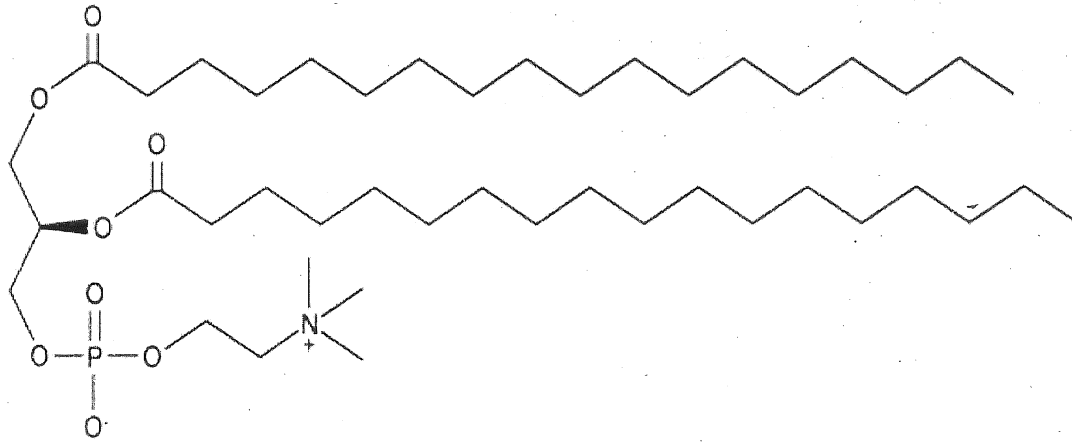
HÌNH 27.



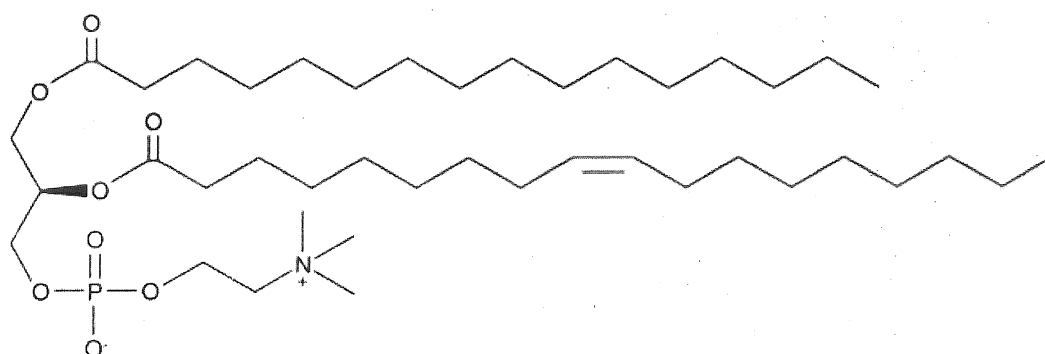
HÌNH 28.



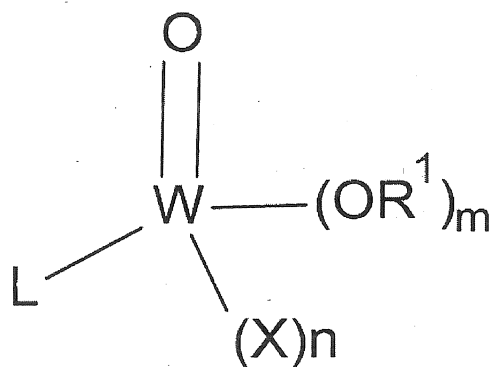
HÌNH 29.



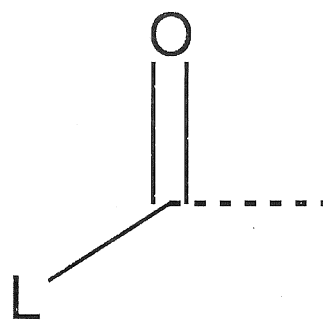
HÌNH 30.



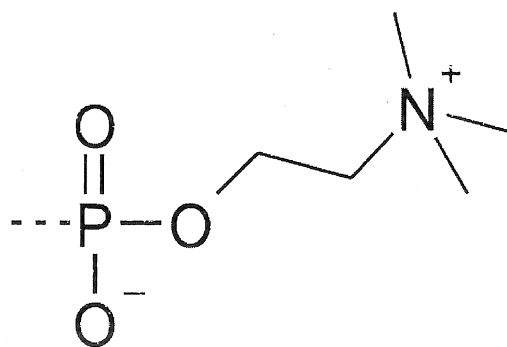
HÌNH 31.



a.



b.



c.