



(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)



1-0051315

(51)^{2022.01} C12N 15/67; C12N 15/82

(13) B

(21) 1-2023-01609

(22) 27/09/2021

(86) PCT/US2021/052131 27/09/2021

(87) WO 2022/067152 31/03/2022

(30) 63/084,150 28/09/2020 US; 17/249,437 02/03/2021 US

(45) 25/09/2025 450

(43) 25/09/2023 426A

(73) INARI AGRICULTURE TECHNOLOGY, INC. (US)

One Kendall Square, Building 600/700, Suite 7-501, Cambridge, Massachusetts
02139, United States of America

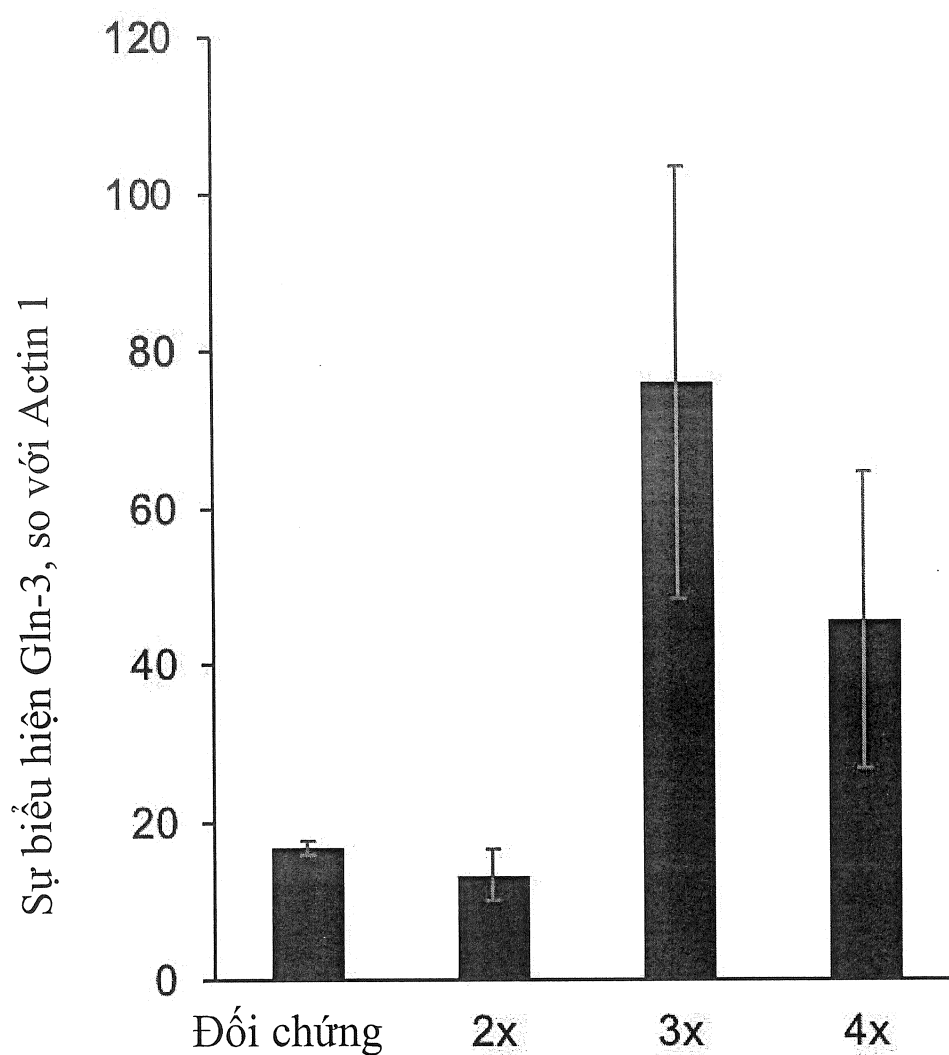
(72) CLAEYS, Hannes Bart (BE).

(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) PHÂN TỬ ADN, TẾ BÀO THỰC VẬT, CÂY, BỘ PHẬN CỦA CÂY BAO GỒM
PHÂN TỬ ADN NÀY, VÀ PHƯƠNG PHÁP SẢN XUẤT VẬT LIỆU THỰC VẬT

(21) 1-2023-01609

(57) Yếu tố làm tăng sự biểu hiện bao gồm trình tự ADN 36 nucleotit mà có thể được sử dụng để làm tăng sự biểu hiện của các gen, và cụ thể là để làm tăng sự biểu hiện của các gen thực vật nội sinh, ở các cây được bộc lộ. Sáng chế cũng đề cập đến các cây, bộ phận của cây, và sản phẩm thực vật hàng hóa bao gồm yếu tố tăng cường sự biểu hiện cùng với các phương pháp có liên quan sử dụng yếu tố làm tăng sự biểu hiện và cây bao gồm yếu tố làm tăng sự biểu hiện này.



Hình 2

THAM CHIẾU ĐẾN DANH MỤC TRÌNH TỰ ĐƯỢC NỘP THEO HÌNH THỨC ĐIỆN TỬ

[0001] Danh mục trình tự được chứa trong tệp có tên là "10074WO1_ST25.txt" mà được nộp theo hình thức điện tử với đơn này được kết hợp toàn bộ ở đây bằng cách viện dẫn.

LĨNH VỰC KỸ THUẬT ĐƯỢC ĐỀ CẬP

[0002] Sáng chế đề cập đến trình tự ADN có thể hoạt động dưới dạng yếu tố làm tăng biểu hiện của gen ở các thực vật chuyển gen và được chỉnh sửa bộ gen. Sáng chế còn đề cập đến cây, bộ phận của cây, và sản phẩm thực vật hàng hóa chứa yếu tố làm tăng biểu hiện của gen cùng với các phương pháp liên quan để sử dụng yếu tố làm tăng biểu hiện của gen và cây chứa yếu tố làm tăng biểu hiện của gen này.

TÌNH TRẠNG KỸ THUẬT CỦA SÁNG CHẾ

[0003] Phương pháp sử dụng công nghệ CRISPR, Nucleaza ngón tay kềm, và Nucleaza yếu tố tác động giống yếu tố kích hoạt phiên mã (TALEN) để chỉnh sửa bộ gen ở thực vật được bộc lộ trong các tài liệu US 20150082478, US 2015/0059010A1, và Bortesi et al., 2015, Biotechnology Advances, pp. 41-52, Vol. 33, No. 1. Ellis et al., 1987, EMBO J. (6):11:3203-3208, bộc lộ yếu tố tăng cường gen octopin synthaza của vi khuẩn gồm 16 cặp bazơ mà có thể làm tăng sự biểu hiện của các gen ngoại sinh ở các tế bào trần của ngô và thuốc lá ở các thử nghiệm biểu hiện nhất thời. Đơn sáng chế PCT WO 2018/140899 bộc lộ sự chèn các yếu tố tăng cường sự biểu hiện có độ đồng nhất với yếu tố tăng cường gen octopin synthaza của vi khuẩn ở vùng trình tự khởi động của gen Lc ngô ở bộ gen tế bào trần của ngô để làm tăng sự biểu hiện của gen đó. Vẫn đang cần có các trình tự thu được từ thực vật mà có khả năng làm tăng đáng kể sự phiên mã của các gen thực vật.

BẢN CHẤT KỸ THUẬT CỦA SÁNG CHẾ

[0004] Các phân tử ADN bao gồm trình tự polynucleotit của SEQ ID NO:3 được đề xuất. Cũng được đề xuất là các mẫu sinh học, các nhiễm sắc thể thực vật, tế bào thực

vật, thể nuôi cấy mô của các tế bào có thể tái sinh được bao gồm các tế bào thực vật, các bộ phận của cây, và cây bao gồm trình tự polynucleotit của SEQ ID NO:3. Theo các phương án nhất định đã nêu trên, trình tự polynucleotit của SEQ ID NO:3 được liên kết hoạt động với trình tự polynucleotit bao gồm trình tự khởi động, trong đó trình tự khởi động này tùy ý là trình tự khởi động nội sinh, và trong đó trình tự khởi động nội sinh là tùy ý được định vị trong nhiễm sắc thể của thực vật. Cũng được đề xuất là việc sử dụng các phân tử ADN đã nêu trên bao gồm trình tự polynucleotit của SEQ ID NO:3 để: (i) làm tăng sự biểu hiện của một hoặc nhiều yếu tố được mã hóa bởi đơn vị phiên mã mà được liên kết hoạt động với trình tự khởi động ở thực vật; (ii) truyền tính trạng hữu dụng cho cây bao gồm phân tử ADN tái tổ hợp, trong đó tính trạng hữu dụng này là khả năng chống chịu áp lực phi sinh học, cấu trúc, áp lực sinh học, khả năng quang hợp, hoặc phân vùng tài nguyên tùy ý được cải thiện so với cây đối chứng thiếu phân tử ADN tái tổ hợp; (iii) thu cây hoặc hạt của cây thể hiện tính trạng hữu dụng của mục (ii); hoặc (iv) trồng quần thể thực vật thể hiện tính trạng hữu dụng của mục (ii); tùy ý trong đó cây của mục (i), (ii), (iii), hoặc (iv) là cây ngô, đậu tương, bông, hoặc cải dầu; hoặc tùy ý trong đó hạt của mục (iii) là cây ngô, đậu tương, bông, hoặc cải dầu.

[0005] Sáng chế đề xuất phương pháp sản xuất hạt thực vật bao gồm bước lai các cây chứa phân tử ADN đã nêu bao gồm trình tự polynucleotit của SEQ ID NO:3 với cây thứ hai để tạo ra hạt thực vật và tùy ý thu hoạch hạt này.

[0006] Sáng chế đề xuất phương pháp sản xuất hạt thực vật bao gồm bước tự phối cây chứa phân tử ADN đã nêu bao gồm trình tự polynucleotit của SEQ ID NO:3 để tạo ra hạt thực vật và tùy ý thu hoạch hạt này.

[0007] Sáng chế đề xuất phương pháp sản xuất cây chứa tính trạng mong muốn được bổ sung bao gồm bước đưa gen chuyển, sự thay đổi di truyền được hướng đích, hoặc locus gen mang tính trạng mong muốn vào trong cây bao gồm các phân tử ADN đã nêu bao gồm trình tự polynucleotit của SEQ ID NO:3.

[0008] Sáng chế đề xuất phương pháp sản xuất sản phẩm thực vật hàng hóa bao gồm bước xử lý cây hoặc hạt chứa các phân tử ADN đã nêu bao gồm trình tự polynucleotit của SEQ ID NO:3 và thu hồi sản phẩm thực vật hàng hóa từ cây hoặc hạt đã xử lý.

[0009] Sáng chế đề xuất phương pháp sản xuất vật liệu thực vật, phương pháp này bao gồm bước trồng cây có yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3 được liên

kết hoạt động với polynucleotit mã hóa bản phiên mã, trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã ở cây đã nêu được tăng khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện.

[0010] Phương pháp sản xuất vật liệu thực vật này bao gồm bước: (a) cung cấp cây có yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã, trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã được tăng ở cây đã nêu khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện; và (b) trồng cây dưới các điều kiện mà cho phép đối với sự biểu hiện của polynucleotit thúc đẩy bản phiên mã được đề xuất.

[0011] Sáng chế đề xuất phương pháp xác định các mẫu sinh học bao gồm polynucleotit bao gồm gen thực vật được cải biến bao gồm bước phát hiện sự biểu hiện của SEQ ID NO:3 ở mẫu sinh học.

[0012] Sáng chế đề xuất phương pháp sản xuất hạt thực vật được xử lý bao gồm bước cho hạt bao gồm các phân tử ADN đã nêu bao gồm trình tự polynucleotit của SEQ ID NO:3 tiếp xúc với chế phẩm chứa tác nhân sinh học, chất diệt côn trùng, hoặc chất diệt nấm.

[0013] Phương pháp làm tăng sự biểu hiện của phân tử ARN ở thực vật bao gồm bước cho phân tử ARN được mã hóa bởi phân tử ADN biểu hiện ở thực vật, trong đó phân tử ADN mã hóa phân tử ARN được liên kết hoạt động với một hoặc nhiều phân tử ADN bao gồm: (i) yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3; (ii) trình tự khởi động; và tùy ý (iii) các phân tử ADN mã hóa vùng không được dịch mã đầu 5' (5' UTR), intron, exon, 3' UTR, vị trí polyadenyl hóa, hoặc tổ hợp của chúng; trong đó sự biểu hiện của ARN được tăng khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện được đề xuất.

MÔ TẢ VĂN TẮT CÁC HÌNH VẼ

[0014] Hình 1 mô tả hiệu quả chèn của dime (SEQ ID NO:2), trime (SEQ ID NO:3), hoặc tetrame (SEQ ID NO:4) của yếu tố lõi 12 nucleotit (SEQ ID NO:1) được chèn vào trong trình tự khởi động ZmGln1-3.

[0015] Hình 2 mô tả tác dụng của dime (SEQ ID NO:2), trime (SEQ ID NO:3), hoặc tetrame (SEQ ID NO:4) của trình tự 12 nucleotit yếu tố lõi (SEQ ID NO:1) được chèn

vào trong trình tự khởi động ZmGln1-3 đối với sự biểu hiện ZmGln1-3, khi so sánh với đối chứng.

[0016] Hình 3 mô tả sự biểu hiện của Lc (bảng thứ nhất), Gln1-3 (bảng ở giữa) và Gln1-4 (bảng bên phải) ở các tế bào trần với sự chèn của yếu tố tăng cường 36 bp (SEQ ID NO:3) trong các trình tự khởi động này (các thanh màu đen), khi so sánh với sự biểu hiện ở tế bào trần mà không có yếu tố tăng cường 36 bp (các thanh màu trắng). Ba lần lặp sinh học cho mỗi loại được thể hiện riêng rẽ.

[0017] Hình 4 mô tả sự biểu hiện của Gln1-3 (so với ZmActin1) ở các sự kiện T0 với sự chèn yếu tố tăng cường 36 bp (SEQ ID NO:3) ở trình tự khởi động Gln1-3 (thanh màu đen), khi so sánh với sự biểu hiện ở sự kiện T0 mà không có sự chèn yếu tố tăng cường hoặc đối chứng B104 kiểu dại (các thanh màu trắng).

[0018] Hình 5 mô tả sự biểu hiện của Gln1-3 (so với ZmAct1) ở các cá thể T1 với sự chèn yếu tố tăng cường 36 bp (SEQ ID NO:3) trong trình tự khởi động Gln1-3 khi so sánh với các cá thể cùng máu mà không chèn yếu tố tăng cường hoặc các kiểu dại, ở rễ chính, lông sơ khởi, và rễ.

MÔ TẢ CHI TIẾT SÁNG CHẾ

[0019] Cụm từ "biến thể alen" như được sử dụng ở đây để chỉ biến thể polynucleotit hoặc trình tự polypeptit mà xuất hiện ở chủng, giống, hoặc dòng phân lập khác của sinh vật đã cho.

[0020] Thuật ngữ "và/hoặc" khi được sử dụng ở đây được lấy là sự bộc lộ cụ thể của mỗi dấu hiệu hoặc thành phần trong số hai dấu hiệu hoặc thành phần đã chỉ rõ mà có hoặc không có dấu hiệu hoặc thành phần khác. Do đó, thuật ngữ "và/hoặc" khi được sử dụng trong cụm từ như "A và/hoặc B" ở đây được nhằm để bao gồm "A và B," "A hoặc B," "A" (một mình), và "B" (một mình). Tương tự, thuật ngữ "và/hoặc" như được sử dụng trong cụm từ như "A, B, và/hoặc C" được nhằm để bao hàm từng phương án trong số các phương án sau đây: A, B, và C; A, B, hoặc C; A hoặc C; A hoặc B; B hoặc C; A và C; A và B; B và C; A (một mình); B (một mình); và C (một mình).

[0021] Như được sử dụng ở đây, cụm từ "mẫu sinh học" để chỉ mô thực vật nguyên vẹn hoặc không nguyên vẹn (ví dụ, hạt hoặc mô thực vật được nghiền, mô thực vật được băm nhỏ, mô được đông khô). Nó cũng có thể là dịch chiết bao gồm mô thực vật hoặc

hạt nguyên vẹn hoặc không nguyên vẹn. Mẫu sinh học có thể bao gồm bột, bột thô, phiên, sirô, dầu, tinh bột, và ngũ cốc được sản xuất nguyên phần hoặc từng phần để chứa phụ phẩm của cây trồng. Theo các phương án nhất định, mẫu sinh học là "không thể tái sinh được" (*tức là*, không có khả năng để tái sinh thành cây hoặc bộ phận của cây).

[0022] Như được sử dụng ở đây, các cụm từ "trình tự khởi động nội sinh," "gen nội sinh," "đơn vị phiên mã thực vật nội sinh" và các dạng tương tự để chỉ dạng bản sinh của trình tự khởi động, gen, hoặc đơn vị phiên mã thực vật ở vị trí tự nhiên của nó ở sinh vật hoặc ở bộ gen của sinh vật.

[0023] Thuật ngữ "ngoại sinh" như được sử dụng ở đây liên quan đến phân tử ADN, nucleotit, hoặc polynucleotit được chèn vào trong bộ gen thực vật để chỉ phân tử ADN, nucleotit, hoặc polynucleotit bất kỳ mà tổng hợp hoặc đã được loại bỏ khỏi vị trí bản sinh của nó và đã được chèn vào trong vị trí bộ gen mới.

[0024] Thuật ngữ "được phân lập" như được sử dụng ở đây có nghĩa là đã được loại bỏ khỏi môi trường tự nhiên của nó.

[0025] Như được sử dụng ở đây, thuật ngữ "bao gồm" được hiểu là ít nhất có các dấu hiệu mà chúng đề cập đến trong khi không loại trừ bất kỳ các dấu hiệu chưa được chỉ rõ khác.

[0026] Như được sử dụng ở đây, cụm từ "được liên kết hoạt động" để chỉ vị trí kề nhau trong đó các thành phần nên được mô tả là trong mối quan hệ cho phép chúng hoạt động theo cách được dự tính của chúng. Ví dụ, trình tự khởi động được liên kết hoạt động với trình tự mã hóa nếu trình tự khởi động tác động đến sự phiên mã hoặc sự biểu hiện của nó. Ở ví dụ không làm giới hạn khác, "yếu tố làm tăng sự biểu hiện" (ví dụ, yếu tố tăng cường phiên mã) được liên kết hoạt động với trình tự khởi động nếu yếu tố làm tăng sự biểu hiện này làm tăng hoạt tính của trình tự khởi động (ví dụ, như được đo bằng sự tích lũy được điều khiển bởi trình tự khởi động của bản phiên mã hoặc protein được mã hóa bởi bản phiên mã).

[0027] Như được sử dụng ở đây, thuật ngữ "thuộc thể tương đồng khác loài" hoặc "thể tương đồng khác loài" được sử dụng để mô tả các gen hoặc protein được mã hóa bởi các gen đó mà thu được từ các loài khác nhau nhưng có chức năng giống nhau (*ví dụ*, mã hóa enzym mà xúc tác cùng các phản ứng hoặc mã hóa các yếu tố điều hòa phiên mã mà kiểm soát sự biểu hiện gen với các chức năng tương tự). Các gen tương đồng khác loài

sẽ thông thường mã hóa các protein với mức độ tương đồng trình tự nào đó (*ví dụ*, tương đồng trình tự ít nhất 40%, 50%, 60%, 70%, 80%, 90%, hoặc 95%, sự bảo toàn các môtip trình tự, và/hoặc sự bảo toàn các đặc điểm cấu trúc).

[0028] Như được sử dụng ở đây, thuật ngữ "thực vật" bao gồm toàn bộ cây và thể hệ con, tế bào, mô, hoặc bộ phận bất kỳ của cây. Thuật ngữ "các bộ phận của cây" bao gồm (các) bộ phận bất kỳ của cây, bao gồm, ví dụ nhưng không bị giới hạn ở: hạt (bao gồm hạt trưởng thành và hạt chưa trưởng thành); phần cắt của cây; tế bào thực vật; thể nuôi cấy tế bào thực vật; hoặc cơ quan của cây (ví dụ, hạt phấn, phôi, hoa, quả, chồi, lá, rễ, thân, và mô cây). Mô thực vật hoặc cơ quan của cây có thể là hạt, tế bào trần, mô sẹo, hoặc nhóm tế bào thực vật bất kỳ khác mà được sắp xếp thành đơn vị cấu trúc hoặc chức năng. Thể nuôi cấy tế bào hoặc mô thực vật có thể có khả năng tái sinh thành cây có các đặc điểm sinh lý và hình thái học của cây mà tế bào hoặc mô thu được từ đó, và có khả năng tái sinh thành cây về cơ bản có cùng kiểu gen như cây này. Các tế bào có thể tái sinh trong nuôi cấy tế bào hoặc mô thực vật có thể là phôi, tế bào trần, tế bào mô phân sinh, mô sẹo, hạt phấn, lá, bao phấn, rễ, đầu rễ, tơ, hoa, nhân, bông, lõi, lá bao, hoặc thân. Ngược lại, một số thực vật thực vật không có khả năng được tái sinh để tạo thành cây và được gọi ở đây là tế bào thực vật "không thể tái sinh được".

[0029] Thuật ngữ "được tinh chế," như được sử dụng ở đây xác định sự phân tách phân tử hoặc hợp chất ở dạng mà về cơ bản không chứa các tạp chất thường có liên quan đến phân tử hoặc hợp chất này ở môi trường bẩm sinh hoặc tự nhiên và có nghĩa là đã được làm tăng độ tinh khiết là kết quả của việc tách từ các bộ phận khác của chế phẩm gốc. Thuật ngữ "phân tử ADN được tinh chế" được sử dụng ở đây để mô tả trình tự axit nucleic mà đã được tách từ các hợp chất khác bao gồm, nhưng không bị giới hạn ở các polypeptit, lipid và carbohydrat.

[0030] Trong phạm vi mà bất kỳ định nghĩa trong số các định nghĩa đã nêu trên không thống nhất với các định nghĩa được đưa ra ở tài liệu tham khảo dạng sáng chế hoặc phi sáng chế bất kỳ được kết hợp ở đây bằng cách viện dẫn, tài liệu tham khảo dạng sáng chế hoặc phi sáng chế bất kỳ được trích dẫn ở đây, hoặc ở tài liệu tham khảo dạng sáng chế hoặc phi sáng chế bất kỳ được tìm thấy ở đâu đó, cần hiểu rằng các định nghĩa đã nêu sẽ được sử dụng ở đây.

[0031] Yếu tố ADN làm tăng sự biểu hiện mà có thể được sử dụng để làm tăng sự biểu hiện của các polynucleotit mã hóa bản phiên mã ở thực vật bao gồm polynucleotit 5'-GTAAGCGCTTACGTAAGCGCTTACGTAAGCGCTTAC-3' (SEQ ID NO:3) được bộc lộ. Sáng chế cũng bộc lộ cây, bộ phận của cây (ví dụ, hạt, lá, rễ, thân), và các tế bào thực vật có khả năng tái sinh hoặc không có khả năng tái sinh bao gồm SEQ ID NO:3. Các mẫu sinh học bao gồm bột thô của hạt mà bao gồm các polynucleotit bao gồm SEQ ID NO:3 cũng được bộc lộ. Các phương pháp có liên quan để sử dụng cây, bộ phận của cây, và tế bào thực vật bao gồm SEQ ID NO:3 để sản xuất vật liệu thực vật bao gồm hạt hoặc để sản xuất sản phẩm hàng hóa thực vật bao gồm SEQ ID NO:3 cũng được đề xuất.

[0032] Sự biểu hiện của gen thực vật nội sinh có thể được làm tăng bằng cách chèn hoặc sự tạo thành của SEQ ID NO:3 trong gen thực vật nội sinh sao cho SEQ ID NO:3 được liên kết hoạt động với trình tự khởi động nội sinh của gen thực vật. Theo các phương án nhất định, sự liên kết hoạt động với trình tự khởi động nội sinh đạt được bằng việc chèn hoặc sự tạo thành của một hoặc nhiều bản sao của SEQ ID NO:3 ở một hoặc nhiều loại trong số trình tự khởi động nội sinh, vùng không được dịch mã đầu 5' (5'UTR), intron, và/hoặc vùng không được dịch mã đầu 3' của gen thực vật nội sinh. Vì SEQ ID NO:3 bao gồm vùng đối xứng xuôi ngược, sự chèn hoặc sự tạo thành của yếu tố làm tăng sự biểu hiện trong gen nội sinh là không phụ thuộc vào hướng. SEQ ID NO:3 bao gồm sự lặp đoạn ba của đơn vị lặp đối xứng xuôi ngược 12 nucleotit yếu tố lõi GTAAGCGCTTAC (SEQ ID NO:1).

[0033] Trình tự nucleotit yếu tố lõi 12 nucleotit của SEQ ID NO:1 có mặt tại một số vị trí trong bộ gen ngô. Ví dụ, nó có thể được phát hiện tại một số vị trí nhiễm sắc thể của giống ngô B73. Theo phiên bản B73v4 của các trình tự bộ gen ngô (sẵn có trên trang web maizegdb.org/genome/assembly/Zm-B73-REFERENCE-GRAMENE-4.0 và sau đây được gọi là "bộ gen ngô B73v4"), SEQ ID NO:1 có thể được phát hiện trên các tọa độ Chr3 1,063,395..1,063,406 (intron của Zm00001d039287), trên các tọa độ Chr3 12,253,969..12,253,980 (ngay xuôi dòng của Zm00001d039695), trên các tọa độ Chr3 12,265,615..12,265,626 (intron của Zm00001d039695), trên các tọa độ Chr3 12,277,428..12,277,439 (intron của Zm00001d039695), trên các tọa độ Chr3 147,698,750..147,698,761 (không có bên trong 2 kb của mô hình gen được chú thích), trên Chr6, các tọa độ 107,132,183..107,132,194 (khoảng 2 kb xuôi dòng của Zm00001d036949), trên Chr10, các tọa độ 53,761,662..53,761,673 (không có bên trong

2 kb của mô hình gen được chú thích). Vì SEQ ID NO: 1 là đối xứng xuôi ngược, nó cũng được phát hiện trên sợi bổ sung của các tọa độ đó.

[0034] Trình tự nucleotit yếu tố lõi 12 nucleotit của SEQ ID NO:1 cũng có mặt tại một số vị trí trong bộ gen đậu tương. Ví dụ, nó có thể được phát hiện tại một số vị trí nhiễm sắc thể của giống đậu tương W82 (Williams 82). Theo phiên bản của các trình tự bộ gen đậu tương (sẵn có trên trang web ncbi.nlm.nih.gov/genome mục "*Glycine_max_v2.1.1*" và sau đây được gọi là "bộ gen đậu tương W82"), SEQ ID NO:1 có thể được phát hiện trên Nhiễm sắc thể 1, các tọa độ từ 44968347 đến 44968358, trên Nhiễm sắc thể 3, các tọa độ từ 3983755 đến 398873766, trên Nhiễm sắc thể 6, các tọa độ từ 26074553 đến 26074564, và trên Nhiễm sắc thể 19, các tọa độ từ 44574051 đến 44574062, 44575338 đến 44575349 (chồng lắp gen GLYMA_19G187100), và 44581677 đến 44581688. Vì SEQ ID NO: 1 là đối xứng xuôi ngược, nó cũng được phát hiện trên sợi bổ sung của các tọa độ đó.

[0035] Theo các phương án nhất định, polynucleotit, gen thực vật, hoặc thực vật bao gồm SEQ ID NO:3 có thể còn bao gồm ít nhất 1, 2, 3, hoặc nhiều bản sao bổ sung của SEQ ID NO:1. Các gen thực vật nội sinh mà được hướng đích đối với sự chèn hoặc sự tạo thành của SEQ ID NO:3 có thể bao gồm trình tự khởi động mà được nhận biết bởi ARN polymeraza II và được liên kết hoạt động với đơn vị phiên mã. Theo các phương án nhất định, các đơn vị phiên mã này có thể bao gồm các yếu tố đơn vị phiên mã bao gồm vùng không được dịch mã đầu 5' (5' UTR), intron, exon, vùng mã hóa microARN, vùng mã hóa tiền chất microARN, 3' UTR, vị trí polyadenyl hóa, hoặc tổ hợp của chúng. Theo các phương án nhất định, trình tự polynucleotit của SEQ ID NO:3 được chèn hoặc được tạo thành khoảng 10, 20, 30, hoặc 40 cặp bazơ (bp) đến khoảng 100, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp từ vị trí bắt đầu phiên mã (TSS) của đơn vị phiên mã (*tức là*, vị trí mũ đầu 5' của bản phiên mã được tạo ra bởi đơn vị phiên mã). Theo các phương án nhất định, trình tự polynucleotit của SEQ ID NO:3 được chèn hoặc được tạo thành khoảng 40 cặp bazơ (bp) đến khoảng 100, 240, 300, hoặc 400 bp từ vị trí bắt đầu phiên mã (TSS) của đơn vị phiên mã. Sự chèn hoặc sự tạo thành của SEQ ID NO:3 có thể xảy ra ở đầu 5' (*tức là*, ngược dòng) hoặc 3' (*tức là*, xuôi dòng) đối với TSS của đơn vị phiên mã. Khi trình tự polynucleotit của SEQ ID NO:3 được chèn hoặc được tạo thành ở đầu 5' đối với TSS, trình tự polynucleotit của SEQ ID NO: 3 do đó được định vị ở các vùng không được phiên mã của trình tự khởi động nội sinh ngược dòng

của TSS. Theo các phương án nhất định, sự chèn hoặc sự tạo thành của SEQ ID NO:3 là ở trình tự khởi động của gen nội sinh (ví dụ, khoảng 40 cặp bazơ (bp) đến khoảng 100, 240, 300, hoặc 400 bp 5' (tức là, ngược dòng) của TSS). Sự biểu hiện của gen chuyển mà được tích hợp vào trong bộ gen thực vật cũng có thể được tăng lên bởi sự chèn hoặc tạo thành của SEQ ID NO:3 ở gen chuyển sao cho nó được liên kết hoạt động với trình tự khởi động mà được liên kết hoạt động với gen chuyển này. Theo các phương án nhất định, sự chèn hoặc sự tạo thành của SEQ ID NO:3 là ở trình tự khởi động của gen nội sinh tại khoảng 50 bp đến khoảng 280 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 240 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 200 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 180 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 160 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 140 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 120 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 100 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 80 bp 5' đối với TSS, tại khoảng 50 bp đến khoảng 70 bp 5' đối với TSS, hoặc tại khoảng 50 bp đến khoảng 60 bp 5' đối với TSS. Theo các phương án nhất định, sự chèn hoặc sự tạo thành của SEQ ID NO:3 là ở trình tự khởi động của gen nội sinh tại khoảng 40 bp đến khoảng 280 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 240 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 200 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 180 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 160 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 140 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 120 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 100 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 80 bp 5' đối với TSS, tại khoảng 40 bp đến khoảng 70 bp 5' đối với TSS, hoặc tại khoảng 40 bp đến khoảng 60 bp 5' đối với TSS. Theo các phương án nhất định, sự chèn hoặc sự tạo thành của SEQ ID NO:3 là ở trình tự khởi động của gen nội sinh tại khoảng 35 bp đến khoảng 280 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 240 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 200 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 180 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 160 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 140 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 120 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 100 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 80 bp 5' đối với TSS, tại khoảng 35 bp đến khoảng 70 bp 5' đối với TSS, hoặc tại khoảng 35 bp đến khoảng 60 bp 5' đối với TSS. Sự biểu hiện của gen chuyển

cũng có thể được tăng bởi sự chèn hoặc sự tạo thành *in vitro* của SEQ ID NO:3 ở gen chuyển sao cho nó được liên kết hoạt động với trình tự khởi động mà được liên kết hoạt động với gen chuyển và sau đó đưa gen chuyển vào trong bộ gen thực vật (*ví dụ*, bởi sự biến nạp qua trung gian *Agrobacterium* hoặc bắn hạt sinh học).

[0036] Sự biểu hiện của các polynucleotit mã hóa bản phiên mã mà được liên kết hoạt động với yếu tố làm tăng bao gồm SEQ ID NO:3 có thể được tăng khi so sánh với cây đối chứng chứa các polynucleotit mã hóa bản phiên mã nhưng thiếu yếu tố làm tăng sự biểu hiện. Sự tăng này ở sự biểu hiện mà được trung gian bởi SEQ ID NO:3 có thể được đo bằng nhiều phương pháp. Theo các phương án nhất định, tính trạng được truyền bởi sự biểu hiện tăng của polynucleotit mã hóa bản phiên mã được đo ở cây chứa SEQ ID NO:3 và được so sánh với các cây đối chứng thiếu SEQ ID NO:3. Các ví dụ về các tính trạng như vậy mà có thể được đo và được so sánh theo cách này bao gồm áp lực phi sinh học, áp lực sinh học, kiến trúc, quang hợp, hoặc phân vùng tài nguyên được cải thiện. Những sự cải thiện ở các tính trạng này có thể được đánh giá bằng cách so sánh số đo bất kỳ của chính tính trạng này (*ví dụ*, hiệu quả sử dụng nước, kháng bệnh, giảm chiều cao cây, khả năng quang hợp được cải thiện, hoặc hiệu quả sử dụng nitơ) hoặc yếu tố đại diện của tính trạng (*ví dụ*, năng suất hạt và/hoặc sinh khối khác tính theo kg/hecta) ở các cây chứa SEQ ID NO:3 và được so sánh với các cây đối chứng thiếu SEQ ID NO:3. Theo các phương án nhất định, sự biểu hiện tăng của chính bản phiên mã được mã hóa được đo trực tiếp bằng việc xác định các lượng của bản phiên mã (*ví dụ*, mARN hoặc ARN không mã hóa) ở các cây chứa SEQ ID NO:3 và được so sánh với việc xác định các lượng của polynucleotit mã hóa bản phiên mã ở các cây đối chứng thiếu SEQ ID NO:3. Các lượng của bản phiên mã có thể được xác định bằng nhiều kỹ thuật bao gồm các kỹ thuật dựa trên PCR (*ví dụ*, PCR phiên mã ngược định lượng; qRT-PCR), sự lai hóa, CRISPR, và/hoặc giải tương tự (Khodakov et al., doi.org/10.1016/j.addr.2016.04.005; Gootenberg, et al. doi: 10.1126/science.aag0179). Theo các phương án nhất định, sự biểu hiện của polynucleotit mã hóa bản phiên mã cũng có thể được xác định bằng cách đo các lượng protein được mã hóa bởi bản phiên mã ở các cây chứa và so sánh với các lượng protein ở các cây đối chứng thiếu SEQ ID NO:3. Các lượng protein có thể được xác định bằng nhiều kỹ thuật bao gồm các thử nghiệm enzym (*ví dụ*, trong đó protein là enzym), các kỹ thuật dựa trên miễn dịch, và dựa trên phép đo phối khối (Chen et al. doi: 10.1186/s12967-015-0537-6; Bruce et al.

doi: 10.1002/0471250953.bi1321s41). Độ lớn sự việc tăng sản xuất bản phiên mã có thể phụ thuộc vào mức biểu hiện đường cơ sở của polynucleotit mã hóa bản phiên mã nội sinh chưa được cải biến ở các tế bào hoặc mô tương ứng. Ví dụ không làm giới hạn là, độ lớn của sự tăng biểu hiện của gen nội sinh được cải biến bởi sự chèn hoặc sự tạo thành của SEQ ID NO: 3 vượt qua các mức biểu hiện đường cơ sở của gen nội sinh chưa được cải biến sẽ là lớn nhất trong đó các mức biểu hiện đường cơ sở là thấp. Theo các phương án nhất định, sự biểu hiện của gen nội sinh được cải biến bởi sự chèn hoặc sự tạo thành của SEQ ID NO: 3 có thể được tăng bằng ít nhất 1,2, 1,5, 2, 3, 4, hoặc 5 lần so với các mức biểu hiện đường cơ sở của gen nội sinh chưa được cải biến. Theo các phương án nhất định, sự biểu hiện của gen nội sinh được cải biến bởi sự chèn hoặc sự tạo thành của SEQ ID NO: 3 có thể được tăng bằng ít nhất khoảng 1,2 hoặc 1,5 lần đến khoảng 2, 3, 4, 5, 6 lần hoặc cao hơn so với các mức biểu hiện đường cơ sở của gen nội sinh chưa được cải biến.

[0037] Theo các phương án nhất định, sẽ mong muốn sử dụng các phân tử chỉnh sửa bộ gen để đưa vào và tạo thành SEQ ID NO:3 trong bộ gen thực vật. Các phân tử chỉnh sửa gen để sử dụng trong các phương pháp được đưa ra ở đây bao gồm các phân tử có khả năng đưa sự đứt gãy sợi kép (double-strand break - "DSB") hoặc sự đứt gãy sợi đơn (single-strand break - "SSB") tại vị trí cụ thể hoặc trình tự ở ADN sợi kép, như ở ADN bộ gen hoặc ở gen đích được định vị bên trong ADN bộ gen cũng như ARN dẫn kèm theo hoặc các polynucleotit khuôn ADN cho hoặc khác. Các ví dụ về các phân tử chỉnh sửa gen như vậy bao gồm: (a) nucleaza bao gồm nucleaza được dẫn bởi ARN, ADN endonucleaza được dẫn bởi ARN hoặc ADN endonucleaza được hướng dẫn bởi ARN (RdDe), hệ nucleaza loại CRISPR lớp 1, nucleaza Cas loại II, Cas9, nCas9 nickaza, nucleaza Cas loại V, nucleaza Cas12a, nickaza nCas12a, Cas12d (CasY), Cas12e (CasX), Cas12b (C2c1), Cas12c (C2c3), Cas12i, Cas12j, Cas14, nucleaza được thiết kế di truyền, nucleaza được tối ưu hóa codon, nucleaza ngón tay kềm (ZFN) hoặc nickaza, nucleaza yếu tố tác động giống yếu tố hoạt hóa sự phiên mã (nucleaza yếu tố tác động TAL hoặc TALEN) hoặc nickaza (nickaza TALE, Argonaute, và meganucleaza hoặc meganucleaza được thiết kế di truyền; (b) polynucleotit mã hóa một hoặc nhiều nucleaza có khả năng thực hiện sự thay đổi đặc hiệu vị trí (bao gồm việc đưa DSB hoặc SSB vào) của trình tự nucleotit đích; (c) ARN dẫn (gARN) để sử dụng với nucleaza được dẫn bởi ARN, hoặc ADN mã hóa gARN để sử dụng với nucleaza được dẫn bởi ARN; (d) các

polynucleotit khuôn ADN cho thích hợp để chèn tại chỗ đứt gãy ở ADN bộ gen bằng sự sửa chữa được định hướng bởi thể tương đồng (HDR) hoặc cơ chế ghép nối đầu qua trung gian vi tương đồng (MMEJ); và (e) các khuôn ADN khác (ví dụ, dsADN, ssADN, hoặc các tổ hợp của chúng) thích hợp cho việc chèn tại chỗ đứt gãy ở ADN bộ gen (ví dụ, bằng cơ chế ghép nối đầu không tương đồng (NHEJ)).

[0038] Công nghệ CRISPR để chỉnh sửa các gen của sinh vật nhân chuẩn được bộc lộ trong các Công bố đơn sáng chế Mỹ 2016/0138008A1 và US2015/0344912A1, và trong các bằng sáng chế Mỹ 8,697,359, 8,771,945, 8,945,839, 8,999,641, 8,993,233, 8,895,308, 8,865,406, 8,889,418, 8,871,445, 8,889,356, 8,932,814, 8,795,965, và 8,906,616. Cpf1 endonucleaza và các ARN dẫn và các vị trí PAM tương ứng được bộc lộ trong Công bố đơn sáng chế Mỹ 2016/0208243 A1. Các trình tự khởi động ARN thực vật cho việc biểu hiện ARN dẫn CRISPR và endonucleaza CRISPR Cas9 được tối ưu hóa codon của thực vật được bộc lộ trong Đơn sáng chế Quốc tế PCT/US2015/018104 (được công bố là WO 2015/131101 và yêu cầu hưởng quyền ưu tiên của Đơn sáng chế tạm thời Mỹ 61/945,700). Các phương pháp sử dụng công nghệ CRISPR để chỉnh sửa bộ gen ở thực vật được bộc lộ trong các Công bố đơn sáng chế Mỹ US 2015/0082478A1 và US 2015/0059010A1 và trong Đơn sáng chế Quốc tế PCT/US2015/038767 A1 (được công bố là WO 2016/007347 và yêu cầu hưởng quyền ưu tiên của Đơn sáng chế tạm thời Mỹ 62/023,246). Theo các phương án nhất định, endonucleaza được dẫn bởi ARN mà để lại đầu cùn sau sự phân tách của vị trí đích được sử dụng. Các endonucleaza được dẫn bởi ARN cắt đầu cùn bao gồm Cas9, Cas12c, Cas12i, và Cas 12h (Yan et al., 2019). Theo các phương án nhất định, endonucleaza được dẫn bởi ARN mà để lại đầu treo ADN sợi đơn so le sau sự phân tách của vị trí đích được sử dụng. Các endonucleaza cắt đầu so le được dẫn bởi ARN bao gồm Cas12a, Cas12b, và Cas12e. Tất cả các công bố sáng chế được đề cập đến trong đoạn này được kết hợp toàn bộ ở đây bằng cách viện dẫn.

[0039] Sự chỉnh sửa bộ gen loại CRISPR có thể được làm thích ứng để sử dụng ở tế bào thực vật và các phương pháp được đưa ra ở đây theo một số cách. Các yếu tố CRISPR, ví dụ, phân tử chỉnh sửa gen bao gồm các endonucleaza CRISPR và các ARN dẫn CRISPR bao gồm các ARN dẫn đơn hoặc các ARN dẫn kết hợp với các tracrARN hoặc scoutARN, hoặc các polynucleotit mã hóa chúng, là hữu dụng trong thực hiện việc chỉnh sửa bộ gen mà không để lại di tích của các yếu tố CRISPR hoặc các chỉ thị gen chọn lọc

xuất hiện ở thể hệ sau. Theo các phương án nhất định, các yếu tố CRISPR được cung cấp trực tiếp cho tế bào nhân chuẩn (ví dụ, tế bào thực vật), các hệ thống, các phương pháp, và các chế phẩm ở dạng các phân tử được phân lập, ở dạng các chế phẩm được phân lập hoặc bán tinh khiết của quy trình tổng hợp không tế bào (ví dụ, sự dịch mã in vitro), hoặc ở dạng các sản phẩm được phân lập hoặc bán tinh khiết ở quy trình tổng hợp dựa trên tế bào (ví dụ, như ở chất dung giải tế bào vi khuẩn hoặc tế bào khác). Theo các phương án nhất định, thực vật và tế bào thực vật được sử dụng trong các hệ thống, phương pháp và chế phẩm được đưa ra ở đây có thể bao gồm gen chuyển mà biểu hiện endonucleaza CRISPR (ví dụ, endonucleaza Cas9, loại Cpf1 hoặc CRISPR khác). Theo các phương án nhất định, một hoặc nhiều endonucleaza CRISPR với các vị trí nhận biết PAM độc nhất có thể được sử dụng. Các ARN dẫn (các sgARN hoặc các crARN và tracrARN) để tạo thành phức hợp endonucleaza được dẫn bởi ARN/ARN dẫn mà có thể liên kết đặc hiệu với các trình tự trong vị trí đích gADN mà liền kề với trình tự môtip liền kề tiền vùng đệm (PAM). Loại endonucleaza được dẫn bởi ARN thông thường thông tin vị trí của các vị trí PAM thích hợp và thiết kế của các crARN hoặc sgARN. Các vị trí PAM giàu G, ví dụ, 5'-NGG thông thường được hướng đích cho thiết kế của các crARN hoặc sgARN được sử dụng với các protein Cas9. Các ví dụ về các trình tự PAM bao gồm 5'-NGG (*Streptococcus pyogenes*), 5'-NNAGAA (*Streptococcus thermophilus* CRISPR1), 5'-NGGNG (*Streptococcus thermophilus* CRISPR3), 5'-NNGRRT hoặc 5'-NNGRR (*Staphylococcus aureus* Cas9, SaCas9), và 5'-NNNGATT (*Neisseria meningitidis*). Các vị trí PAM giàu T (ví dụ, 5'-TTN hoặc 5'-TTTV, trong đó "V" là A, C, hoặc G) thông thường được hướng đích cho thiết kế của các crARN hoặc sgARN được sử dụng với các protein Cas12a. Trong một số trường hợp, Cas12a cũng có thể nhận biết môtip 5'-CTA PAM. Các ví dụ khác về các trình tự Cas12a PAM tiềm năng bao gồm TTN, CTN, TCN, CCN, TTTN, TCTN, TTCN, CTTN, ATTN, TCCN, TTGN, GTTN, CCCN, CCTN, TTAN, TCGN, CTCN, ACTN, GCTN, TCAN, GCCN, và CCGN (trong đó N được xác định là nucleotit bất kỳ). Endonucleaza Cpf1 và các ARN và các vị trí PAM tương ứng được bộc lộ trong Công bố đơn sáng chế Mỹ 2016/0208243 A1, mà được kết hợp ở đây bằng cách viện dẫn đối với sự bộc lộ của nó về ADN mã hóa các endonucleaza Cpf1 và các ARN dẫn và các vị trí PAM.

[0040] Theo các phương án nhất định, các nucleaza ngón tay kềm hoặc các nickaza ngón tay kềm cũng có thể được sử dụng trong các phương pháp được đề xuất ở đây. Các

nucleaza ngón tay kẽm là các endonucleaza đặc hiệu vị trí bao gồm hai miền protein: miền liên kết ADN, bao gồm nhiều đoạn lặp ngón tay kẽm riêng rẽ mà mỗi đoạn lặp này nhận biết giữa 9 và 18 cặp bazơ, và miền phân tách ADN mà bao gồm miền nucleaza (thông thường là FokI). Miền phân tách dime hóa để phân tách ADN; do đó, cần một cặp ZFN để hướng đích các polynucleotit đích không đối xứng xuôi ngược. Theo các phương án nhất định, các phương pháp thiết kế nucleaza ngón tay kẽm và nickaza ngón tay kẽm mà đã được mô tả (Urnov et al. (2010) *Nature Rev. Genet.*, 11:636 – 646; Mohanta et al. (2017) *Genes* vol. 8,12: 399; Ramirez et al. *Nucleic Acids Res.* (2012); 40(12): 5560–5568; Liu et al. (2013) *Nature Communications*, 4: 2565) có thể được làm thích ứng để sử dụng trong các phương pháp được nêu ở đây. Các miền liên kết ngón tay kẽm của nucleaza hoặc nickaza ngón tay cung cấp độ đặc hiệu và có thể được thiết kế để nhận biết một cách đặc hiệu trình tự ADN đích mong muốn bất kỳ. Các miền liên kết ADN ngón tay kẽm có nguồn gốc từ miền liên kết ADN của lớp lớn các yếu tố phiên mã sinh vật nhân chuẩn được gọi là protein ngón tay kẽm (ZFP). Miền liên kết ADN của các ZFP thông thường chứa mảng nối tiếp gồm ít nhất ba "ngón tay" kẽm, mỗi loại nhận biết bộ ba cụ thể của ADN. Nhiều chiến lược có thể được sử dụng để thiết kế độ đặc hiệu liên kết của miền liên kết ngón tay kẽm. Một phương pháp tiếp cận, được gọi là "sự lắp ráp môđun", dựa vào tính tự trị chức năng của các ngón tay kẽm riêng rẽ với ADN. Trong phương pháp này, trình tự đã cho được hướng đích bằng việc xác định các ngón tay kẽm cho mỗi bộ ba thành phần trong trình tự và liên kết chúng thành peptit đa ngón tay. Một số chiến lược thay thế để thiết kế các miền liên kết ADN ngón tay kẽm cũng đã được phát triển. Các phương pháp này được thiết kế để điều tiết khả năng của các ngón tay kẽm để tiếp xúc với các ngón tay lân cận cũng như các bazơ nucleotit ở bên ngoài bộ ba đích của chúng. Thông thường, miền liên kết ADN ngón tay kẽm được thiết kế di truyền có độ đặc hiệu liên kết mới, so với protein ngón tay kẽm xuất hiện tự nhiên. Các phương pháp thiết kế di truyền bao gồm, ví dụ, thiết kế hợp lý và các loại chọn lọc khác nhau. Thiết kế hợp lý bao gồm, ví dụ, sử dụng cơ sở dữ liệu của các trình tự nucleotit bộ ba (hoặc bộ tứ) và các trình tự axit amin ngón tay kẽm, trong đó mỗi trình tự nucleotit bộ ba hoặc bộ tứ có liên quan đến một hoặc nhiều trình tự axit amin của các ngón tay kẽm mà liên kết với trình tự bộ ba hoặc bộ tứ cụ thể. Xem, ví dụ, các bảng sáng chế Mỹ 6,453,242 và 6,534,261, cả hai bằng được kết hợp toàn bộ ở đây bằng cách viện dẫn. Các phương pháp chọn lọc làm ví dụ (ví dụ, biểu hiện thực khuẩn thể và hệ thống

hai thể lai nấm men) có thể được làm thích ứng để sử dụng trong các phương pháp được mô tả ở đây. Ngoài ra, sự tăng cường độ đặc hiệu liên kết đối với các miền liên kết ngón tay kẽm đã được mô tả trong bằng sáng chế Mỹ 6,794,136, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Ngoài ra, các miền ngón tay kẽm riêng rẽ có thể được liên kết với nhau sử dụng các trình tự cầu nối thích hợp bất kỳ. Các ví dụ về các trình tự cầu nối là đã biết công khai, ví dụ, xem các bằng sáng chế Mỹ 6,479,626; 6,903,185; và 7,153,949, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Miền phân tách axit nucleic là không đặc hiệu và thông thường là endonucleaza giới hạn, như FokI. Endonucleaza này phải dime hóa để phân tách ADN. Do đó, sự phân tách bằng FokI là bộ phận của ZFN yêu cầu hai sự kiện liên kết liên kề và độc lập, mà phải xuất hiện ở cả sự định hướng đúng và với sự đặt cách thích hợp để cho phép sự tạo thành dime. Yêu cầu đối với hai sự kiện liên kết ADN cho phép hướng đích đặc hiệu hơn của các vị trí nhận biết dài và độc nhất tiềm năng. Các biến thể FokI với các hoạt tính được tăng cường đã được mô tả và có thể được làm thích ứng để sử dụng trong các phương pháp được mô tả ở đây; xem, ví dụ, Guo et al. (2010) *J. Mol. Biol.*, 400:96 - 107.

[0041] Các yếu tố giống yếu tố hoạt hóa sự phiên mã (TALE) là các protein được tiết bởi loài *Xanthomonas* nhất định để điều biến sự biểu hiện gen ở thực vật chủ và để thúc đẩy sự tạo khuẩn lạc và sự sống sót của vi khuẩn. Các TALE hoạt động như các yếu tố phiên mã và điều biến sự biểu hiện của các gen kháng ở thực vật. Các kết quả gần đây của các TALE đã bộc lộ liên kết lõi vùng lặp của các TALE với các vị trí liên kết ADN đích của chúng. Các TALE bao gồm vùng bảo toàn và lặp lại cao gồm các đoạn lặp nối tiếp của hầu như là 33 hoặc 34 đoạn axit amin. Các monome lặp khác nhau chủ yếu ở các vị trí axit amin 12 và 13. Sự tương quan chặt giữa các cặp axit amin độc nhất tại các vị trí 12 và 13 và nucleotit tương ứng ở vị trí liên kết TALE đã được phát hiện. Mỗi quan hệ đơn giản giữa trình tự axit amin và sự nhận biết ADN của miền liên kết TALE cho phép đối với thiết kế của các miền liên kết ADN có độ đặc hiệu mong muốn bất kỳ. Các TALE có thể được liên kết với miền phân tách ADN không đặc hiệu để tạo ra các protein chỉnh sửa bộ gen, được gọi là các nucleaza yếu tố tác động TAL hoặc các TALEN. Như trong trường hợp của các ZFN, endonucleaza giới hạn, như FokI, có thể được sử dụng một cách thuận tiện. Các phương pháp để sử dụng các TALEN ở thực vật đã được mô tả và có thể được làm thích ứng để sử dụng ở các phương pháp được mô tả ở đây, xem Mahfouz et al. (2011) *Proc. Natl. Acad. Sci. USA*, 108:2623 – 2628;

Mahfouz (2011) GM Crops, 2:99 – 103; và Mohanta et al. (2017) Genes vol. 8,12: 399). Các nickaza TALE cũng có thể được mô tả và có thể được làm thích ứng để sử dụng ở các phương pháp được mô tả ở đây (Wu et al.; Biochem Biophys Res Commun. (2014);446(1):261-6; Luo et al; Scientific Reports 6, Article number: 20657 (2016)).

[0042] Theo các phương án nhất định trong đó SEQ ID NO:3 được chèn vào trong bộ gen tại một vị trí của sự đứt gãy sợi kép ở bộ gen thực vật được đưa vào bởi một hoặc nhiều nucleaza hoặc nickaza (ví dụ, phức hợp CRISPR/ARN dẫn với endonucleaza hoặc nickaza đặc hiệu vị trí, nucleaza hoặc nickaza aZF, và/hoặc nucleaza hoặc nickaza TALE), khuôn ADN cho hoặc khuôn ADN khác bao gồm SEQ ID NO:3. Theo các phương án nhất định trong đó SEQ ID NO:3 được tạo thành trong bộ gen tại một vị trí của sự đứt gãy sợi kép ở bộ gen thực vật được đưa vào bằng nucleaza, khuôn ADN cho hoặc khuôn ADN khác có thể bao gồm ít hơn tập hoàn chỉnh gồm 36 nucleotit hoặc cặp bazơ của trình tự SEQ ID NO:3 và ADN bộ gen tại vị trí tích hợp có thể đóng góp các nucleotit hoặc các cặp bazơ của SEQ ID NO:3 mà mất mặt khỏi khuôn ADN cho hoặc khuôn ADN khác. Theo các phương án nhất định trong đó SEQ ID NO:3 được tạo thành ở ADN bộ gen, khuôn ADN cho hoặc khuôn ADN khác có thể bao gồm 24 đến 35 nucleotit tiếp giáp hoặc các cặp bazơ của SEQ ID NO:3, ADN bộ gen tại vị trí tích hợp có thể đóng góp từ 1 đến 12 nucleotit hoặc cặp bazơ của trình tự SEQ ID NO:3 mà thiếu khỏi khuôn ADN cho hoặc khuôn ADN khác, và trình tự 36 cặp bazơ hoàn chỉnh của SEQ ID NO:3 được tạo thành tại vị trí tích hợp trong bộ gen. Các phân tử khuôn ADN cho được sử dụng ở các phương pháp được đề xuất ở đây bao gồm các phân tử ADN bao gồm, từ đầu 5' đến 3', nhánh tương đồng thứ nhất, ADN thay thế, và nhánh tương đồng thứ hai, trong đó các nhánh tương đồng chứa các trình tự mà tương đồng một phần hoặc hoàn toàn với các trình tự ADN bộ gen (gADN) gần sườn vị trí phân tách endonucleaza đặc hiệu vị trí đích ở gADN. Theo các phương án nhất định, ADN thay thế có thể bao gồm sự chèn, sự làm khuyết hoặc sự biểu hiện 1 hoặc nhiều cặp bazơ của ADN so với gADN đích. Theo một phương án, phân tử khuôn ADN cho là sợi kép và tốt hơn là được bắt cặp bazơ qua toàn bộ hoặc hầu hết chiều dài của nó, với sự loại trừ có thể có của các nucleotit không ghép cặp bất kỳ tại một trong hai đầu hoặc cả hai đầu. Theo phương án khác, phân tử khuôn ADN cho là sợi kép và bao gồm một hoặc nhiều sự bắt cặp nhằm không phải đầu cùng hoặc nucleotit không ghép cặp không phải ở đầu cùng bên trong bộ đôi sợi kép khác. Theo một phương án, phân tử khuôn ADN cho mà

được tích hợp tại vị trí của ít nhất một sự đứt gãy sợi kép (DSB) bao gồm giữa 2-20 nucleotit ở một sợi (nếu là sợi đơn) hoặc ở cả hai sợi (nếu là sợi kép), *ví dụ*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, hoặc 20 nucleotit trên một hoặc trên cả hai sợi, mỗi sợi trong số chúng có thể được bắt cặp bazơ với nucleotit trên sợi đối diện của vị trí tích hợp được hướng đích (trong trường hợp tốt hơn là phân tử nucleotit sợi kép được bắt cặp bazơ). Các khuôn ADN cho như vậy có thể được tích hợp trong ADN bộ gen chứa các đứt gãy ADN sợi kép đầu cùn và/hoặc so le bởi đoạn lặp được dẫn hướng tương đồng (HDR) hoặc cơ chế ghép nối đầu qua trung gian vi tương đồng (MMEJ). Theo các phương án nhất định, nhánh tương đồng khuôn ADN cho có thể có chiều dài khoảng 20, 50, 100, 200, 400, hoặc 600 đến khoảng 800, hoặc 1000 cặp bazơ. Theo các phương án nhất định, phân tử khuôn ADN cho có thể được phân phối đến tế bào thực vật ở dạng phân tử ADN mạch vòng (*ví dụ*, *plasmid* hoặc vectơ virus bao gồm vectơ geminivirus) hoặc dạng mạch thẳng. Theo các phương án nhất định, phân tử ADN mạch vòng hoặc mạch thẳng mà có thể được sử dụng bao gồm phân tử khuôn ADN cho được cải biến bao gồm, từ đầu 5' đến 3', bản sao thứ nhất của trình tự vị trí phân tách endonuclease đặc hiệu trình tự đích, nhánh tương đồng thứ nhất, ADN thay thế, nhánh tương đồng thứ hai, và bản sao thứ hai của trình tự vị trí phân tách endonuclease đặc hiệu trình tự đích. Theo các phương án khác, các khuôn ADN thích hợp cho việc chèn NHEJ sẽ thiếu các nhánh tương đồng mà tương đồng một phần hoặc hoàn toàn với trình tự ADN bộ gen (gADN) gần sườn vị trí phân tách endonuclease đặc hiệu vị trí đích ở gADN. Theo các phương án nhất định, khuôn ADN bao gồm tất cả SEQ ID NO:3 (*ví dụ*, dsADN, ssADN, hoặc các tổ hợp của nó) có thể được chèn tại chỗ đứt gãy sợi kép ở gADN bởi cơ chế ghép nối đầu không tương đồng (NHEJ). Theo các phương án nhất định, khuôn ADN bao gồm ít hơn bộ hoàn chỉnh gồm 36 nucleotit hoặc cặp bazơ của trình tự SEQ ID NO:3 (*ví dụ*, dsADN, ssADN, hoặc các tổ hợp của chúng) có thể được chèn tại chỗ đứt gãy sợi kép in gADN bởi cơ chế ghép nối đầu không tương đồng (NHEJ), gADN tại vị trí chèn có thể đóng góp các nucleotit hoặc các cặp bazơ của SEQ ID NO:3 mà vắng mặt khỏi khuôn ADN, và SEQ ID NO:3 có thể được tạo thành tại vị trí của sự đứt gãy sợi kép ở gADN.

[0043] Theo một số phương án, yếu tố tăng cường của SEQ ID NO:3 thay thế hoặc thay thế lớn trình tự tương ứng ở gen thực vật, như trình tự khởi động. Do đó, sự thay thế thay vì sự chèn để lại sự định vị các yếu tố khác không thay đổi. Vị trí đích thay thế có

thể được chọn bởi độ tương tự với SEQ ID NO: 3. Khuôn thay thế có thể được sử dụng ở quy trình HDR, và/hoặc việc chỉnh sửa bazơ ADN và/hoặc chỉnh sửa bộ gen có thể được sử dụng để tạo ra vùng thay thế mong muốn mà tương ứng với SEQ ID NO: 3. Các yếu tố chỉnh sửa bazơ bao gồm, ví dụ, chỉnh sửa bazơ đặc hiệu vị trí được trung gian bởi các enzym deaminaza chỉnh sửa bazơ từ C*G thành T·A hoặc A·T thành G*C (Gaudelli et al., "Programmable base editing of A·T to G*C in genomic ADN without ADN cleavage." *Nature* (2017); Nishida et al. "Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems." *Science* 353 (6305) (2016); Komor et al. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage." *Nature* 533 (7603) (2016):420-4. dCas9 chết xúc tác được dung hợp với protein xytidin deaminaza hoặc adenin deaminaza trở thành yếu tố chỉnh sửa bazơ đặc hiệu mà có thể thay đổi các bazơ ADN mà gây ra sự đứt gãy ADN. Các yếu tố chỉnh sửa bazơ chuyển đổi C->T (hoặc G->A trên sợi đối diện) hoặc yếu tố chỉnh sửa bazơ adenin mà sẽ chuyển đổi adenin thành inosin, dẫn đến sự thay đổi A->G với ngưỡng chỉnh sửa được định cụ thể bởi gARN.

[0044] Các xử lý khác nhau có thể được sử dụng để phân phối các phân tử chỉnh sửa gen và/hoặc các phân tử khác đến tế bào thực vật. Theo các phương án nhất định, một hoặc nhiều cách xử lý được sử dụng để phân phối phân tử chỉnh sửa gen hoặc các phân tử khác (ví dụ, bao gồm polynucleotit, polypeptit hoặc tổ hợp của chúng) vào trong tế bào sinh vật nhân chuẩn hoặc tế bào thực vật, ví dụ, qua các rào chắn như thành tế bào, màng huyết tương, vỏ bọc nhân, và/hoặc lớp kép lipid khác. Theo các phương án nhất định, chế phẩm chứa polynucleotit, polypeptit, hoặc RNP (ribonucleoprotein) bao gồm các phân tử được phân phối một cách trực tiếp, ví dụ bằng sự tiếp xúc trực tiếp của chế phẩm với tế bào thực vật. Các chế phẩm đã nêu trên có thể được cung cấp ở dạng chất lỏng, dung dịch, huyền phù, nhũ tương, nhũ tương nghịch, dạng keo, dạng phân tán, dạng gel, liposom, mixen, vật liệu có thể tiêm được, sol khí, chất rắn, dạng bột, dạng hạt, dạng hạt nano, hoặc tổ hợp của chúng có thể được áp dụng trực tiếp đến cây, bộ phận của cây, tế bào thực vật, hoặc mô cây thực vật (ví dụ, qua việc làm trầy vỏ hoặc đâm thủng hoặc sự đứt đoạn khác của thành tế bào hoặc màng tế bào, bằng cách phun, nhúng hoặc ngâm hoặc cách tiếp xúc trực tiếp theo, bằng cách vi tiêm). Ví dụ, tế bào thực vật hoặc tế bào trần của thực vật được ngâm trong chế phẩm chứa phân tử chỉnh sửa bộ gen dạng lỏng, nhờ vậy tác nhân được phân phối đến tế bào thực vật. Theo các phương án nhất định,

chế phẩm chứa tác nhân này được phân phối bằng cách sử dụng áp suất âm hoặc dương, ví dụ, sử dụng sự thâm nhiễm chân không hoặc áp dụng áp suất thủy động hoặc áp suất chất lưu. Theo các phương án nhất định, chế phẩm chứa tác nhân này được đưa vào trong tế bào thực vật hoặc tế bào trần của thực vật, *ví dụ*, bằng vi tiêm hoặc bằng cách làm đứt đoạn hoặc làm biến dạng thành tế bào hoặc màng tế bào, ví dụ bằng xử lý vật lý như bằng cách áp dụng áp suất âm hoặc dương, lực trượt, hoặc xử lý bằng tác nhân phân phối hóa học hoặc vật lý như chất hoạt động bề mặt, liposom, hoặc hạt nano; xem, *ví dụ*, sự phân phối vật liệu đến tế bào sử dụng dòng vi chất lưu qua sự co thắt biến dạng tế bào như được mô tả trong Đơn sáng chế được công bố của Mỹ 2014/0287509, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Các kỹ thuật khác sử dụng để phân phối chế phẩm chứa tác nhân đến tế bào nhân chuẩn, tế bào thực vật hoặc tế bào trần của thực vật bao gồm: siêu âm hoặc khuấy siêu âm; sự chấn động, ma sát, ứng suất trượt, rung lắc, sự tạo hang; ly tâm hoặc áp dụng lực cơ học; sự biến dạng hoặc sự đứt vỡ thành tế bào hoặc màng tế bào bằng cơ học; sự đứt vỡ hoặc sự thẩm hóa thành tế bào hoặc màng tế bào bằng enzym; sự làm trầy hoặc rạch cơ học (*ví dụ*, làm trầy bằng carborundum hoặc chất làm trầy dạng hạt khác hoặc sự rạch bằng rũa hoặc giấy giáp) hoặc sự rạch hóa học (*ví dụ*, xử lý bằng tác nhân axit hoặc ăn da); và xung điện. Theo các phương án nhất định, chế phẩm chứa tác nhân được cung cấp bởi sự chuyển nạp qua trung gian vi khuẩn (*ví dụ*, *Agrobacterium* sp., *Rhizobium* sp., *Sinorhizobium* sp., *Mesorhizobium* sp., *Bradyrhizobium* sp., *Azobacter* sp., *Phyllobacterium* sp.) của tế bào thực vật hoặc tế bào trần của thực vật với polynucleotit mã hóa các phân tử chỉnh sửa bộ gen (*ví dụ*, endonucleaza ADN phụ thuộc ARN, protein liên kết ADN phụ thuộc vào ARN, nickaza phụ thuộc ARN, ABE, hoặc CBE, và/hoặc ARN dẫn); xem, *ví dụ*, Broothaerts *et al.* (2005) *Nature*, 433:629 – 633). Kỹ thuật bất kỳ trong số các kỹ thuật này hoặc tổ hợp của chúng được sử dụng thay thế trên mô cấy thực vật, bộ phận hoặc mô của cây hoặc cây nguyên vẹn (hoặc hạt) mà tế bào thực vật tùy ý về cơ bản được thu hoặc được tách ra từ đó; theo các phương án nhất định, chế phẩm chứa tác nhân được phân phối ở bước riêng rẽ sau khi tế bào thực vật đã được tách.

[0045] Các kỹ thuật để thực hiện việc chỉnh sửa bộ gen ở cây trồng (ví dụ, ngô,) bao gồm việc sử dụng các yếu tố tạo hình morphogen như Wuschel (WUS), Protein phát triển noãn (ODP), và/hoặc Babyboom (BBM) mà có thể cải thiện hiệu quả của việc thu hồi cây có các chỉnh sửa bộ gen mong muốn. Theo một số khía cạnh, yếu tố tạo hình

morphogen bao gồm WUS1, WUS2, WUS3, WOX2A, WOX4, WOX5, WOX9, BBM2, BMN2, BMN3, và/hoặc ODP2. Theo các phương án nhất định, các chế phẩm và phương pháp để sử dụng WUS, BBM, và/hoặc ODP, cũng như các kỹ thuật khác mà có thể được làm thích ứng để thực hiện chỉnh sửa bộ gen ở thực vật hoặc các nguồn gen khác, được nêu trong các tài liệu US 20030082813, US 20080134353, US 20090328252, US 20100100981, US 20110165679, US 20140157453, US 20140173775, và US 20170240911, mà mỗi tài liệu được kết hợp toàn bộ ở đây bằng cách viện dẫn. Theo các phương án nhất định, những sự chỉnh sửa bộ gen có thể được thực hiện ở các bộ phận của thực vật có thể tái sinh được (*ví dụ*: phôi thực vật) của cây trồng bằng sự cung cấp nhất thời các phân tử chỉnh sửa gen hoặc polynucleotit mã hóa chúng và không nhất thiết yêu cầu việc kết hợp gen chỉ thị chọn lọc vào trong bộ gen thực vật (*ví dụ*, US 20160208271 và US 20180273960, cả hai tài liệu này được kết hợp toàn bộ ở đây bằng cách viện dẫn; Svitashchev et al. Nat Commun. 2016; 7:13274).

[0046] Yếu tố tăng cường sự biểu hiện bao gồm SEQ ID NO:3 có thể được liên kết hoạt động với các gen mã hóa bản phiên mã ngoại sinh (*tức là*, gen chuyển được mã hóa) hoặc nội sinh mà bao gồm các gen truyền các tính trạng như năng suất được cải thiện, các đặc tính thực phẩm và/hoặc thức ăn được cải thiện (*ví dụ*, chất lượng hoặc số lượng dầu, tinh bột, protein, hoặc axit amin được cải thiện), hiệu quả sử dụng nitơ được cải thiện (*ví dụ*, gen glutamin synthetase), các đặc tính sử dụng nhiên liệu sinh học được cải thiện (*ví dụ*, sự sản sinh ethanol được cải thiện), sự ra hoa được trì hoãn, không ra hoa, kháng áp lực sinh học tăng (*ví dụ*, kháng sự hư hỏng do côn trùng, tuyến trùng, vi khuẩn, hoặc nấm), kháng áp lực phi sinh học tăng (*ví dụ*, kháng hạn, lạnh, nhiệt, kim loại, hoặc muối), tính chống đổ được tăng cường, tốc độ sinh trưởng được tăng cường, sinh khối tăng, sự đâm chồi tăng, đẻ nhánh tăng, thời gian ra hoa được trì hoãn, sự già hóa bị trì hoãn, số lượng hoa tăng, kiến trúc tăng đối với trồng cây mật độ cao, khả năng quang hợp được cải thiện, khối lượng rễ tăng, số lượng tế bào tăng, sức sống cây con được cải thiện, kích thước cây con được cải thiện, tốc độ phân chia tế bào tăng, hiệu quả chuyển hóa được cải thiện, và/hoặc kích thích mô phân sinh tăng) khi so sánh với cây đối chứng thiếu sự thay đổi di truyền được hướng đích này. Các polynucleotit bao gồm SEQ ID NO:3 có thể được liên kết hoạt động với các gen như vậy bằng sự chèn các polynucleotit trong bộ gen thực vật mà dẫn đến sự chèn hoặc sự tạo thành SEQ ID NO:3 ở gen thực vật nội sinh. Các vị trí ở gen thực vật nội sinh thích hợp cho sự chèn hoặc sự tạo thành

của SEQ ID NO:3 bao gồm các vùng khởi động, mã hóa, và không mã hóa (ví dụ, các 5' UTR, intron, và 3' UTR). Các thực vật đích thích hợp cho sự chèn hoặc sự tạo thành SEQ ID NO: 3 bao gồm các thực vật và tế bào thực vật thuộc các loài quan tâm bất kỳ, bao gồm thực vật hai lá mầm và thực vật một lá mầm. Thực vật quan tâm bao gồm cây trồng theo hàng, cây và cây thân gỗ lấy quả, cây rau, cây thân gỗ, và cây trang trí bao gồm hoa trang trí, cây bụi, cây thân gỗ, cây che phủ mặt đất, và cây cỏ thảm. Các ví dụ về các cây trồng, cây thân gỗ, và cây canh tác quan trọng về mặt thương mại bao gồm: cỏ linh lăng (*Medicago sativa*), cây hạnh nhân (*Prunus dulcis*), cây táo (*Malus x domestica*), cây mơ (*Prunus armeniaca*, *P. brigantiae*, *P. mandshurica*, *P. mume*, *P. sibirica*), măng tây (*Asparagus officinalis*), chuối (*Musa* spp.), lúa mạch (*Hordeum vulgare*), cây đậu (*Phaseolus* spp.), cây việt quất và nam việt quất (*Vaccinium* spp.), cây cacao (*Theobroma cacao*), cây cải dầu và cây hạt cải dầu hoặc cây hạt cải lấy dầu, (*Brassica napus*), cây cẩm chướng (*Dianthus caryophyllus*), cây cà rốt (*Daucus carota sativus*), cây sắn (*Manihot esculentum*), cây anh đào (*Prunus avium*), cây đậu gà (*Cicer arietinum*), rau diếp xoăn (*Cichorium intybus*), cây ớt và cây ớt chuông khác (*Capsicum annuum*, *C. frutescens*, *C. chinense*, *C. pubescens*, *C. baccatum*), cây hoa cúc (*Chrysanthemum* spp.), cây dừa (*Cocos nucifera*), cây cà phê (*Coffea* spp. bao gồm *Coffea arabica* và *Coffea canephora*), cây bông (*Gossypium hirsutum* L.), cây đậu dài (*Vigna unguiculata*), cây dưa chuột (*Cucumis sativus*), cây lý chua và cây lý gai (*Ribes* spp.), cây cà tím hoặc cây cà quả tím (*Solanum melongena*), cây khuynh diệp (*Eucalyptus* spp.), cây lanh (*Linum usitatissimum* L.), cây phong lữ (*Pelargonium* spp.), cây bưởi chùm (*Citrus x paradisi*), cây nho (*Vitis* spp.) bao gồm nho làm rượu (*Vitis vinifera*), cây ổi (*Psidium guajava*), cây hoa bia (*Humulus lupulus*), cây gai dầu và cây cần sa (*Cannabis sativa* và *Cannabis* spp.), cây irit (*Iris* spp.), cây chanh (*Citrus limon*), cây rau diếp (*Lactuca sativa*), chanh lá cam (*Citrus* spp.), cây ngô (*Zea mays* L.), cây xoài (*Mangifera indica*), cây măng cụt (*Garcinia mangostana*), cây dưa lưới (*Cucumis melo*), cây kê (*Setaria* spp, *Echinochloa* spp, *Eleusine* spp, *Panicum* spp., *Pennisetum* spp.), cây yến mạch (*Avena sativa*), cây cọ dầu (*Ellis quineensis*), cây ôliu (*Olea europaea*), cây hành (*Allium cepa*), cây cam (*Citrus sinensis*), cây đu đủ (*Carica papaya*), cây đào và cây xuân đào (*Prunus persica*), cây lê (*Pyrus* spp.), cây đậu Hà Lan (*Pisa sativum*), cây lạc (*Arachis hypogaea*), cây mẫu đơn (*Paeonia* spp.), cây thuốc lá cảnh (*Petunia* spp.), cây dứa (*Ananas comosus*), cây mã đề (*Musa* spp.), cây mận

(*Prunus domestica*), cây trạch nguyên (*Euphorbia pulcherrima*), cây cải dầu Ba Lan (*Brassica rapa*), cây bạch dương (*Populus* spp.), cây khoai tây (*Solanum tuberosum*), cây bí đỏ (*Cucurbita pepo*), cây lúa nước (*Oryza sativa* L.), cây hoa hồng (*Rosa* spp.), cây cao su (*Hevea brasiliensis*), cây lúa mạch đen (*Secale cereale*), cây rum (*Carthamus tinctorius* L.), hạt vừng (*Sesame indium*), cây cao lương (*Sorghum bicolor*), cây đậu tương (*Glycine max* L.), cây bí xanh (*Cucurbita pepo*), cây dâu tây (*Fragaria* spp., *Fragaria x ananassa*), củ cải đường (*Beta vulgaris*), cây mía đường (*Saccharum* spp.), cây hoa hướng dương (*Helianthus annuus*), cây khoai lang (*Ipomoea batatas*), cây quýt (*Citrus tangerina*), cây chè (*Camellia sinensis*), cây thuốc lá (*Nicotiana tabacum* L.), cây cà chua (*Lycopersicon esculentum*), cây hoa tulip (*Tulipa* spp.), cây củ cải (*Brassica rapa rapa*), cây óc chó (*Juglans* spp. L.), cây dưa hấu (*Citrulus lanatus*), cây lúa mì (*Tritium aestivum*), và cây khoai mỡ (*Discorea* spp.).

[0047] Các loại gen nói chung được quan tâm bao gồm, nhưng không bị giới hạn ở, các gen liên quan đến thông tin, như các yếu tố phiên mã bao gồm các yếu tố phiên mã chứa ngón tay kẽm, các loại gen liên quan đến sự truyền thông, như các kinaza và/hoặc các yếu tố tải nạp tín hiệu khác, và các gen liên quan đến việc giữ nhà, như các protein sốc nhiệt. Các loại cụ thể hơn, ví dụ, bao gồm, nhưng không bị giới hạn ở, các gen mã hóa các tính trạng quan trọng đối với nông học (ví dụ, năng suất tăng, hiệu quả sử dụng nito được cải thiện), chống chịu áp lực vô học (ví dụ, hiệu quả sử dụng nước tăng, chống chịu nóng, lạnh, và hạn), chống chịu áp lực sinh học (ví dụ, kháng côn trùng, kháng bệnh nấm hoặc vi khuẩn, và kháng tuyến trùng), kháng thuốc trừ cỏ, các đặc tính bất dục, hạt hoặc hạt, và các sản phẩm thương mại. Các gen quan tâm bao gồm, nói chung là, các gen liên quan đến sự chuyển hóa dầu, tinh bột, carbohydrat, hoặc dinh dưỡng cũng như các gen ảnh hưởng đến kích thước hạt, sự phát triển của cây, sự điều hòa sinh trưởng của cây, và sự cải thiện năng suất. Sự phát triển và sự điều hòa sinh trưởng của cây cũng để chỉ sự phát triển và sự điều hòa sinh trưởng của các bộ phận khác nhau của cây, như hoa, hạt, rễ, lá và chồi.

[0048] Các gen kháng bệnh và /hoặc côn trùng có thể mã hóa tính kháng sinh vật gây hại mà có sự giảm năng suất lớn như, ví dụ, Bệnh cháy lá ngô miền Bắc, bệnh đen đầu, bệnh thán thư, virus khảm đậu tương, tuyến trùng nang đậu tương, tuyến trùng gây bướu rễ, bệnh đốm lá nâu, Bệnh sương mai, bệnh hạt tím, bệnh thối hạt và các bệnh cây con gây ra thông thường bởi các loại nấm - *Pythium* sp., *Phytophthora* sp., *Rhizoctonia* sp.,

Diaporthe sp.. Bệnh bạc lá vi khuẩn gây ra bởi vi khuẩn *Pseudomonas syringae* pv. *Glycinea*. Các gen truyền tính kháng côn trùng bao gồm, ví dụ, các gen protein độc *Bacillus thuringiensis* (Bằng sáng chế Mỹ số 5,366,892; 5,747,450; 5,737,514; 5,723,756; 5,593,881; và Geiser et al (1986) *Gene* 48:109); các lectin (Van Damme et al. (1994) *Plant Mol. Biol.* 24:825); và các loại tương tự.

[0049] Tính trạng kháng thuốc diệt cỏ có thể bao gồm các gen mã hóa cho tính kháng thuốc diệt cỏ mà tác động để ức chế hoạt động của axetolactat synthaza (ALS), cụ thể là các thuốc diệt cỏ loại sulfonylure (ví dụ, gen axetolactat synthaza ALS chứa các đột biến dẫn đến tính kháng này, cụ thể là các đột biến S4 và/hoặc HRA). Các thể đột biến gen ALS mã hóa tính kháng đối với thuốc diệt cỏ clorsulfuron. Glyphosat axetyl transferaza (GAT) là N-axetyltransferaza từ *Bacillus licheniformis* mà được tối ưu hóa bằng sự xáo trộn gen cho sự axetyl hóa của thuốc diệt cỏ phổ rộng, glyphosat, tạo thành cơ sở của cơ chế dung nạp glyphosat mới ở thực vật chuyển gen (Castle et al. (2004) *Science* 304, 1 151 -1 154).

[0050] Các gen liên quan đến sự sinh trưởng và phát triển của thực vật đã được xác định ở thực vật. Một gen như vậy, mà liên quan đến sự sinh tổng hợp xytokinin, là isopentenyl transferaza (IPT). Xytokinin đóng vai trò quan trọng ở sự sinh trưởng và phát triển của thực vật bởi sự kích thích phân chia tế bào và sự biệt hóa tế bào (Sun et al. (2003), *Plant Physiol.* 131 : 167-176).

[0051] Theo các phương án nhất định, sáng chế dự tính sự chèn trình tự yếu tố tăng cường vào trong tế bào nhận tại nhiều hơn một locut có lợi.

[0052] Sản phẩm thực vật hàng hóa thu được từ cây hoặc các bộ phận của cây bao gồm SEQ ID NO:3 cũng như các phương pháp để tạo ra các sản phẩm như vậy được đề xuất. Theo các phương án nhất định, các sản phẩm hàng hóa là các sản phẩm được xử lý được tạo ra từ cây hoặc hạt của nó, bao gồm: (a) bột thô ngô, đậu tương, bông, hoặc hạt cải dầu (được khử béo hoặc không được khử béo); (b) các protein, dầu, đường, sirô, và tinh bột được tách chiết; (c) các sản phẩm lên men; (d) các sản phẩm thức ăn chăn nuôi hoặc thành phần cho người (ví dụ, thức ăn chăn nuôi và thực phẩm bao gồm bột thô ngô, đậu tương, bông, hoặc hạt cải dầu (được khử béo hoặc không được khử béo) và các thành phần khác (ví dụ, các hạt ngũ cốc khác, bột thô hạt khác, bột thô protein khác, dầu khác, tinh bột khác, đường khác, chất gắn kết, chất bảo quản, chất hút ẩm, vitamin, và/hoặc

khoáng chất); (e) sản phẩm được lý; (f) sinh khối thô hoặc đã được xử lý (*ví dụ*, vật liệu xenluloza và/hoặc lignoxenluloza; thức ăn ủ xilô); và (g) các sản phẩm công nghiệp khác.

[0053] Sáng chế cũng đề xuất các phương pháp để phát hiện SEQ ID NO:3 ở loại bất kỳ trong số các mẫu sinh học và các sản phẩm hàng hóa đã nêu trên. Sự phát hiện các phân tử ADN bao gồm SEQ ID NO:3 có thể đạt được bằng tổ hợp bất kỳ trong số các kỹ thuật dựa trên sự khuếch đại axit nucleic (*ví dụ*, sự khuếch đại PCR), sự lai hóa, giải trình tự, và/hoặc đo phổ khối. Các phương pháp đã nêu để phát hiện các axit nucleic ngoại lai ở các locut gen chuyển được nêu trong tài liệu US 20190136331 và US 9,738,904, cả hai tài liệu này được kết hợp toàn bộ ở đây bằng cách viện dẫn, có thể được làm thích ứng để sử dụng trong việc phát hiện các axit nucleic được đưa ra ở đây. Theo các phương án nhất định, sự phát hiện này đạt được bằng các phương pháp phát hiện dựa trên sự khuếch đại và/hoặc sự lai hóa sử dụng phương pháp (*ví dụ*, các môi khuếch đại chọn lọc) và/hoặc mẫu dò (*ví dụ*, có khả năng lai hóa hoặc sản sinh một cách chọn lọc sản phẩm kéo dài môi đặc hiệu) mà nhận biết một cách đặc hiệu phân tử ADN đích (*ví dụ*, vị trí cắt locut gen chuyển) nhưng không nhận biết ADN từ locut gen chuyển không được cải biến. Theo các phương án nhất định, các mẫu dò sự lai hóa (*ví dụ*, các polynucleotit bao gồm ít nhất từ 15 đến 36 nucleotit của SEQ ID NO:3) có thể bao gồm các chất đánh dấu phát hiện được (*ví dụ*, các chất đánh dấu phát huỳnh quang, phóng xạ, epitop, và phát quang hóa học). Theo các phương án nhất định, thử nghiệm phát hiện sự đa hình nucleotit đơn có thể được làm thích ứng để phát hiện phân tử ADN đích (*ví dụ*, vị trí chèn hoặc tạo thành SEQ ID NO:3 ở bộ gen thực vật).

[0054] Các cây và hạt lai cùng dòng và lai bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã được đưa ra ở đây cùng với các phương pháp tạo ra và sử dụng hạt lai và hạt lai cùng dòng này. Các phương pháp sản xuất hạt lai cùng dòng bao gồm cây lai cùng dòng cùng máu và hạn chế thụ phân chéo bởi cây bất kỳ không phải cây lai cùng dòng. Các phương pháp sản xuất hạt lai như vậy có thể bao gồm bước lai các dòng cây trồng ưu tú mà ít nhất một trong số thể cho hoặc thể nhận phân bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã. Theo các phương án nhất định, thể cho và nhận phân sẽ bao gồm nguồn gen của các nhóm ưu thế lai khác biệt và cung cấp hạt và cây lai thể hiện ưu thế lai. Theo các phương án nhất định, cây lai cùng dòng, cây lai, thể cho phân và/hoặc thể nhận phân

mỗi loại có thể bao gồm locut chuyển gen khác biệt mà truyền tính trạng khác biệt (ví dụ, dung nạp thuốc diệt cỏ hoặc kháng côn trùng), loại tính trạng khác (ví dụ, dung nạp với các thuốc diệt cỏ khác biệt hoặc chống chịu côn trùng khác biệt như côn trùng cánh cứng hoặc côn trùng cánh vảy), hoặc chế độ hoạt động khác nhau đối với cùng một tính trạng (ví dụ, kháng côn trùng cánh cứng bởi hai cơ chế hoạt động khác biệt hoặc kháng côn trùng cánh vảy bởi hai cơ chế hoạt động khác biệt). Các gen chuyển mà có thể được đưa vào trong dòng thực vật bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã bằng cách chọn tạo giống hoặc bằng sự biến nạp trực tiếp bao gồm: (i) các gen chuyển mà truyền tính kháng côn trùng (ví dụ, các gen chuyển mà sản xuất các protein *Bacillus thuringiensis* bao gồm Cry1Ab, Cry1Ac, Cry1F, Cry2Ab, Cry2Ae, Cry3A, Cry3Bb, Cry9c, Cry34, Cry35, VIP3A, và các biến thể của chúng; các gen chuyển mà cảm ứng sự đáp ứng ARNi ức chế côn trùng bao gồm *dvsnf7*); và (ii) các gen chuyển mà truyền khả năng dung nạp thuốc trừ cỏ (ví dụ, CP4-EPSPS hoặc các gen EPSPS khác mà truyền khả năng dung nạp glyphosat; các gen PAT hoặc BAR mà truyền tính kháng với thuốc diệt cỏ glufosinat; các gen *aad-1* mà truyền tính kháng với thuốc diệt cỏ 2,4-D và aryloxyphenoxypropionat; các gen DMO mà truyền tính kháng với thuốc diệt cỏ dicamba). Các ví dụ về các sự kiện cây ngô, đậu tương, bông, và cải dầu chuyển gen mà truyền tính trạng như dung nạp thuốc diệt cỏ và/hoặc chống chịu sinh vật gây hại được bộc lộ trong Bằng sáng chế Mỹ số 7323556, 8575434, 6040497, 10316330; 8618358, 8212113, 9428765, 8455720, 7897748, 8273959, 8093453, 8901378, 8466346, RE44962, 9540655, 9738904, 8680363, 8049071, 9447428, 9944945, 8592650, 10184134, 7179965, 7371940, 9133473, 8735661, 7381861, 8048632, và 9738903, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Các gen chuyển mà có thể được sử dụng để truyền tính kháng côn trùng ở ngô bao gồm các gen được bộc lộ trong Công bố đơn sáng chế Mỹ số 20150361446 và 20200190533, được kết hợp toàn bộ ở đây bằng cách viện dẫn, cũng như các gen được bộc lộ trong Bằng sáng chế Mỹ số 6342660, 6852915, 7323556, 7695914, 7705216, 7897748, 8212113, 8455720, 8466346, 8575434, 8901378, 9428765, và 10316330, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Các gen chuyển mà có thể được sử dụng để truyền khả năng dung nạp thuốc diệt cỏ ở ngô bao gồm các gen được bộc lộ trong Công bố đơn sáng chế Mỹ số 20120244533 và 20200190533, được kết hợp toàn bộ ở đây bằng cách viện dẫn, cũng như các gen được bộc lộ trong Bằng sáng chế Mỹ số

6040497, 6852915, 8273959, 8618358, 8759618, và 9994863, được kết hợp toàn bộ ở đây bằng cách viện dẫn. Theo các phương án nhất định, thể nhận phần sẽ là bất dục được tạo ra hoặc bất dục được có điều kiện. Các phương pháp để cảm ứng tính bất dục được hoặc tính bất dục được có điều kiện bao gồm khử dục (ví dụ, khử còi), tính bất dục được tế bào chất, các tác nhân hoặc các hệ thống lai hóa học, các gen chuyển hoặc các hệ gen chuyển, và/hoặc (các) đột biến ở một hoặc nhiều gen thực vật nội sinh. Những mô tả về các hệ thống bất dục được khác nhau mà có thể được làm thích ứng để sử dụng với cây trồng được đưa ra ở đây được mô tả trong tài liệu Wan et al. *Molecular Plant*; 12, 3, (2019):321-342 cũng như trong US 8,618,358; US 20130031674; và US 2003188347.

[0055] Theo các phương án nhất định, các thực vật được đưa ra ở đây mà bao gồm SEQ ID NO: 3 có thể còn bao gồm một hoặc nhiều sự thay đổi di truyền được hướng đích được đưa vào bằng một hoặc nhiều phân tử hoặc hệ thống chỉnh sửa gen. Sự thay đổi di truyền được hướng đích như vậy bao gồm các thay đổi mang lại các tình trạng như năng suất được cải thiện, các đặc tính thực phẩm và/hoặc thức ăn được cải thiện (ví dụ, chất lượng hoặc số lượng dầu, tinh bột, protein, hoặc axit amin được cải thiện), hiệu quả sử dụng nitơ được cải thiện, các đặc tính sử dụng nhiên liệu sinh học được cải thiện (ví dụ, sự sản xuất ethanol được cải thiện), các hệ thống bất dục được/bất dục được có điều kiện (ví dụ, bằng cách hướng đích các gen MS26, MS45 và MSCA1 nội sinh), khả năng dung nạp thuốc diệt cỏ (ví dụ, bằng cách hướng đích các gen ALS, EPSPS, HPPD, hoặc các gen đích thuốc diệt cỏ khác), sự ra hoa được trì hoãn, không ra hoa, tính kháng áp lực sinh học tăng (ví dụ, tính kháng hư hại do côn trùng, tuyến trùng, vi khuẩn, hoặc nấm), tính kháng áp lực phi sinh học tăng (ví dụ, kháng hạn, lạnh, nhiệt, kim loại, hoặc muối), tính kháng đổ được tăng cường, tốc độ sinh trưởng được tăng cường, sinh khối được tăng cường, khả năng đâm chồi được tăng cường, khả năng đẻ nhánh được tăng cường, thời gian ra hoa được trì hoãn, sự già hóa được trì hoãn, số lượng hoa tăng, kiến trúc được cải thiện để trồng mật độ cao, khả năng quang hợp được cải thiện, khối lượng rễ tăng, số lượng tế bào tăng, sức sống cây con được cải thiện, kích thước cây con được cải thiện, tốc độ phân chia tế bào tăng, hiệu quả chuyển hóa được cải thiện, và kích thước mô phân sinh tăng khi so sánh với cây đối chứng thiếu sự thay đổi di truyền được hướng đích. Các loại thay đổi di truyền được hướng đích mà có thể được đưa vào bao gồm các sự chèn, làm khuyết, và thế một hoặc nhiều nucleotit ở bộ gen cây trồng. Các vị trí ở các gen thực vật nội sinh cho sự thay đổi di truyền được hướng đích bao gồm các vùng trình

tự khởi động, mã hóa, và không mã hóa (ví dụ, các 5' UTR, intron, các vị trí cho và nhận ghép nối và 3' UTR). Theo các phương án nhất định, sự thay đổi di truyền được hướng đích bao gồm sự chèn trình tự điều hòa hoặc trình tự ADN khác trong gen thực vật nội sinh. Các ví dụ không giới hạn về các trình tự điều hòa mà có thể được chèn vào trong các gen thực vật nội sinh với các phân tử chỉnh sửa gen để thực hiện các thay đổi di truyền được hướng đích mà đem lại các kiểu hình hữu dụng bao gồm các trình tự được nêu trong Công bố đơn sáng chế Mỹ 20190352655, mà được kết hợp toàn bộ ở đây bằng cách viện dẫn, như: (a) trình tự yếu tố đáp ứng auxin (AuxRE); (b) ít nhất một trình tự D1-4 (Ulmasov et al. (1997) Plant Cell, 9:1963-1971), (c) ít nhất một trình tự DR5 (Ulmasov et al. (1997) Plant Cell, 9:1963-1971); (d) ít nhất một trình tự m5-DR5 (Ulmasov et al. (1997) Plant Cell, 9:1963-1971); (e) ít nhất một trình tự P3; (f) trình tự vị trí nhận biết ARN nhỏ được liên kết bởi ARN nhỏ tương ứng (ví dụ, siARN, microARN (miARN), siARN hoạt động trans như được mô tả trong Bằng sáng chế Mỹ số 8,030,473, hoặc sARN được chia pha như được mô tả trong Bằng sáng chế Mỹ số 8,404,928; cả hai bằng sáng chế được trích dẫn này được kết hợp ở đây bằng cách viện dẫn); (g) trình tự vị trí nhận biết microARN (miARN); (h) trình tự nhận biết microARN (miARN) đối với miARN được thiết kế di truyền trong đó tác nhân liên kết đặc hiệu là miARN trưởng thành được thiết kế tương ứng; (i) trình tự nhận biết gen nhảy; (j) trình tự được nhận biết bởi môtip ức chế lường tính liên quan đến yếu tố liên kết yếu tố đáp ứng etylen (EAR); (k) trình tự vị trí cắt nối (ví dụ, vị trí cho, vị trí phân nhánh, hoặc vị trí nhận; xem, ví dụ, các vị trí cắt nối và các tín hiệu cắt nối được nêu trên trang web [lemur\[dot\]amu\[dot\]edu\[dot\]pl/share/ERISdb/home.html](http://lemur[dot]amu[dot]edu[dot]pl/share/ERISdb/home.html)); (l) trình tự vị trí nhận biết recombinaza mà được nhận biết bởi recombinaza đặc hiệu vị trí; (m) trình tự mã hóa aptame ARN hoặc axit amin hoặc công tắc ribo ARN, tác nhân liên kết đặc hiệu là phối tử tương ứng, và thay đổi ở sự biểu hiện là sự điều hòa tăng hoặc điều hòa giảm; (n) yếu tố đáp ứng hormon được nhận biết bởi thụ thể nhân hoặc miền liên kết hormon của chúng; (o) trình tự liên kết yếu tố phiên mã; và (p) yếu tố đáp ứng dạng đa lược (xem Xiao et al. (2017) Nature Genetics, 49:1546-1552, doi: 10.1038/ng.3937). Các ví dụ không giới hạn về các gen ngô đích mà có thể được cho tiến hành chỉnh sửa gen hướng đích để mang lại các tính trạng hữu dụng bao gồm: (a) ZmIPK1 (ngô dung nạp thuốc trừ cỏ và khử phytat; Shukla et al., Nature. 2009;459:437–41); (b) ZmGL2 (sáp tâm mô sừng ngoài giảm ở lá; Char et al. Plant Biotechnol J. 2015;13:1002); (c) ZmMTL (cảm

ứng cách cây đơn bội; Kelliher et al. *Nature*. 2017;542:105); (d) Wx1 (hàm lượng amylopectin cao; US 20190032070; được kết hợp toàn bộ ở đây bằng cách viện dẫn); (e) TMS5 (bất dục đực mẫn cảm nhiệt; Li et al. *J Genet Genomics*. 2017;44:465–8); (f) ALS (dung nạp thuốc trừ cỏ; Svitashv et al.; *Plant Physiol*. 2015;169:931–45); và (g) ARGOS8 (chống chịu áp lực hạn; Shi et al., *Plant Biotechnol J*. 2017;15:207–16). Các ví dụ không giới hạn về các gen đậu tương đích mà có thể được cho tiến hành chỉnh sửa gen được hướng đích để đem lại các tính trạng hữu dụng bao gồm: (a) FAD2-1A, FAD2-1B (hàm lượng axit oleic tăng; Haun et al.; *Plant Biotechnol J*. 2014;12:934–40); (b) FAD2-1A, FAD2-1B, FAD3A (hàm lượng axit oleic tăng và hàm lượng linolenic giảm; Demorest et al., *BMC Plant Biol*. 2016;16:225); và (c) ALS (dung nạp thuốc trừ cỏ; Svitashv et al.; *Plant Physiol*. 2015;169:931–45). Các ví dụ không giới hạn về các gen *Brassica* đích mà có thể được cho tiến hành chỉnh sửa gen được hướng đích để đem lại các tính trạng hữu dụng bao gồm: (a) gen FRIGIDA để đem lại sự ra hoa sớm (Sun Z, et al. *J Integr Plant Biol*. 2013;55:1092–103); và (b) ALS (dung nạp thuốc trừ cỏ; US 20160138040, được kết hợp toàn bộ ở đây bằng cách viện dẫn). Các ví dụ không giới hạn về các gen đích ở cây trồng bao gồm ngô và đậu tương mà có thể được cho tiến hành các thay đổi di truyền được hướng đích mà đem lại các kiểu hình hữu dụng bao gồm các gen được nêu trong Công bố đơn sáng chế Mỹ số 20190352655, 20200199609, 20200157554, và 20200231982, mà được kết hợp toàn bộ ở đây bằng cách viện dẫn; và Zhang et al. (*Genome Biol*. 2018; 19: 210). Theo các phương án nhất định, các thay đổi di truyền được hướng đích có thể được kết hợp với các cây mà chứa SEQ ID NO: 3 bằng các kỹ thuật tạo giống. Các kỹ thuật tạo giống như vậy bao gồm lai và/hoặc sự nhập nội gen bằng cách lai lại với cây mẹ hồi quy. Trong các phép lai này, các cây mà chứa SEQ ID NO: 3 có thể là thể cho hoặc thể nhận phần. Theo các phương án nhất định, các cây mà chứa SEQ ID NO: 3 có thể được sử dụng làm cây mẹ hồi quy ở các phép lai lại này để nhập nội sự thay đổi di truyền được hướng đích vào trong nguồn gen thực vật bao gồm SEQ ID NO: 3. Theo các phương án nhất định, các cây mà bao gồm (các) thay đổi di truyền đích có thể được sử dụng làm cây mẹ hồi quy trong các phép lai lại để nhập nội vùng bộ gen bao gồm SEQ ID NO: 3 vào trong nguồn gen thực vật bao gồm (các) thay đổi di truyền đích.

[0056] Theo các phương án nhất định, các thực vật được đưa ra ở đây mà bao gồm SEQ ID NO: 3 có thể còn bao gồm một hoặc nhiều locut di truyền đem lại các tính trạng như

năng suất được cải thiện, các đặc tính thực phẩm và/hoặc thức ăn được cải thiện (ví dụ, chất lượng hoặc số lượng dầu, tinh bột, protein hoặc axit amin được cải thiện), hiệu quả sử dụng nitơ được cải thiện, các đặc tính sử dụng nhiên liệu sử dụng được cải thiện (ví dụ, sự sản xuất etanol được cải thiện), các hệ thống bất dục đực/bất dục đực có điều kiện (ví dụ, bằng cách hướng đích các gen MS26, MS45 và MSCA1 nội sinh), khả năng dung nạp thuốc trừ cỏ (ví dụ, bằng cách hướng đích các gen ALS, EPSPS, HPPD nội sinh, hoặc các gen đích thuốc diệt cỏ khác), sự ra hoa được trì hoãn, không ra hoa, tính kháng áp lực sinh học tăng (ví dụ, tính kháng hư hại do côn trùng, tuyến trùng, vi khuẩn, hoặc nấm), tính kháng áp lực phi sinh học tăng (ví dụ, kháng hạn, lạnh, nhiệt, kim loại, hoặc muối), tính kháng đổ được tăng cường, tốc độ sinh trưởng được tăng cường, sinh khối được tăng cường, khả năng đâm chồi được tăng cường, khả năng đẻ nhánh được tăng cường, thời gian ra hoa được trì hoãn, sự già hóa được trì hoãn, số lượng hoa tăng, kiến trúc được cải thiện để trồng mật độ cao, khả năng quang hợp được cải thiện, khối lượng rễ tăng, số lượng tế bào tăng, sức sống cây con được cải thiện, kích thước cây con được cải thiện, tốc độ phân chia tế bào tăng, hiệu quả chuyển hóa được cải thiện, và kích thước mô phân sinh tăng khi so sánh với cây đối chứng thiếu sự thay đổi di truyền được hướng đích. Các nguồn của các locut di truyền này bao gồm các giống canh tác ưu tú, các họ hàng hoang hại tương thích sinh dục (ví dụ, *Glycine soja*), nguồn gen thực vật mà đã được tiến hành gây đột biến ngẫu nhiên (ví dụ, với tác nhân gây đột biến hóa học như EMS hoặc với sự gây đột biến bằng tia gamma), và loại tương tự. Theo các phương án nhất định, các locut di truyền như vậy có thể được kết hợp với các cây mà bao gồm SEQ ID NO: 3 bằng các kỹ thuật tạo giống. Các kỹ thuật tạo giống như vậy bao gồm lai và/hoặc sự nhập nội gen bằng cách lai lại với cây mẹ hồi quy. Trong các phép lai này, các cây mà chứa SEQ ID NO: 3 có thể là thể cho hoặc thể nhận phấn. Theo các phương án nhất định, các cây mà chứa SEQ ID NO: 3 có thể được sử dụng làm cây mẹ hồi quy trong các phép lai lại này để nhập nội locut di truyền vào trong nguồn gen thực vật chứa SEQ ID NO: 3. Theo các phương án nhất định, các cây mà bao gồm locut hoặc các locut di truyền có thể được sử dụng làm cây mẹ hồi quy ở các phép lai lại như vậy để nhập nội vùng bộ gen chứa SEQ ID NO: 3 vào trong nguồn gen thực vật bao gồm locut hoặc các locut di truyền.

[0057] Sáng chế cũng đề xuất phương pháp để sản xuất sản phẩm thực vật hàng hóa hoặc vật liệu thực vật bao gồm bước trồng cây bất kỳ trong số các cây đã nêu trên chứa SEQ

ID NO:3 hoặc trồng các cây từ hạt chứa SEQ ID NO:3. Theo các phương án nhất định, các cây và/hoặc hạt như vậy được tưới, bón phân, và/hoặc được xử lý bằng tác nhân sinh học (ví dụ, vi sinh vật có lợi cho thực vật bao gồm *Bacillus* sp., a *Rhizobium* sp., và loại tương tự), chất diệt tuyến trùng (ví dụ, chất diệt côn trùng carbamat hoặc organophosphat), chất diệt côn trùng (ví dụ, chất diệt côn trùng neonicotinoit, pyrethroit, carbamat, hoặc organophosphat) và/hoặc chất diệt nấm (ví dụ, chất diệt nấm benzimidazol, imidazol, hoặc strobilurin). Các cây có thể được xử lý bằng các phân bón, tác nhân sinh học, chất diệt tuyến trùng, chất diệt côn trùng, và chất diệt nấm như vậy bằng các phương pháp bao gồm phun, xông khói, và/hoặc làm ướt đất. Các hạt có thể được xử lý bằng phân bón, tác nhân sinh học, chất diệt tuyến trùng, chất diệt côn trùng, và chất diệt nấm như vậy bằng các phương pháp bao gồm áp dụng theo luống hoặc bằng cách phủ (ví dụ, với thiết bị phủ trồng, thiết bị phủ quay, trồng nhào trộn, thiết bị tăng sôi, và/hoặc tăng phun). Các phương pháp và chế phẩm bao gồm các chất gắn kết, chất độn, chất phủ màng, và các thành phần hoạt tính như phân bón, chất hoạt động bề mặt, chất điều hòa sinh trưởng, chất làm khô cây trồng, chất diệt nấm, chất diệt vi khuẩn, chất kìm hãm vi khuẩn, chất diệt côn trùng, và chất xua đuổi côn trùng để phủ hạt mà có thể được làm thích ứng để sử dụng với các hạt được đưa ra ở đây được bộc lộ trong Bảng sáng chế Mỹ số 10745578, mà được kết hợp toàn bộ ở đây bằng cách viện dẫn.

Các phương án

[0058] Các phương án khác nhau về các phân tử ADN, cây, bộ phận của cây, bộ gen, nhiễm sắc thể, phương pháp, mẫu sinh học, và các chế phẩm khác được mô tả ở đây được nêu trong tập hợp các phương án được đánh số sau đây.

[0059] 1. Phân tử ADN bao gồm trình tự polynucleotit của SEQ ID NO:3.

[0060] 2. Phân tử ADN theo phương án 1, trong đó phân tử ADN này là phân tử ADN được phân lập, tổng hợp, và/hoặc tái tổ hợp.

[0061] 3. Phân tử AADN theo phương án 1 hoặc 2, trong đó trình tự polynucleotit của SEQ ID NO:3 được liên kết hoạt động với trình tự polynucleotit bao gồm trình tự khởi động, tùy ý trong đó trình tự khởi động này là trình tự khởi động nội sinh, và tùy ý trong đó trong đó trình tự khởi động nội sinh này được định vị ở trong nhiễm sắc thể của thực vật.

[0062] 4. Phân tử AADN theo phương án 1, 2, hoặc 3, trong đó trình tự polynucleotit của SEQ ID NO:3 và trình tự khởi động được liên kết hoạt động với đơn vị phiên mã và làm tăng sự biểu hiện của bản phiên mã hoặc protein được mã hóa bởi đơn vị phiên mã này ở tế bào thực vật so với tế bào thực vật đối chứng bao gồm trình tự khởi động trong sự vắng mặt của trình tự polynucleotit của SEQ ID NO:3.

[0063] 5. Phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 4, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị từ khoảng 10, 20, 30, 35, 40, hoặc 50 cặp bazơ (bp) đến khoảng 70, 80, 100, 120, 140, 160, 180, 200, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp của vị trí bắt đầu phiên mã của đơn vị phiên mã, và tùy ý trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị tại đầu 5' đối với TSS.

[0064] 6. Phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 5, trong đó trình tự khởi động là trình tự khởi động thực vật nội sinh và đơn vị phiên mã là đơn vị phiên mã thực vật nội sinh, trong đó trình tự khởi động và đơn vị phiên mã được định vị trong nhiễm sắc thể thực vật, và tùy ý trong đó trình tự khởi động thực vật nội sinh và yếu tố đơn vị phiên mã thực vật nội sinh là trình tự khởi động và đơn vị phiên mã thực vật của ngô, đậu tương, bông, hoặc cây cải dầu.

[0065] 7. Mẫu sinh học bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6.

[0066] 8. Mẫu sinh học theo phương án 7, trong đó mẫu này bao gồm: (i) bột, sirô, dầu, hoặc tinh bột; hoặc (ii) bột thô hoặc phiến mỏng của hạt ngô, đậu tương, bông, hoặc cải dầu.

[0067] 9. Nhiễm sắc thể thực vật bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6, tùy ý trong đó nhiễm sắc thể thực vật là nhiễm sắc thể cây ngô, đậu tương, bông, hoặc cây cải dầu.

[0068] 10. Nhiễm sắc thể thực vật theo phương án 9, trong đó nhiễm sắc thể thực vật này là nhiễm sắc thể của cây ngô.

[0069] 11. Nhiễm sắc thể thực vật theo phương án 9 hoặc 10, trong đó một hoặc nhiều nucleotit của SEQ ID NO:3 là nucleotit nội sinh của nhiễm sắc thể thực vật, tùy ý trong

đó ít nhất 24 nucleotit của SEQ ID NO:3 là các nucleotit ngoại sinh được chèn vào trong nhiễm sắc thể thực vật.

[0070] 12. Tế bào thực vật bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6, tùy ý trong đó tế bào thực vật này là tế bào cây ngô, đậu tương, bông, hoặc cây cải dầu.

[0071] 13. Tế bào thực vật theo phương án 12, trong đó trình tự polynucleotit của SEQ ID NO:3 được liên kết hoạt động với trình tự polynucleotit bao gồm trình tự khởi động, trong đó trình tự khởi động này là trình tự khởi động nội sinh, và trong đó trình tự khởi động nội sinh này được định vị trong nhiễm sắc thể thực vật.

[0072] 14. Tế bào thực vật theo phương án 12 hoặc 13, trong đó tế bào thực vật này là tế bào cây ngô.

[0073] 15. Thể nuôi cấy mô của tế bào có thể tái sinh được bao gồm tế bào thực vật theo phương án 12.

[0074] 16. Cây chứa phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6, tùy ý trong đó cây này là cây ngô, đậu tương, bông hoặc cây cải dầu.

[0075] 17. Cây theo phương án 16, trong đó cây này là cây ngô, đậu tương, bông, hoặc cây cải dầu.

[0076] 18. Cây theo phương án 16 hoặc 17, trong đó trình tự polynucleotit của SEQ ID NO:3 được liên kết hoạt động với trình tự polynucleotit bao gồm trình tự khởi động, trong đó trình tự khởi động này là trình tự khởi động nội sinh, và trong đó trình tự khởi động nội sinh này được định vị trong nhiễm sắc thể thực vật.

[0077] 19. Bộ phận của cây bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6, tùy ý trong đó bộ phận của cây là bộ phận của cây ngô, đậu tương, bông, hoặc cây cải dầu.

[0078] 20. Bộ phận của cây theo phương án 19, trong đó bộ phận của cây là bộ phận của cây ngô.

[0079] 21. Bộ phận của cây theo phương án 19 hoặc 20, trong đó bộ phận của cây là hạt, tùy ý trong đó hạt này là hạt ngô.

[0080] 22. Bộ phận của cây theo phương án 19, 20, hoặc 21, trong đó hạt này là hạt lai, tùy ý trong đó hạt lai này là hạt lai F1, và tùy ý trong đó hạt lai F1 này thể hiện ưu thế lai.

[0081] 23. Bộ phận của cây theo phương án bất kỳ trong số các phương án từ 19 đến 22, trong đó hạt này là hạt lai cùng dòng.

[0082] 24. Phương pháp sản xuất hạt thực vật, phương pháp này bao gồm bước lai cây theo phương án bất kỳ trong số các phương án từ 16 đến 18 với cây thứ hai để tạo ra hạt thực vật và tùy ý thu hoạch hạt này.

[0083] 25. Phương pháp theo phương án 24, trong đó kiểu gen của cây này và kiểu gen của cây thứ hai là các kiểu gen khác nhau và hạt là hạt lai.

[0084] 26. Phương pháp theo phương án 24 hoặc 25, trong đó cây này hoặc cây thứ hai còn bao gồm gen chuyển, sự thay đổi di truyền được hướng đích, hoặc locut di truyền mang lại tính trạng mong muốn và hạt được thu hoạch bao gồm gen chuyển, sự thay đổi di truyền được hướng đích, hoặc locut di truyền này.

[0085] 27. Phương pháp sản xuất hạt thực vật, phương pháp này bao gồm bước tự phối cây theo phương án bất kỳ trong số các phương án từ 16 đến 18 để sản xuất hạt thực vật và tùy ý thu hoạch hạt này.

[0086] 28. Phương pháp theo phương án 27, trong đó cây và hạt này được lai cùng dòng.

[0087] 29. Phương pháp sản xuất cây bao gồm tính trạng mong muốn được bổ sung, phương pháp này bao gồm bước đưa gen chuyển, sự thay đổi di truyền được hướng đích, hoặc locut di truyền mang lại tính trạng mong muốn vào trong cây theo phương án 16, 17, hoặc 18.

[0088] 30. Phương pháp sản xuất sản phẩm thực vật hàng hóa, phương pháp này bao gồm bước xử lý cây hoặc hạt bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6 và thu hồi sản phẩm thực vật hàng hóa từ cây hoặc hạt đã được xử lý.

[0089] 31. Phương pháp theo phương án 30, sản phẩm thực vật hàng hóa là bột thô hạt, tinh bột, sirô, thức ăn ủ xilô, dầu, hoặc protein.

[0090] 32. Phương pháp theo phương án 30 hoặc 31, trong đó sản phẩm thực vật hàng hóa bao gồm lượng có thể phát hiện được của phân tử ADN.

[0091] 33. Phương pháp sản xuất vật liệu thực vật, phương pháp này bao gồm bước trồng cây có yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã, trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã ở cây đã nêu được tăng khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện.

[0092] 34. Phương pháp theo phương án 33, trong đó việc trồng cây bao gồm ít nhất một trong số các bước gieo hạt mà nảy mầm và tạo thành cây, tưới nước cho hạt hoặc cây, và/hoặc xử lý cây hoặc hạt bằng tác nhân sinh học, chất diệt cỏ, chất diệt côn trùng, hoặc chất diệt nấm.

[0093] 35. Phương pháp sản xuất vật liệu thực vật, phương pháp này bao gồm bước:

(a) cung cấp cây có yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3 được liên kết hoạt động với polynucleotit mã hóa bản phiên mã, trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã được tăng ở cây đã nêu khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện; và,

(b) trồng cây dưới các điều kiện mà cho phép sự biểu hiện của polynucleotit thúc đẩy bản phiên mã.

[0094] 36. Phương pháp theo phương án bất kỳ trong số các phương án từ 33 đến 35, trong đó vật liệu thực vật bao gồm hạt, tùy ý trong đó phương pháp này bao gồm bước thu hoạch hạt từ cây này.

[0095] 37. Phương pháp theo phương án bất kỳ trong số các phương án từ 33 đến 36, trong đó yếu tố làm tăng sự biểu hiện là trong trình tự khởi động nội sinh mà được liên kết hoạt động với polynucleotit mã hóa bản phiên mã hoặc trong 5' UTR, intron, hoặc 3' UTR của polynucleotit mã hóa bản phiên mã và trong đó trình tự khởi động nội sinh, 5' UTR, intron, hoặc 3' UTR đã nêu được định vị ở nhiễm sắc thể thực vật.

[0096] 38. Phương pháp theo phương án bất kỳ trong số các phương án từ 33 đến 37, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị từ khoảng 10, 20, 30, 40, hoặc 50 cặp bazơ (bp) đến khoảng 70, 80, 100, 120, 140, 160, 180, 200, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp từ vị trí bắt đầu phiên mã của đơn vị phiên mã, và tùy ý trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị tại đầu 5' đối với TSS.

[0097] 39. Phương pháp theo phương án bất kỳ trong số các phương án từ 33 đến 38, trong đó polynucleotit mã hóa bản phiên mã là đơn vị phiên mã thực vật nội sinh, trong đó trình tự khởi động và đơn vị phiên mã được định vị ở nhiễm sắc thể thực vật, tùy ý trong đó polynucleotit mã hóa bản phiên mã là polynucleotit mã hóa bản phiên mã của cây ngô, đậu tương, bông, hoặc cây cải dầu.

[0098] 40. Phương pháp theo phương án bất kỳ trong số các phương án từ 33 đến 39, trong đó cây này là cây ngô và/hoặc trong đó hạt là hạt của cây ngô.

[0099] 41. Phương pháp xác định mẫu sinh học bao gồm polynucleotit bao gồm gen thực vật được cải biến, bao gồm bước phát hiện sự biểu hiện của SEQ ID NO:3 ở mẫu sinh học.

[00100] 42. Phương pháp theo phương án 41, trong đó mẫu sinh học này thu được từ tế bào thực vật, cây, hoặc bộ phận của cây, tùy ý trong đó bộ phận của cây là hạt, và/hoặc tùy ý trong đó cây này là cây ngô, đậu tương, bông, hoặc cây cải dầu.

[00101] 43. Phương pháp sản xuất các axit nucleic bao gồm trình tự polynucleotit của SEQ ID NO:3, phương pháp này bao gồm bước tách các axit nucleic từ cây theo phương án 16, 17, hoặc 18 hoặc từ bộ phận của cây theo phương án bất kỳ trong số các phương án từ 19 đến 23.

[00102] 44. Phương pháp sản xuất hạt thực vật được xử lý bao gồm bước cho hạt bao gồm phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6 tiếp xúc với chế phẩm chứa tác nhân sinh học, chất diệt tuyến trùng, chất diệt côn trùng, hoặc chất diệt nấm.

[00103] 45. Phương pháp theo phương án 44, trong đó hạt là hạt của cây ngô, đậu tương, bông, hoặc cây cải dầu.

[00104] 46. Phương pháp làm tăng sự biểu hiện của trình tự polynucleotit ở thực vật, phương pháp này bao gồm bước cho biểu hiện trình tự polynucleotit mà liên kết hoạt động với yếu tố tăng cường sự biểu hiện bao gồm SEQ ID NO:3 và với trình tự khởi động, trong đó sự biểu hiện của polynucleotit được tăng khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện.

[00105] 47. Phương pháp theo phương án 46, trong đó cây này là cây ngô, đậu tương, bông, hoặc cây cải dầu, tùy ý trong đó trình tự polynucleotit bao gồm gen mà mang lại

năng suất được cải thiện, các đặc tính thực phẩm và/hoặc thức ăn được cải thiện (*ví dụ*, chất lượng hoặc số lượng dầu, tinh bột, protein, hoặc axit amin được cải thiện), hiệu quả sử dụng nitơ được cải thiện, các đặc tính sử dụng nhiên liệu sinh học được cải thiện (*ví dụ*, sự sản sinh etanol được cải thiện), sự ra hoa được trì hoãn, không ra hoa, kháng áp lực sinh học tăng (*ví dụ*, kháng sự hư hỏng do côn trùng, tuyến trùng, vi khuẩn, hoặc nấm), kháng áp lực phi sinh học tăng (*ví dụ*, kháng hạn, lạnh, nhiệt, kim loại, hoặc muối), tính chống đổ được tăng cường, tốc độ sinh trưởng được tăng cường, sinh khối tăng, sự đâm chồi tăng, đẻ nhánh tăng, thời gian ra hoa được trì hoãn, sự già hóa bị trì hoãn, số lượng hoa tăng, kiến trúc tăng đối với trồng cây mật độ cao, khả năng quang hợp được cải thiện, khối lượng rễ tăng, số lượng tế bào tăng, sức sống cây con được cải thiện, kích thước cây con được cải thiện, tốc độ phân chia tế bào tăng, hiệu quả chuyển hóa được cải thiện, và/hoặc kích thích mô phân sinh tăng

[00106] 48. Phương pháp theo phương án 46 hoặc 47, trong đó sự biểu hiện được tăng ở bộ phận của cây, tùy ý trong đó bộ phận này là hạt, lá, thân, hoa, rễ, thân, hoặc tế bào.

[00107] 49. Sử dụng phân tử ADN theo phương án bất kỳ trong số các phương án từ 1 đến 6 để: (i) làm tăng sự biểu hiện của một hoặc nhiều yếu tố được mã hóa bởi đơn vị phiên mã mà được liên kết hoạt động với trình tự khởi động ở thực vật; (ii) truyền tính trạng hữu dụng cho cây bao gồm phân tử ADN hoặc phân tử ADN tái tổ hợp, trong đó tính trạng hữu dụng này tùy ý là năng suất được cải thiện, các đặc tính thực phẩm và/hoặc thức ăn được cải thiện (*ví dụ*, chất lượng hoặc số lượng dầu, tinh bột, protein hoặc axit amin được cải thiện), hiệu quả sử dụng nitơ được cải thiện, các đặc tính sử dụng nhiên liệu sử dụng được cải thiện (*ví dụ*, sự sản xuất etanol được cải thiện), sự ra hoa được trì hoãn, không ra hoa, kháng áp lực sinh học tăng (*ví dụ*, kháng sự hư hỏng do côn trùng, tuyến trùng, vi khuẩn, hoặc nấm), kháng áp lực phi sinh học tăng (*ví dụ*, kháng hạn, lạnh, nhiệt, kim loại, hoặc muối), tính kháng đổ được tăng cường, tốc độ sinh trưởng được tăng cường, sinh khối được tăng cường, khả năng đâm chồi được tăng cường, khả năng đẻ nhánh được tăng cường, thời gian ra hoa được trì hoãn, sự già hóa được trì hoãn, số lượng hoa tăng, kiến trúc được cải thiện để trồng mật độ cao, khả năng quang hợp được cải thiện, khối lượng rễ tăng, số lượng tế bào tăng, sức sống cây con được cải thiện, kích thước cây con được cải thiện, tốc độ phân chia tế bào tăng, hiệu quả chuyển hóa được cải thiện, và/hoặc kích thích mô phân sinh tăng so với cây đối chứng thiếu phân tử ADN tái tổ hợp; (iii) thu cây hoặc hạt từ đó thể hiện tính trạng hữu dụng của (ii); hoặc

(iv) để trồng quần thể thực vật thể hiện tính trạng hữu dụng của (ii); tùy ý trong đó cây của (i), (ii), (iii), hoặc (iv) là cây ngô, đậu tương, bông, hoặc cây cải dầu; hoặc tùy ý trong đó hạt của (iii) là cây ngô, đậu tương, bông, hoặc cây cải dầu.

Ví dụ thực hiện sáng chế

[00108] Ví dụ 1.

[00109] Sự chèn yếu tố tăng cường phiên mã vào trong đích đơn ở tế bào trần của ngô bằng NHEJ được minh họa trong ví dụ này. Các tế bào trần của ngô được tạo ra bằng cách sử dụng các phương pháp đã được thiết lập (xem US20190352655, mà được kết hợp toàn bộ ở đây bằng cách viện dẫn).

[00110] Các tế bào trần của ngô được tách từ vật liệu lá từ các cây con B104, và được chuyển nạp với các oligo sợi kép và các RNP gồm protein Cas9 được tạo phức với bộ đôi tracrARN:crARN hướng đích trình tự khởi động ZmGln1-3 (trình tự crARN: UACACGUACGAUUACAACCAGUUUAGAGCUAUGCU; SEQ ID NO:5) và các oligo sợi kép được mô tả dưới đây. Phức hợp Cas9/tracrARN/crARN dẫn đến sự đứt gãy sợi kép ở ADN 66 bp ngược dòng của vị trí bắt đầu phiên mã của ZmGln1-3 (*tức là*, gen Zm00001d017958 theo bộ gen ngô B73v4 mã hóa polypeptit của SEQ ID NO:6). Ba loại oligo sợi kép được thử nghiệm: cấu hình 24 nt chứa hai đoạn lặp của trình tự yếu tố tăng cường, cấu hình 36 bp chứa ba đoạn lặp của trình tự yếu tố tăng cường, và cấu hình 48 nt chứa bốn đoạn lặp của trình tự yếu tố tăng cường. Các oligo được phosphoryl đầu 5', và chứa hai liên kết phosphorothioat tại cả hai đầu. Là đối chứng, các tế bào trần được chuyển nạp với chỉ các RNP, tức là không có các oligo bổ sung. Mỗi sự chuyển nạp được thực hiện ba lần lặp. Sau khi chuyển nạp, các tế bào được rửa, được ủ trong 48 giờ. Khi kết thúc thời gian ủ, các tế bào được thu hoạch cho sự chuẩn bị gADN and ARN.

[00111] Để phân tích các vị trí chèn, gADN được sử dụng làm khuôn mà trên đó PCR được thực hiện với các mồi gắn sườn vị trí chèn. Sau khi thu dọn hạt, các sản phẩm khuếch đại thu được được giải trình tự bằng phương pháp giải trình tự thế hệ mới. Các đoạn đọc được căn thẳng hàng với locut đích, và tỉ lệ phần trăm các đoạn đọc chứa trình tự yếu tố tăng cường được sử dụng làm đại diện cho phần tế bào mà được tích hợp trình tự tăng cường. Số liệu được tổng hợp trên Hình 1. Đối với sự chuyển nạp với yếu tố tăng cường kép, 7% số tế bào được ước lượng là đã tích hợp yếu tố tăng cường tại locut đích

Gln1-3. Đối với các tế bào được chuyển nạp với yếu tố tăng cường bộ ba hoặc bộ bốn, lần lượt là 44% và 41% số tế bào đã được tích hợp yếu tố tăng cường tại locut đích. Tất cả các số được tính trung bình qua 3 lần lặp sinh học. Số liệu này thể hiện rằng yếu tố bộ ba thực hiện tốt nhất đối với hiệu quả chèn.

[00112] Để đánh giá tác động của trình tự yếu tố tăng cường được chèn đối với mức biểu hiện của Gln1-3, ARN khối được chiết từ các tế bào trần được thu hoạch, và được chuyển đổi thành cADN. Sử dụng qRT-PCR, mức biểu hiện của ZmGln1-3 được đo so với mức biểu hiện của ZmAct1, gen tham chiếu đã biết rõ, số liệu được tổng hợp trên Hình 2. Ở các tế bào được chuyển nạp với yếu tố tăng cường kép, mức biểu hiện của ZmGln1-3 là không phân biệt được với các tế bào đối chứng được chuyển nạp với chỉ RNP. Tuy nhiên, ở các tế bào được chuyển nạp với yếu tố tăng cường bộ ba, mức biểu hiện của ZmGln1-3 là cao hơn khoảng 4,5 lần so với các đối chứng. Đối với các tế bào được chuyển nạp với yếu tố tăng cường bộ bốn, nó chỉ tăng bằng 2,7 lần. Số liệu này chỉ ra rằng cả yếu tố tăng cường bộ ba và bộ bốn đều tăng cường sự biểu hiện của gen đích khi được tích hợp vào trong trình tự khởi động, nhưng yếu tố tăng cường bộ ba hiệu quả hơn nhiều.

[00113] Ví dụ 2.

[00114] Sự phiên mã được tăng cường ở các tế bào trần của ngô tại nhiều đích gen với yếu tố tăng cường 36 bp được minh họa trong ví dụ này. Các tế bào trần của ngô được tách từ vật liệu lá từ cây con B104, và được chuyển nạp với các RNP gồm Cas9 được tạo phức với bộ đôi tracrARN:crARN, hoặc với hoặc không với các oligo yếu tố tăng cường 36-nt được phosphoryl hóa đầu 5' sợi kép (SEQ ID NO:3) với hai liên kết phosphorothioat trên cả hai đầu với trình tự yếu tố tăng cường 36 nucleotit (SEQ ID NO:3). Ba crARN được sử dụng trong thí nghiệm này crARN thứ nhất (UACACGUACGAUUACAACCAGUUUUAGAGCUAUGCU; SEQ ID NO:5) hướng đích trình tự khởi động ZmGln1-3 (*tức là*, trình tự khởi động của gen Zm00001d017958 theo bộ gen ngô B73v4 mã hóa polypeptit của SEQ ID NO:6). crARN thứ hai (UGUAUCCGUAUUUAUACGUGGUUUUAGAGCUAUGCU; SEQ ID NO:7) hướng đích trình tự khởi động ZmGln1-4 (*tức là*, trình tự khởi động của gen Zm00001d051804 theo bộ gen ngô B73v4 mã hóa polypeptit của SEQ ID NO:8). crARN thứ ba (CUCCAAGUGACCGAGCAAGAGUUUUAGAGCUAUGCU; SEQ ID NO:9) hướng đích trình tự khởi động ZmLc (*tức là*, trình tự khởi động của gen

Zm00001d026147 theo bộ gen ngô B73v4 mã hóa polypeptit của SEQ ID NO:10). Mỗi sự chuyển nạp được thực hiện ba lần lặp. Sau khi chuyển nạp, các tế bào được rửa, được ủ trong 48 giờ. Khi kết thúc thời gian ủ, các tế bào được thu hoạch cho sự chuẩn bị gADN and ARN.

[00115] Để phân tích các vị trí chèn, gADN được sử dụng làm khuôn cho PCR được thực hiện với các môi gắn sườn vị trí chèn, và sau khi thu dọn hạt, các sản phẩm khuếch đại thu được được giải trình tự bằng phương pháp giải trình tự thế hệ mới. Các đoạn đọc được căn thẳng hàng với locut đích, tỉ lệ phần trăm các đoạn đọc chứa trình tự yếu tố tăng cường được sử dụng làm đại diện cho phần tế bào mà được tích hợp trình tự tăng cường. Trên các đoạn đọc mà không có trình tự yếu tố tăng cường, số lượng các đoạn đọc với các indel do NHEJ được xác định. Hiệu quả chèn ở cả ba trường hợp là khoảng một nửa của tổng hiệu quả chỉnh sửa (kết hợp các đoạn đọc với chỉ các indel và các đoạn đọc với oligonucleotit được chèn), nhưng hiệu quả chèn tuyệt đối khác giữa các đích, và là 31% đối với ZmGln1-3, 9% đối với Gln1-4, và 12% đối với Lc.

[00116] Để đánh giá tác động của trình tự yếu tố tăng cường SEQ ID NO:3 được chèn đối với mức biểu hiện của Gln1-3, ARN khối được chiết từ các tế bào trần được thu hoạch, và được chuyển đổi thành cADN. Sử dụng qRT-PCR, mức biểu hiện của gen đích (ZmGln1-3, ZmGln1-4, hoặc ZmLc) được đo so với mức biểu hiện của ZmAct1, gen tham chiếu đã biết rõ. Các kết quả này được tổng hợp trên Hình 3, thể hiện sự tăng biểu hiện của Lc bằng >400 lần, của Gln1-3 bằng 2,3 lần, và của Gln1-4 bằng 1,9 lần do sự chèn trình tự yếu tố tăng cường ngô trong các trình tự khởi động của các gen này. Xem xét rằng chỉ phần tế bào đã tích hợp trình tự yếu tố tăng cường, như được thảo luận ở trên, sự tăng biểu hiện gen ở các tế bào với yếu tố tăng cường thậm chí cao hơn, chỉ ra rằng yếu tố tăng cường hoạt động tốt để làm tăng mức biểu hiện của nhiều đích.

[00117] Ví dụ 3.

[00118] Việc chỉnh sửa và biến nạp bộ gen ngô được minh họa trong ví dụ này. Các phôi chưa trưởng thành từ dòng ngô biểu hiện phổ biến nucleaza CRISPR/Cas ở nền B104 được bắn phá bằng các hạt vàng được phủ ARN dẫn hướng đích trình tự khởi động Gln1-3 ở gen Gln1.3 (*tức là*, trình tự khởi động của gen Zm00001d017958 theo bộ gen ngô B73v4 mã hóa polypeptit của SEQ ID NO:6) và các oligonucleotit được phosphoryl hóa đầu 5' 5' sợi kép với 2 liên kết phosphorothioat tại cả hai đầu với trình

tự yếu tố tăng cường 36 nucleotit (SEQ ID NO:3), và plasmit mã hóa gen kháng pat và protein phát huỳnh quang. Qua thực hành nuôi cấy mô tiêu chuẩn, với sự chọn lọc đối với tính kháng pat và sự phát huỳnh quang, các cây được tái sinh chứa yếu tố tăng cường 36 bp đầy đủ hoặc một phần (SEQ ID NO:3). Mô lá được lấy từ các cây được tái sinh cho sự tách chiết gADN để kiểm tra sự có mặt của sự chèn sử dụng các môi gắn sườn vị trí đích gARN. Việc tăng kích thước sản phẩm khuếch đại cho phép chọn lọc các sự kiện với sự chèn có mặt. Giải trình tự sản phẩm khuếch đại khẳng định rằng sự chèn oligo xảy ra tại các vị trí đích được dự đoán, mà đối với Gln1-3 là 66 bp ngược dòng của vị trí bắt đầu phiên mã (TSS), với tính hiệu quả nằm trong khoảng từ 0,4% đến 1,1%, như được chỉ ra trong Bảng 1.

[00119] Bảng 1. Hiệu quả của sự chèn yếu tố tăng cường bộ ba vào trong trình tự khởi động ZmGln1.3 qua sự bắn phá.

Thí nghiệm	Phôi được bắn phá	Các sự kiện được tái sinh	Các sự kiện với sự chèn	% hiệu quả
ZM035	271	7	1	0,4%
ZM037	306	4	3	1,0%
ZM041	348	6	4	1,1%
ZM072	454	15	3	0,7%
ZM081	383	3	2	0,5%

[00120] ARN được chuẩn bị từ các mẫu lá được lấy từ các thể hệ con non từ thí nghiệm bắn phá đơn được hỗ trợ để chèn yếu tố tăng cường ngô trong trình tự khởi động Gln1-3 và được chuyển đổi thành cADN. Sử dụng qRT-PCR dựa trên SYBR, mức biểu hiện của Gln1-3 và ZmAct1 đối chứng được đo. Cùng các cây cũng được đánh giá kiểu gen đối với sự có mặt của yếu tố tăng cường. Các kết quả được tổng hợp trên Hình 4, chỉ ra các cây từ các sự kiện mà không có yếu tố tăng cường được chèn có các mức

Gln1-3 thấp, tương tự với đối chứng không được biến nạp B104, trong khi các cây từ các sự kiện với yếu tố tăng cường 36 bp (SEQ ID NO:3) có mặt có các mức mà cao hơn 4-10 lần so với đối chứng.

[00121] Các cây chứa sự chèn được trồng đến khi trưởng thành, và được lai với các cây B104 kiểu dại. Quần thể T1BC1 được sinh ra theo cách này đến từ các dị hợp tử của cây đối với sự chèn yếu tố tăng cường 36 bp chiều dài đầy đủ sau đó được trồng cho phân tích khác. Từ các cây này, mô lá được lấy mẫu để đánh giá kiểu gen đối với sự chèn bằng PCR, và lá, lông sơ khởi, và mô rễ chính được thu hoạch cho sự tách chiết ARN, sự tổng hợp cADN, và qRT-PCR đối với Gln1-3 và ZmAct1. Các kết quả của các thí nghiệm này được thể hiện trên Hình 5. Các cá thể với sự chèn được so sánh với cây cùng máu mà không thừa hưởng alen chèn yếu tố tăng cường, và với các cây B104 kiểu dại được trồng kế bên các cây T1BC1. Ở cả rễ và lông sơ khởi, có sự tăng gấp 2 lần và đáng kể ($p < 0,05$, kiểm định t Student hai phía) của sự biểu hiện Gln1-3 ở các cây có sự chèn khi so sánh với các cây cùng máu mà không có sự chèn và với các cây B104. Ở các lá, không có sự chênh lệch đáng kể ở sự biểu hiện có thể được phát hiện, có thể được liên kết với thực tế rằng sự biểu hiện Gln1-3 rõ ràng rất cao ở lá khi so sánh với rễ và lông sơ khởi.

[00122] Ví dụ 4.

[00123] Ví dụ này minh họa sự tăng cường của sự biểu hiện bản phiên mã ở tế bào đậu tương. Các tế bào trần của đậu tương được tạo ra bằng cách sử dụng các phương pháp đã được thiết lập (xem US20190352655,).

[00124] Hai plasmit được chuẩn bị cho sự chuyển nạp tế bào trần. Plasmit đối chứng có cat-xet biểu hiện với trình tự khởi động gen Tfl1b đậu tương, 5' UTR, và codon bắt đầu (SEQ ID NO: 11) điều khiển sự biểu hiện của chỉ thị phát huỳnh quang GFP. Plasmit thử nghiệm khác với plasmit đối chứng chỉ bởi sự chèn của yếu tố tăng cường 36bp của SEQ ID NO: 3, sự chèn này là 484 nucleotit từ Codon bắt đầu đối với protein Tfl1b (SEQ ID NO: 12).

[00125] Bảng 2. Các mức biểu hiện chỉ thị đậu tương (các đơn vị định lượng sự phát huỳnh quang tùy ý) tại các thời điểm khác nhau sau sự chuyển nạp

Plasmit	16 giờ	24 giờ	40 giờ
đối chứng (trung bình/độ lệch chuẩn)	0,09/0,02	0,12/0,04	0,02/0,01
thử nghiệm (trung bình/độ lệch chuẩn)	0,61/0,10	0,74/0,08	0,52/0,14

[00126] Sự biểu hiện tăng của gen chỉ thị GFP được quan sát thấy ở tế bào trần của đậu tương được chuyển nạp với phương án thử nghiệm bao gồm sự chèn của yếu tố tăng cường 36bp của SEQ ID NO: 3 ở trình tự khởi động gen Tlf1b đậu tương so với các đối chứng được chuyển nạp với plasmit bao gồm trình tự khởi động gen Tlf1b đậu tương tại 16 giờ, 24 giờ, và 40 giờ sau chuyển nạp.

[00127] Độ rộng và phạm vi của sáng không nên bị giới hạn bởi phương án bất kỳ trong số các phương án làm ví dụ được mô tả ở trên, mà chỉ nên được xác định theo các điểm yêu cầu bảo hộ sau đây và các tương đương của chúng.

YÊU CẦU BẢO HỘ

1. Phân tử ADN bao gồm trình tự polynucleotit của SEQ ID NO:3.
2. Phân tử ADN theo điểm 1, trong đó phân tử ADN này là phân tử ADN được phân lập, tổng hợp, và/hoặc tái tổ hợp.
3. Phân tử ADN theo điểm 1, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị trong trình tự polynucleotit bao gồm trình tự khởi động.
4. Phân tử ADN theo điểm 3, trong đó trình tự polynucleotit của SEQ ID NO:3 và trình tự khởi động được liên kết hoạt động với đơn vị phiên mã và làm tăng sự biểu hiện của bản phiên mã hoặc protein được mã hóa bởi đơn vị phiên mã này ở tế bào thực vật so với tế bào thực vật đối chứng bao gồm trình tự khởi động trong sự vắng mặt của trình tự polynucleotit của SEQ ID NO:3.
5. Phân tử ADN theo điểm 4, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị từ khoảng 10, 20, 30, hoặc 40 cặp bazơ (bp) đến khoảng 100, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp của vị trí bắt đầu phiên mã của đơn vị phiên mã.
6. Phân tử ADN theo điểm 4, trong đó trình tự khởi động là trình tự khởi động thực vật và đơn vị phiên mã là đơn vị phiên mã thực vật, và tùy ý trong đó trình tự khởi động thực vật và đơn vị phiên mã thực vật là trình tự khởi động và đơn vị phiên mã của cây ngô hoặc cây đậu tương.
7. Tế bào thực vật bao gồm phân tử ADN theo điểm 3, trong đó tế bào thực vật này tùy ý là tế bào thực vật không thể tái sinh được mà không có khả năng được tái sinh để tạo ra cây.
8. Tế bào thực vật theo điểm 7, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị trong trình tự polynucleotit bao gồm trình tự khởi động, trong đó trình tự khởi động này là trình tự khởi động nội sinh, và trong đó trình tự khởi động nội sinh này được định vị ở vị trí tự nhiên của nó ở bộ gen của thực vật.
9. Tế bào thực vật theo điểm 7, trong đó tế bào thực vật này là tế bào cây ngô hoặc đậu tương.
10. Cây bao gồm phân tử ADN theo điểm 3.

11. Cây theo điểm 10, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị ở trong trình tự polynucleotit bao gồm trình tự khởi động, trong đó trình tự khởi động này là trình tự khởi động nội sinh, và trong đó trình tự khởi động nội sinh này được định vị ở vị trí tự nhiên của nó trong bộ gen của thực vật.
12. Bộ phận của cây bao gồm phân tử ADN theo điểm 3.
13. Bộ phận của cây theo điểm 12, trong đó bộ phận của cây là hạt.
14. Phương pháp sản xuất vật liệu thực vật, phương pháp này bao gồm bước trồng cây có yếu tố làm tăng sự biểu hiện bao gồm SEQ ID NO:3 được định vị trong trình tự khởi động, trong đó yếu tố làm tăng sự biểu hiện của SEQ ID NO:3 và trình tự khởi động được liên kết hoạt động với polynucleotit mã hóa bản phiên mã, và trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã trong cây đã nêu được tăng khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện.
15. Phương pháp theo điểm 14, trong đó việc trồng cây bao gồm ít nhất một trong số các bước gieo hạt mà nảy mầm và tạo thành cây, tưới nước cho hạt hoặc cây, và/hoặc xử lý cây hoặc hạt bằng tác nhân sinh học, chất diệt cỏ, chất diệt côn trùng, hoặc chất diệt nấm.
16. Phương pháp theo điểm 14, trong đó vật liệu thực vật bao gồm hạt và trong đó phương pháp này còn bao gồm bước thu hoạch hạt từ cây này.
17. Phương pháp theo điểm 14, trong đó yếu tố tăng cường sự biểu hiện được định vị trong trình tự khởi động nội sinh mà được liên kết hoạt động với polynucleotit mã hóa bản phiên mã nội sinh và trong đó trình tự khởi động nội sinh và polynucleotit mã hóa bản phiên mã nội sinh đã nêu được định vị ở vị trí tự nhiên của chúng trong bộ gen của thực vật.
18. Phương pháp theo điểm 14, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị từ khoảng 10, 20, 30, hoặc 40 cặp bazơ (bp) đến khoảng 100, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp từ vị trí bắt đầu phiên mã của đơn vị phiên mã.
19. Phương pháp theo điểm 17, trong đó thực vật này là cây ngô hoặc đậu tương, trong đó trình tự khởi động nội sinh và polynucleotit mã hóa bản phiên mã nội sinh là trình tự khởi động nội sinh của cây ngô hoặc đậu tương và polynucleotit mã hóa bản phiên mã

nội sinh được định vị ở vị trí tự nhiên của chúng trong bộ gen của cây ngô hoặc đậu tương.

20. Phương pháp theo điểm 14, trong đó thực vật này là cây ngô hoặc đậu tương.

21. Phương pháp sản xuất vật liệu thực vật, phương pháp này bao gồm bước chèn yếu tố làm tăng sự phiên mã của SEQ ID NO: 3 trong trình tự khởi động của cây, và trồng cây này để tạo ra vật liệu thực vật, trong đó yếu tố làm tăng sự biểu hiện và trình tự khởi động được liên kết hoạt động với polynucleotit mã hóa bản phiên mã, và trong đó sự biểu hiện của polynucleotit mã hóa bản phiên mã trong cây này được tăng lên khi so sánh với cây đối chứng thiếu yếu tố làm tăng sự biểu hiện.

22. Phương pháp theo điểm 21, trong đó việc trồng cây bao gồm ít nhất một trong số các bước gieo hạt mà nảy mầm và tạo thành cây, tưới nước cho hạt hoặc cây, và/hoặc xử lý cây hoặc hạt bằng tác nhân sinh học, chất diệt cỏ, chất diệt côn trùng, hoặc chất diệt nấm.

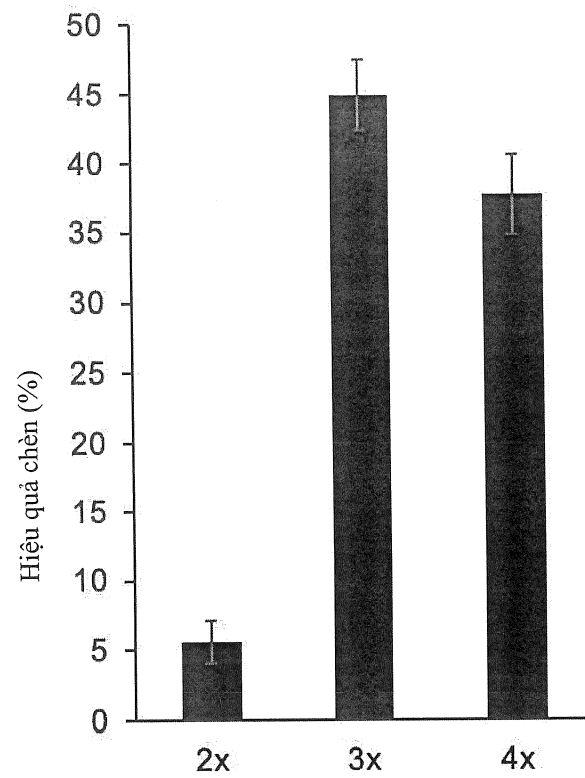
23. Phương pháp theo điểm 21, trong đó vật liệu thực vật bao gồm hạt và trong đó phương pháp này còn bao gồm bước thu hoạch hạt từ cây này.

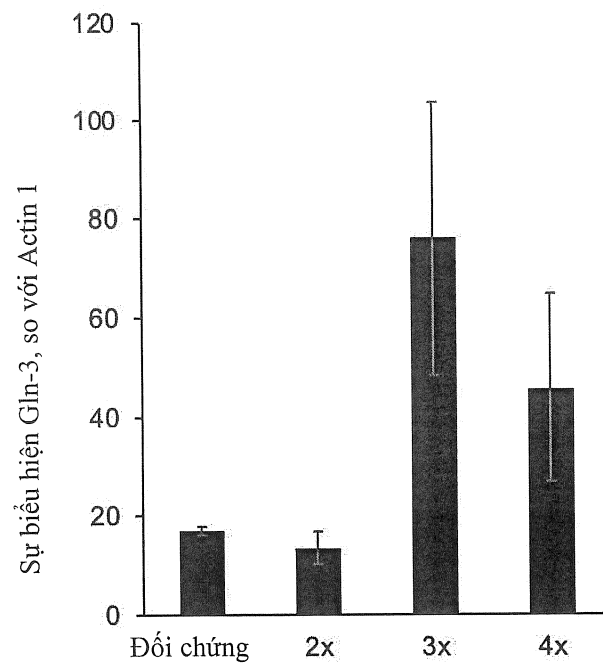
24. Phương pháp theo điểm 21, trong đó yếu tố tăng cường sự biểu hiện được định vị trong trình tự khởi động nội sinh mà được liên kết hoạt động với polynucleotit mã hóa bản phiên mã nội sinh và trong đó trình tự khởi động nội sinh và polynucleotit mã hóa bản phiên mã nội sinh đã nêu được định vị ở vị trí tự nhiên của chúng trong bộ gen của thực vật.

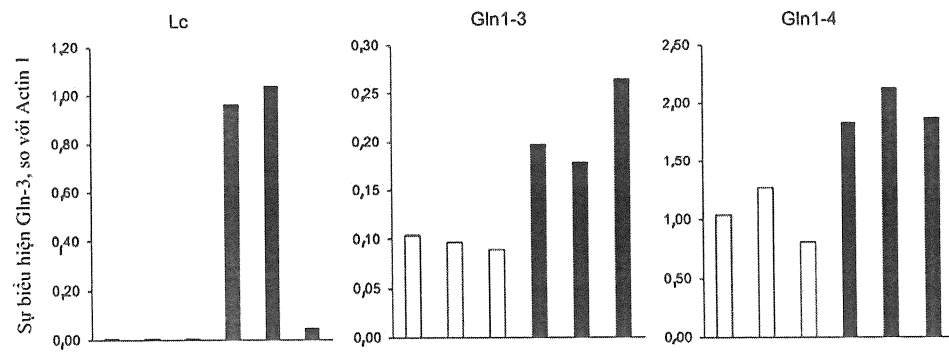
25. Phương pháp theo điểm 21, trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị khoảng 10, 20, 30, hoặc 40 cặp bazơ (bp) đến khoảng 100, 240, 300, 400, 500, 1000, 2000, 3000, hoặc 5000 bp từ vị trí bắt đầu phiên mã của đơn vị phiên mã.

26. Phương pháp theo điểm 21, trong đó thực vật này là cây ngô hoặc đậu tương, trong đó yếu tố làm tăng sự biểu hiện được định vị ở trình tự khởi động ngô hoặc đậu tương nội sinh mà được liên kết hoạt động với polynucleotit mã hóa bản phiên mã nội sinh, và trong đó trình tự khởi động nội sinh và polynucleotit mã hóa bản phiên mã nội sinh đã nêu được định vị trong vị trí tự nhiên của chúng trong bộ gen của cây ngô hoặc đậu tương.

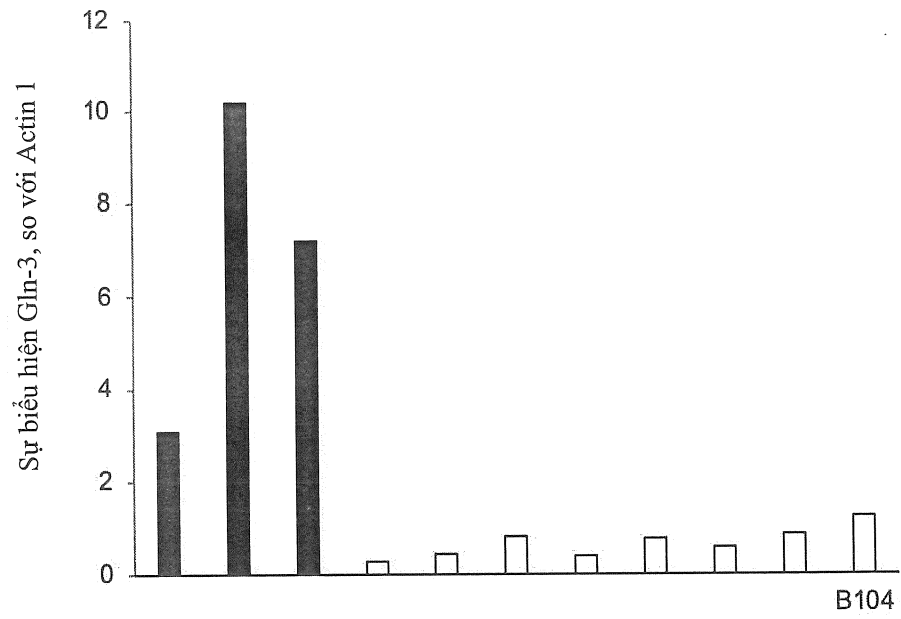
27. Mẫu sinh học bao gồm phân tử ADN theo điểm 1, trong đó mẫu sinh học này tùy ý là "không tái sinh được" và không có khả năng được tái sinh thành cây hoặc bộ phận của cây.
28. Mẫu sinh học theo điểm 27, trong đó mẫu này bao gồm bột thô hạt ngô, đậu tương, bông, hoặc hạt cải dầu.
29. Phương pháp xác định mẫu sinh học theo điểm 27 bao gồm polynucleotit bao gồm gen thực vật được cải biến trong đó trình tự polynucleotit của SEQ ID NO:3 được định vị trong trình tự polynucleotit bao gồm trình tự khởi động, bao gồm bước phát hiện sự có mặt của SEQ ID NO:3 ở mẫu sinh học này.
30. Phương pháp theo điểm 29, trong đó mẫu sinh học này thu được từ tế bào thực vật, cây, hoặc bộ phận của cây, tùy ý trong đó bộ phận của cây là hạt, và/hoặc tùy ý trong đó cây này là cây ngô, đậu tương, bông, hoặc cây cải dầu.

**HÌNH 1**

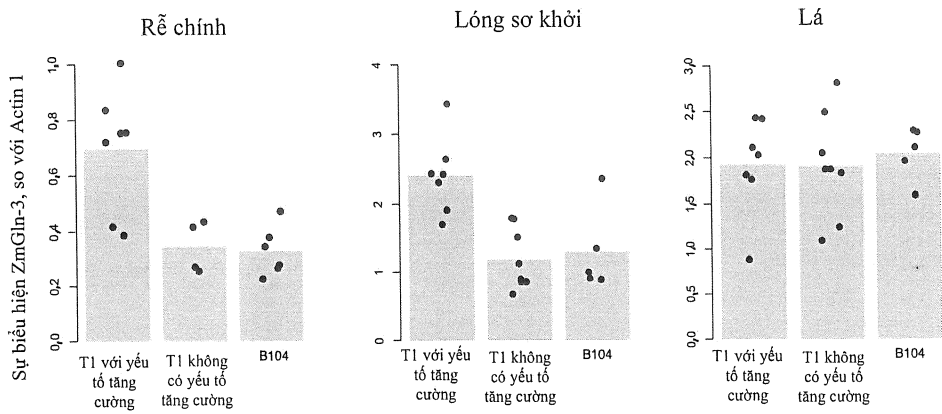
**HÌNH 2**



HÌNH 3



HÌNH 4



HÌNH 5

DANH MỤC TRÌNH TỰ

<110> Claey's, Hannes

<120> PHÂN TỬ ADN, TẾ BÀO THỰC VẬT, CÂY, BỘ PHẬN CỦA CÂY BAO GỒM PHÂN TỬ ADN NÀY VÀ PHƯƠNG PHÁP SẢN XUẤT VẬT LIỆU THỰC VẬT

<130> P13470W000

<150> US 17/249,437

<151> 2021-03-02

<150> US 63/084,150

<151> 2020-09-28

<160> 12

<170> PatentIn phiên bản 3.5

<210> 1

<211> 12

<212> ADN

<213> Nhân tạo

<220>

<223> tổng hợp

<400> 1

gtaagcgctt ac

12

<210> 2

<211> 24

<212> ADN

<213> Nhân tạo

<220>

<223> tổng hợp

<400> 2

gtaagcgctt acgtaagcgc ttac

24

<210> 3

<211> 36

<212> ADN

<213> Nhân tạo

<220>

<223> tổng hợp

<400> 3

gtaagcgctt acgtaagcgc ttacgtaagc gcttac

36

<210> 4
 <211> 48
 <212> ADN
 <213> Nhân tạo

<220>
 <223> tổng hợp

<400> 4
 gtaagcgctt acgtaagcgc ttacgtaagc gcttacgtaa gcgcttac . 48

<210> 5
 <211> 36
 <212> ARN
 <213> Nhân tạo

<220>
 <223> tổng hợp

<400> 5
 uacacguacg auuacaacca guuuuagagc uaugcu 36

<210> 6
 <211> 356
 <212> PRT
 <213> Zea mays

<400> 6

Met Ala Cys Leu Thr Asp Leu Val Asn Leu Asn Leu Ser Asp Asn Thr
 1 5 10 15

Glu Lys Ile Ile Ala Glu Tyr Ile Trp Ile Gly Gly Ser Gly Met Asp
 20 25 30

Leu Arg Ser Lys Ala Arg Thr Leu Ser Gly Pro Val Thr Asp Pro Ser
 35 40 45

Lys Leu Pro Lys Trp Asn Tyr Asp Gly Ser Ser Thr Gly Gln Ala Pro
 50 55 60

Gly Glu Asp Ser Glu Val Ile Leu Tyr Pro Gln Ala Ile Phe Lys Asp
 65 70 75 80

Pro Phe Arg Arg Gly Asn Asn Ile Leu Val Met Cys Asp Cys Tyr Thr
 85 90 95

Pro Ala Gly Glu Pro Ile Pro Thr Asn Lys Arg Tyr Asn Ala Ala Lys
 100 105 110

Ile Phe Ser Ser Pro Glu Val Ala Ala Glu Glu Pro Trp Tyr Gly Ile
 115 120 125

Glu Gln Glu Tyr Thr Leu Leu Gln Lys Asp Thr Asn Trp Pro Leu Gly
 130 135 140

Trp Pro Ile Gly Gly Phe Pro Gly Pro Gln Gly Pro Tyr Tyr Cys Gly
 145 150 155 160

Ile Gly Ala Glu Lys Ser Phe Gly Arg Asp Ile Val Asp Ala His Tyr
 165 170 175

Lys Ala Cys Leu Tyr Ala Gly Ile Asn Ile Ser Gly Ile Asn Gly Glu
 180 185 190

Val Met Pro Gly Gln Trp Glu Phe Gln Val Gly Pro Ser Val Gly Ile
 195 200 205

Ser Ser Gly Asp Gln Val Trp Val Ala Arg Tyr Ile Leu Glu Arg Ile
 210 215 220

Thr Glu Ile Ala Gly Val Val Val Thr Phe Asp Pro Lys Pro Ile Pro
 225 230 235 240

Gly Asp Trp Asn Gly Ala Gly Ala His Thr Asn Tyr Ser Thr Glu Ser
 245 250 255

Met Arg Lys Glu Gly Gly Tyr Glu Val Ile Lys Ala Ala Ile Glu Lys
 260 265 270

Leu Lys Leu Arg His Arg Glu His Ile Ala Ala Tyr Gly Glu Gly Asn
 275 280 285

Glu Arg Arg Leu Thr Gly Arg His Glu Thr Ala Asp Ile Asn Thr Phe

290

295

300

Ser Trp Gly Val Ala Asn Arg Gly Ala Ser Val Arg Val Gly Arg Glu
 305 310 315 320

Thr Glu Gln Asn Gly Lys Gly Tyr Phe Glu Asp Arg Arg Pro Ala Ser
 325 330 335

Asn Met Asp Pro Tyr Val Val Thr Ser Met Ile Ala Glu Thr Thr Ile
 340 345 350

Ile Trp Lys Pro
 355

<210> 7
 <211> 36
 <212> ARN
 <213> Nhân tạo

<220>
 <223> tổng hợp

<400> 7
 uguaucgcu uuuauacgug guuuuagagc uaugcu

36

<210> 8
 <211> 356
 <212> PRT
 <213> Zea mays

<400> 8

Met Ala Cys Leu Thr Asp Leu Val Asn Leu Asn Leu Ser Asp Thr Thr
 1 5 10 15

Glu Lys Ile Ile Ala Glu Tyr Ile Trp Ile Gly Gly Ser Gly Met Asp
 20 25 30

Leu Arg Ser Lys Ala Arg Thr Leu Pro Gly Pro Val Thr Asp Pro Ser
 35 40 45

Lys Leu Pro Lys Trp Asn Tyr Asp Gly Ser Ser Thr Gly Gln Ala Pro
 50 55 60

Gly Glu Asp Ser Glu Val Ile Leu Tyr Pro Gln Ala Ile Phe Lys Asp
 65 70 75 80

Pro Phe Arg Arg Gly Asn Asn Ile Leu Val Met Cys Asp Cys Tyr Thr
 85 90 95

Pro Ala Gly Glu Pro Ile Pro Thr Asn Lys Arg Tyr Ser Ala Ala Lys
 100 105 110

Ile Phe Ser Ser Leu Glu Val Ala Ala Glu Glu Pro Trp Tyr Gly Ile
 115 120 125

Glu Gln Glu Tyr Thr Leu Leu Gln Lys Asp Thr Asn Trp Pro Leu Gly
 130 135 140

Trp Pro Ile Gly Gly Phe Pro Gly Pro Gln Gly Pro Tyr Tyr Cys Gly
 145 150 155 160

Ile Gly Ala Glu Lys Ser Phe Gly Arg Asp Ile Val Asp Ala His Tyr
 165 170 175

Lys Ala Cys Leu Tyr Ala Gly Ile Asn Ile Ser Gly Ile Asn Gly Glu
 180 185 190

Val Met Pro Gly Gln Trp Glu Phe Gln Val Gly Pro Ser Val Gly Ile
 195 200 205

Ser Ser Gly Asp Gln Val Trp Val Ala Arg Tyr Ile Leu Glu Arg Ile
 210 215 220

Thr Glu Ile Ala Gly Val Val Val Thr Phe Asp Pro Lys Pro Ile Pro
 225 230 235 240

Gly Asp Trp Asn Gly Ala Gly Ala His Thr Asn Tyr Ser Thr Glu Ser
 245 250 255

Met Arg Lys Glu Gly Gly Tyr Glu Val Ile Lys Ala Ala Ile Glu Lys
 260 265 270

Leu Lys Leu Arg His Lys Glu His Ile Ala Ala Tyr Gly Glu Gly Asn
 275 280 285

Glu Arg Arg Leu Thr Gly Arg His Glu Thr Ala Asp Ile Asn Thr Phe
 290 295 300

Ser Trp Gly Val Ala Asn Arg Gly Ala Ser Val Arg Val Gly Arg Glu
 305 310 315 320

Thr Glu Gln Asn Gly Lys Gly Tyr Phe Glu Asp Arg Arg Pro Ala Ser
 325 330 335

Asn Met Asp Pro Tyr Val Val Thr Ser Met Ile Ala Glu Thr Thr Ile
 340 345 350

Val Trp Lys Pro
 355

<210> 9
 <211> 36
 <212> ARN
 <213> Nhân tạo

<220>
 <223> tổng hợp

<400> 9
 cuccaaguga ccgagcaaga guuuuagagc uaugcu

36

<210> 10
 <211> 615
 <212> PRT
 <213> Zea mays

<400> 10

Met Ala Leu Ser Ala Ser Arg Val Gln Gln Ala Glu Glu Leu Leu Gln
 1 5 10 15

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

Arg Ser Ile Asn Trp Ser Tyr Ala Leu Phe Trp Ser Ile Ser Asp Thr
 35 40 45

Gln Pro Gly Val Leu Thr Trp Thr Asp Gly Phe Tyr Asn Gly Glu Val
 50 55 60

Lys Thr Arg Lys Ile Ser Asn Ser Val Glu Leu Thr Ser Asp His Leu
 65 70 75 80

Val Met Gln Arg Ser Asp Gln Leu Arg Glu Leu Tyr Glu Ala Leu Leu
 85 90 95

Ser Gly Glu Gly Asp Arg Arg Ala Ala Pro Ala Arg Pro Ala Gly Ser
 100 105 110

Leu Ser Pro Glu Asp Leu Gly Asp Thr Glu Trp Tyr Tyr Val Val Ser
 115 120 125

Met Thr Tyr Ala Phe Arg Pro Gly Gln Gly Leu Pro Gly Arg Ser Phe
 130 135 140

Ala Ser Asp Glu His Val Trp Leu Cys Asn Ala His Leu Ala Gly Ser
 145 150 155 160

Lys Ala Phe Pro Arg Ala Leu Leu Ala Lys Ser Ala Ser Ile Gln Ser
 165 170 175

Ile Leu Cys Ile Pro Val Met Gly Gly Val Leu Glu Leu Gly Thr Thr
 180 185 190

Asp Thr Val Pro Glu Ala Pro Asp Leu Val Ser Arg Ala Thr Ala Ala
 195 200 205

Phe Trp Glu Pro Gln Cys Pro Thr Tyr Ser Glu Glu Pro Ser Ser Ser
 210 215 220

Pro Ser Gly Arg Ala Asn Glu Thr Gly Glu Ala Ala Ala Asp Asp Gly
 225 230 235 240

Thr Phe Ala Phe Glu Glu Leu Asp His Asn Asn Gly Met Asp Ile Glu
 245 250 255

Ala Met Thr Ala Ala Gly Gly His Gly Gln Glu Glu Glu Leu Arg Leu
 260 265 270

Arg Glu Ala Glu Ala Leu Ser Asp Asp Ala Ser Leu Glu His Ile Thr
 275 280 285

Lys Glu Ile Glu Glu Phe Tyr Ser Leu Cys Asp Glu Met Asp Leu Gln
 290 295 300

Ala Leu Pro Leu Pro Leu Glu Asp Gly Trp Thr Val Asp Ala Ser Asn
 305 310 315 320

Phe Glu Val Pro Cys Ser Ser Pro Gln Pro Ala Pro Pro Pro Val Asp
 325 330 335

Arg Ala Thr Ala Asn Val Ala Ala Asp Ala Ser Arg Ala Pro Val Tyr
 340 345 350

Gly Ser Arg Ala Thr Ser Phe Met Ala Trp Thr Arg Ser Ser Gln Gln
 355 360 365

Ser Ser Cys Ser Asp Asp Ala Ala Pro Ala Ala Val Val Pro Ala Ile
 370 375 380

Glu Glu Pro Gln Arg Leu Leu Lys Lys Val Val Ala Gly Gly Gly Ala
 385 390 395 400

Trp Glu Ser Cys Gly Gly Ala Thr Gly Ala Ala Gln Glu Met Ser Gly
 405 410 415

Thr Gly Thr Lys Asn His Val Met Ser Glu Arg Lys Arg Arg Glu Lys
 420 425 430

Leu Asn Glu Met Phe Leu Val Leu Lys Ser Leu Leu Pro Ser Ile His
 435 440 445

Arg Val Asn Lys Ala Ser Ile Leu Ala Glu Thr Ile Ala Tyr Leu Lys
 450 455 460

Glu Leu Gln Arg Arg Val Gln Glu Leu Glu Ser Ser Arg Glu Pro Ala
465 470 475 480

Ser Arg Pro Ser Glu Thr Thr Thr Arg Leu Ile Thr Arg Pro Ser Arg
485 490 495

Gly Asn Asn Glu Ser Val Arg Lys Glu Val Cys Ala Gly Ser Lys Arg
500 505 510

Lys Ser Pro Glu Leu Gly Arg Asp Asp Val Glu Arg Pro Pro Val Leu
515 520 525

Ile Met Asp Ala Gly Thr Ser Asn Val Thr Val Thr Val Ser Asp Lys
530 535 540

Asp Val Leu Leu Glu Val Gln Cys Arg Trp Glu Glu Leu Leu Met Thr
545 550 555 560

Arg Val Phe Asp Ala Ile Lys Ser Leu His Leu Asp Val Leu Ser Val
565 570 575

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

Phe Ala Gly Ser Gly Ala Val Val Pro Trp Met Ile Ser Glu Ala Leu
595 600 605

Arg Lys Ala Ile Gly Lys Arg
610 615

<210> 11

<211> 2011

<212> ADN

<213> Glycine max

<400> 11

ctttttatatt tttctgtatg tatgcatata ttatttgaaa atttaattga aatacactga 60

taatataagg attttttata ctgttggttta atcataaaac atcatatagt gaaaatttgt 120

taaattttat acaataatta ttttaaaaaa ttatctttaa aaattacttt aaaaaaatac 180

atttttggta catataaaat ttatactaac agtatataaa aattaagcta aatataaaaa 240

gtatatTTTaa aaaaattatg ttgtatataa taccaatatg accaacgaat attccatgct	300
taatcggcat cactaacttt taatcaagtt aaacgaagaa agtaggaact aaaagataaa	360
catgagcaac ttagcatTTTt aattataatg atattggaag tttaggaaag aaaacctaaa	420
atatccaaat ttagaaatat gaatatcaac ttataccaaa ataagcaact aacagTTTtca	480
gaccaacaaa tTTTaaaatc atcatctcca tatgccacat aattaataat taagtctcaa	540
attgaaataa tattatattt tcacagtgaac acgataccaa cctgcaattt atttatgcaa	600
taaacggtct atatatTTcca cTTTcaaatc ataatataaa aaaatcaata taacgaatgt	660
ctgcttaatt aattatacgt actaattaaa ctatgccacc accaattagg gctTTTTtTat	720
tttctggaac tagaatttTgt gatctTTTtaa gcacgggatg tgcattgtaca cacacctcaa	780
ccccTTTTtTc tccaacaaaa gaatgcaaac agaacttgaa taacataaaaa tttcacatgc	840
agagatgaag agggcaaaat ggtttcttga taaagaaggt cccccTTcct cTTTttaccaa	900
tattggaact tggctctgcat ctaagagtta agTTTTtatt ctgaaatgac ctcttaggca	960
ctcaaacgct tatactccat aggcgtgtct agccttgaga atttTgtggtg gagttccttt	1020
aattatgccc tcttcatcag gaatcttcca acaaaattaa aagaaagact atggctaagg	1080
gaaaatggag aatcacgcac gcacaaacag aaatttTgcat agtTgtTgTgca aacaattaag	1140
gaaggTtctc tacgcaaggc atacaacaga aatttctgca cgtccatggt actagtagta	1200
tactaaagag aatgctaaaa tTgtgggtgtt caagaaacga agcattTcct agaattttTg	1260
ccttttatgc tTTtagTTta ataataaaac ggcagaagct tgacaactTg caaaggggag	1320
ggaggaccca cacacacagg attacacatg atgctTTTTt tgacctgatc gagctgatgg	1380
gagagaaaaa actgggaaat tttatatTTTt taaatatTTTt ttacaaattt tttcagaact	1440
aggatgggat agaaataaat aaaaaaagta cacaataaga gttgctaata gaataaaaaa	1500
aattctttat tggatgagaa atatccatgc atccggtgta cagccaatca attaagaacc	1560
gattccaagg gccccattct gtgtttatga ccagcggaga gagaaagaga gtggaaaagt	1620
ggaaaggata actcgtcaca tatcatgact aatgtgattt cagcaciaaac atcccagacc	1680
caaaggcaat tgttgagatg agccgaatta agagacaata tattagtagt agtgggagga	1740
cctgcacgga tgcacttgat ttatttctag cctTTTTtTca tTTTttgttt ttcctataga	1800

acttgatatt tctctagcta aatctaaatt atgtattgca acatatttta aaggaggggc 1860
 aaaacacact cgatctacct tcaaagaaac tttaagaca ttataaatat ggagcccaca 1920
 ctcccttctt cctcagtaag atcatcaact tgtattcctc acttccttag ctccctctct 1980
 tccttggtcc tcctcttaca cctgcaggat g 2011

<210> 12
 <211> 2047
 <212> ADN
 <213> Nhân tạo

<220>
 <223> tổng hợp

<400> 12
 ctttttattt ttctgtatg tatgcatata ttatttgaat atttaattga aatacactga 60
 taatataagg attttttata ctgttggtta atcataaaac atcatatagt gaaaatttgt 120
 taaattttat acaataatta ttttaaaaaa ttatcttaaa aaattacttt aaaaaaatac 180
 attttttggt catataaaat ttatactaac agtatataaa aattaagcta aatataaaaa 240
 gtatatattaa aaaaattatg ttgtatataa taccaatatg accaacgaat attccatgct 300
 taatcggcat cactaacttt taatcaagtt aaacgaagaa agtaggaact aaaagataaa 360
 catgagcaac ttagcatttt aattataatg atattggaag ttaggaaag aaaacctaaa 420
 atatccaaat ttagaaatat gaatatcaac ttataccaaa ataagcaact aacagtttca 480
 gaccaacaaa ttttaaaatc atcatctcca tatgccacat aattaataat taagtctcaa 540
 attgaaataa tatttatatt tcacagtga acgataccaa cctgcaattt atttatgcaa 600
 taaacggtct atatattcca ctttcaaatc ataataaaa aaaatcaata taacgaatgt 660
 ctgcttaatt aattatacgt actaattaaa ctatgccacc accaattagg gcttttttat 720
 ttcttggaac tagaatttgt gatcttttaa gcacgggatg tgcattgaca cacacctcaa 780
 cccctttttc tccaacaaaa gaatgcaaac agaacttgaa taacataaaa tttcacatgc 840
 agagatgaag agggcaaaat gggttcttga taaagaaggt ccccttcct cttttaccaa 900
 tattggaact tggctctgcat ctaagagtta agtttttatt ctgaaatgac ctcttaggca 960
 ctcaaacgct tatactccat aggcgtgtct agccttgaga atttgtggtg gagttccttt 1020
 aattatgccc tcttcacag gaatcttcca acaaaattaa aagaaagact atggctaagg 1080

gaaaatggag aatcacgcac gcacaaacag aaatttgcat agtgtgtgca aacaattaag	1140
gaaggttctc tacgcaaggc atacaacaga aattttctgca cgtccatggc actagtagta	1200
tactaaagag aatgctaaaa tgtgggtgtt caagaaacga agcatttcct agaatttttg	1260
ccttttatgc ttttagttta ataataaaac ggacagaagct tgacaacttg caaaggggag	1320
ggaggacca cacacacagg attacacatg atgctttttt tgacctgatc gagctgatgg	1380
gagagaaaaa actgggaaat tttatatatt taaatatatt ttacaaattt tttcagaact	1440
aggatgggat agaaataaat aaaaaaagta cacaataaga gttgctaata gaataaaaaa	1500
aattctttat tggatgagaa atatgtaagc gcttacgtaa gcgcttacgt aagcgcttac	1560
ccatgcatcc ggtgtacagc caatcaatta agaaccgatt ccaagggcc cttctgtgt	1620
ttatgaccag cggagagaga aagagagtgg aaaagtggaa aggataactc gtcacatatc	1680
atgactaatg tgatttcagc acaaacatcc cagacccaaa ggcaattgtt gagatgagcc	1740
gaattaagag acaatatatt agtagtagtg ggaggacctg cacggatgca cttgatttat	1800
ttctagcctt ttttcatttt ttgtttttcc tatagaactt gtattttctc tagctaaatc	1860
taaattatgt attgcaacat attttaaagg aggggcaaaa cacactcgat ctaccttcaa	1920
agaaactttt aagacattat aaatatggag cccacactcc cttcttcctc agtaagatca	1980
tcaacttgta ttcctcactt ccttagctcc tcctcttcct tgttcctcct cttacacctg	2040
caggatg	2047