(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) **1-0052285**

(51)[2020.01] **A01H 1/00**; C12Q 1/68; C12N 9/22; C12N 15/09; C12N 15/63

(13) **B**

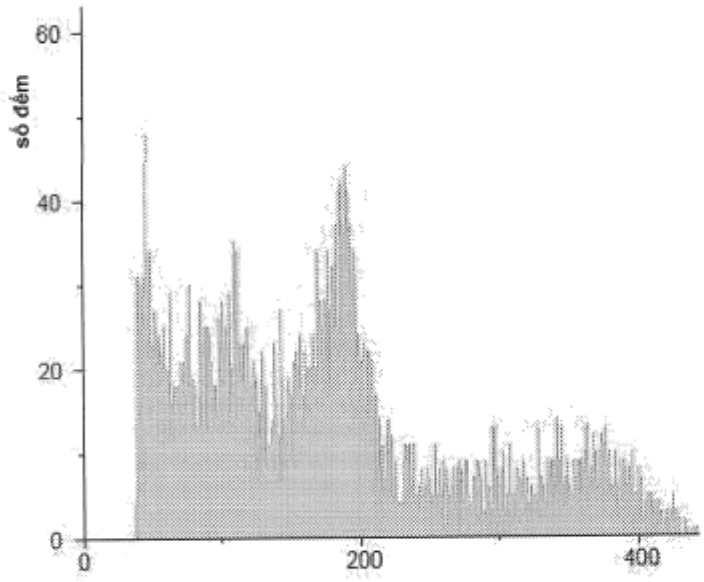
(21) 1-2022-02316
(22) 04/12/2017
(62) 1-2019-03476
(86) PCT/US2017/064512 04/12/2017
(87) WO 2018/102816 07/06/2018
(30) 62/429,260 02/12/2016 US
(45) 27/10/2025 451
(43) 25/07/2022 412A
(73) SYNGENTA CROP PROTECTION AG (CH)
Rosentalstrasse 67, 4058 Basel, Switzerland
(72) KELLIHER, Timothy (US); QUE, Qiudeng (US).
(74) Công ty TNHH Ban Ca (BANCA)

(54) PHƯƠNG PHÁP CHỈNH SỬA ADN HỆ GEN Ở CÂY HAI LÁ MẦM

(21) 1-2022-02316

(57) Sáng chế đề cập đến phương pháp chỉnh sửa ADN hệ gen ở cây hai lá mầm, bao gồm các bước: a) thu lấy cây hai lá mầm thứ nhất bao gồm đột biến trong gen histon 3 của tâm động (CENH3) và tùy ý gen chuyển hoán đổi đuôi CENH3 ở ngô, trong đó cây hai lá mầm thứ nhất biểu hiện enzym cải biến ADN và tùy ý ít nhất một axit nucleic hướng dẫn; b) thu lấy cây hai lá mầm thứ hai, trong đó cây hai lá mầm thứ hai bao gồm ADN hệ gen cây hai lá mầm cần được chỉnh sửa; c) thụ phấn cây hai lá mầm thứ nhất bằng phấn hoa từ cây hai lá mầm thứ hai; và d) chọn lọc ít nhất một con cháu đơn bội được tạo ra bởi sự thụ phấn của bước (c) trong đó con cháu đơn bội bao gồm hệ gen của cây hai lá mầm thứ hai chứ không phải cây hai lá mầm thứ nhất, và hệ gen của con cháu đơn bội đã được cải biến bằng enzym cải biến ADN và tùy ý ít nhất một axit nucleic hướng dẫn được chuyển bởi cây hai lá mầm thứ nhất; trong đó đột biến trong gen CENH3 được chọn từ nhóm bao gồm đột biến mất chức năng, đột biến mất chức năng một phần, đột biến dịch khung được phục hồi, và đột biến mất đoạn trong khung đọc.

Hình 1



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực công nghệ sinh học thực vật, cụ thể là công nghệ sinh học nông nghiệp và chỉnh sửa gen, cũng như là gây giống cây trồng. Đối tượng được bộc lộ liên quan đến việc sử dụng dòng kích tạo đơn bội (dù hiện có hay được tạo ra) và biến nạp dòng kích tạo đơn bội sao cho nó chứa ADN mã hóa cho bộ máy tế bào có khả năng chỉnh sửa gen.

Tình trạng kỹ thuật của sáng chế

Sự gây đột biến hướng đích (còn gọi là "chỉnh sửa gen") là công nghệ rất quan trọng để gây giống cây trồng. Hiện nay có nhiều phương pháp để chỉnh sửa các đích gen cụ thể, bao gồm CRISPR, TALEN, meganucleaza, và ngón tay kẽm. Một phương pháp để đưa bộ máy chỉnh sửa vào trong cây trồng là sử dụng sự biến nạp *Agrobacterium* hoặc súng bắn hạt gen của mô cây trồng. Trong sự biến nạp này, ADN mã hóa cho bộ máy chỉnh sửa (ví dụ, CAS9 và ARN dẫn) được đưa vào thể chai, hạt hoặc mô phôi của cây trồng. Sau đó cây trồng đã được biến nạp ổn định ("biến cố") được thu hồi, tùy ý với sự trợ giúp của thành phần đánh dấu chọn lọc. Nhưng vì việc nuôi cấy mô phụ thuộc vào kiểu gen, con đường này không hiệu quả với tất cả cây trồng, hoặc thậm chí tất cả các giống cây trồng mà nó có hiệu quả với. Chúng được biết đến dưới dạng cây trồng hoặc giống chống lại sự biến nạp. Các cây trồng hoặc giống này có thể được đánh giá cao về năng suất nhưng thách thức đối với công nghệ sinh học là chúng không thể được biến nạp và do đó không thể được chỉnh sửa trực tiếp thông qua sự biến nạp. Đối với các giống chống lại sự biến nạp, một trong hai phương pháp thay thế có thể được sử dụng để đưa các đột biến mong muốn vào. Thứ nhất, người ta có thể đưa vào sự chỉnh sửa thông qua sự du nhập tính trạng. Con đường này tốn tiền, công sức, và thời gian. Cũng có nghĩa là sự không tinh khiết của sản phẩm cuối cùng do sự liên kết di truyền, tức là, có khối liên kết xung quanh sự chỉnh sửa được du nhập, chứa gen và alen từ dòng cho có thể biến nạp. Sự liên kết này có thể là vấn đề nếu bất kỳ gen hoặc alen nào trong số đó ảnh hưởng đến năng suất của dòng chống lại sự biến nạp (còn có thể được đề cập đến dưới dạng "dòng cao cấp"). Thứ hai, người ta có thể đưa bộ máy chỉnh sửa tạm thời

vào cây trồng đang sinh trưởng mà không nuôi cấy mô, chẳng hạn như nhúng hoa để biến nạp *Arabidopsis*. Thách thức là phải đảm bảo việc chỉnh sửa kết thúc ở các tế bào cấu thành dòng mầm, do đó chúng được truyền cho hạt con cháu. Có một số ít các phương pháp thông thường hoặc đã được thiết lập để thực hiện điều này ở cây trồng.

Ở đây chúng tôi trình bày về phương pháp mới để tạm thời đưa bộ máy chỉnh sửa vào trong quá trình kích tạo đơn bội. Sự kích tạo đơn bội ("HI") là loại hiện tượng thực vật đặc trưng bởi sự mất một bộ nhiễm sắc thể của bố mẹ (các nhiễm sắc thể từ bố mẹ tạo đơn bội) từ phôi ở một số thời điểm trong hoặc sau khi thụ tinh, thường là trong quá trình phát triển phôi sớm. Sự kích tạo đơn bội còn được gọi là sự sinh sản đơn tính cái nếu dòng kích tạo được sử dụng làm thể đực trong phép lai, hoặc sinh sản đơn tính đực nếu dòng kích tạo được sử dụng làm thể cái trong phép lai. Sự kích tạo đơn bội đã được quan sát thấy ở nhiều loài thực vật, chẳng hạn như cây lúa miến, cây lúa mạch, cây lúa mỳ, cây ngô, *Arabidopsis*, và nhiều loài khác.

Thông thường, trong quá trình kích tạo đơn bội, cả hai dòng bố mẹ dùng trong phép lai kích tạo đều là thể lưỡng bội, do đó giao tử của chúng (tế bào giao tử cái và tế bào giao tử đực) là thể đơn bội. Sự kích tạo đơn bội thường là tính trạng có độ thâm nhập từ trung bình đến thấp của dòng kích tạo, do đó cây con thu được, tùy thuộc vào loài hoặc tình trạng, có thể là thể lưỡng bội (nếu sự mất hệ gen không xảy ra) hoặc thể đơn bội (nếu sự mất hệ gen thực sự xảy ra). Nếu dòng bố mẹ mà được lai với thể kích tạo đơn bội không phải là thể lưỡng bội, mà là thể tứ bội, thể lục bội, hoặc cây trồng khác có mức bội thể cao hơn, thuật ngữ sự kích tạo đơn bội là cách dùng sai thuật ngữ, vì cây con "đơn bội" được tạo ra sẽ có số lượng nhiễm sắc thể của giao tử, và do đó không thực sự là thể đơn bội, mà là thể lưỡng bội (nếu bố mẹ là thể tứ bội) hoặc thể tam bội (nếu bố mẹ là thể lục bội) và v.v.. Do đó, như dùng trong bản mô tả này, "thể đơn bội" mang một nửa số nhiễm sắc thể của bố hoặc mẹ; do đó thể đơn bội của sinh vật lưỡng bội (ví dụ, cây ngô) thể hiện có một bộ nhiễm sắc thể; thể đơn bội của sinh vật tứ bội (ví dụ, cỏ hoang) thể hiện có hai bộ nhiễm sắc thể; thể đơn bội của sinh vật lục bội (ví dụ, cây lúa mỳ) thể hiện có ba bộ nhiễm sắc thể.

Sự kích tạo đơn bội có thể xảy ra trong quá trình tự thụ phấn hoặc sự lai chéo của hai dòng trong cùng loài, hoặc nó có thể xảy ra trong quá trình lai xa, khi nó có thể được

xem là hàng rào lai hóa, ngăn chặn sự tạo thành của thể lai giữa các loài. Ở cây ngô, phương pháp kích tạo đơn bội thường được sử dụng nhất là thông qua việc sử dụng dòng đực thể kích tạo đơn bội cùng loài, mà chủ yếu được kích hoạt bởi sự tái sắp xếp, các đột biến, và/hoặc sự tái tổ hợp, đột biến cài xen, hoặc đột biến mất trong khu vực của nhiễm sắc thể 1, cụ thể là gen *MATRILINEAL* (*MATL*), còn gọi là *NOT LIKE DAD1* (*NLD1*) và *PHOSPHOLIPASE A1* (*PLA1*) (với ngoại lệ đáng chú ý của sự kích tạo đơn bội kiểu *ig*, mà là kết quả của đột biến trong gen *INDETERMINATE GAMETOPHYTE1* trên nhiễm sắc thể 3). Ở cây lúa mỳ, phương pháp kích tạo đơn bội phổ biến nhất là bằng cách lai xa với phấn ngô – bất kể kiểu gen hoặc dòng của bố mẹ, phương pháp này có hiệu quả với hầu hết các cây lúa mỳ bất kỳ được lai bởi hầu hết phấn ngô bất kỳ.

Dòng ngô HI chứa locut tính trạng định lượng ("QTL") trên nhiễm sắc thể 1 chịu trách nhiệm đối với ít nhất là 66% sự biến đổi trong sự kích tạo đơn bội. QTL gây ra sự kích tạo đơn bội ở tỉ lệ khác nhau khi nó được du nhập vào các nền tảng khác nhau. Tất cả các dòng kích tạo đơn bội ngô dùng trong công nghiệp hạt giống đều có nguồn gốc từ dòng HI nền móng, gọi là Stock6, và đều có đột biến QTL nhiễm sắc thể 1 thể kích tạo đơn bội.

Ở cây ngô, hạt hoặc phôi đơn bội được sản xuất đặc hiệu bằng cách thực hiện phép lai giữa thể đực kích tạo đơn bội (tức là, "phấn hoa kích tạo đơn bội") và gần như là bông bất kỳ mà người ta chọn, bông này có thể thuộc phép lai cùng dòng, thể lai bất kỳ, hoặc chất mầm nguyên sinh khác. Thể đơn bội được tạo ra khi ADN phấn hoa kích tạo đơn bội không được truyền và/hoặc được duy trì hoàn toàn qua các lần phân chia tế bào đầu tiên của phôi. Kiểu hình thu được không thâm nhập hoàn toàn, với một số noãn chứa phôi đơn bội, và các noãn khác chứa phôi lưỡng bội, phôi lệch bội, phôi khảm, hoặc phôi thui. Nhân đơn bội có phôi mà chỉ chứa ADN của mẹ cộng với nội nhũ tam bội bình thường. Sau sự kích tạo đơn bội, phôi hoặc hạt đơn bội thường bị cô lập khỏi thể lưỡng bội và các anh em lệch bội bằng cách sử dụng sàng lọc thành phần đánh dấu kiểu hình hoặc di truyền và trồng hoặc nuôi cấy thành cây đơn bội. Sau đó các cây trồng này được biến đổi tự nhiên hoặc thông qua thao tác hóa học (ví dụ, bằng cách sử dụng chất kháng vi ống chẳng hạn như colchixin) thành cây đơn bội kép ("DH") mà sau đó tạo ra hạt lai cùng dòng.

Việc gây giống cây trồng được làm cho thuận lợi bằng cách sử dụng cây đơn bội kép (DH). Sự sản xuất của cây DH giúp cho người nhân giống cây trồng thu được dòng lai cùng dòng mà không lai cùng dòng nhiều thế hệ, do đó giảm thời gian cần để tạo ra cây đồng hợp tử. Cây DH cung cấp công cụ vô giá cho người nhân giống cây trồng, cụ thể là để tạo ra các dòng lai cùng dòng, lập bản đồ QTL, biến đổi tế bào chất, sự du nhập tính trạng, và sàng lọc F2 để cải thiện tính trạng năng suất cao. Tiết kiệm được nhiều thời gian khi dòng đồng hợp tử về cơ bản được tạo ra trong một thế hệ, không cần dòng dõi hạt đơn lẻ qua nhiều thế hệ (lai cùng dòng thông thường). Cụ thể là, vì cây DH hoàn toàn đồng hợp tử, chúng rất dễ tuân theo nghiên cứu di truyền định lượng. Sự sản xuất hạt đơn bội có tính quyết định đối với quy trình gây giống đơn bội kép. Hạt đơn bội được tạo ra trên chất mầm nguyên sinh theo dòng mẹ khi thụ tinh với phấn từ thể kích tạo theo dòng mẹ, chẳng hạn như Stock 6 và các dòng phát sinh từ Stock 6.

Ở đây, chúng tôi mô tả phương pháp mới trong đó quy trình kích tạo đơn bội in vivo có thể được kết nạp để tạm thời đưa bộ máy chỉnh sửa vào chất mầm nguyên sinh bất kỳ bằng cách bao gồm nó trong bố mẹ tạo đơn bội, được tích hợp ổn định dưới dạng gen chuyển, hoặc được biểu hiện tạm thời. Sự chỉnh sửa đồng thời cộng với sự kích tạo đơn bội có thể được thực hiện ở gần như là bất kỳ cây trồng nào thông qua lai xa hoặc sự kích tạo đơn bội mới ví dụ như thông qua đột biến CENH3 (tức là, thể kích tạo đơn bội được cải biến CENH3; xem tài liệu, ví dụ, tài liệu công bố đơn WO 2017/004375, được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó) hoặc thông qua phun dịch lỏng (xem tài liệu đơn sáng chế PCT số PCT/US2016/62548, được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó). Chúng tôi trình bày các ví dụ về HI ở cây ngô, cả ngô làm thức ăn gia súc và ngô ngọt, bằng cách sử dụng thể đực kích tạo đơn bội làm dòng cho chỉnh sửa. Ngoài ra, chúng tôi trình bày các ví dụ về HI ở *Arabidopsis* bằng cách sử dụng các dòng kích tạo đơn bội được cải biến CENH3.

Chúng tôi cũng trình bày các ví dụ về HI ở cây lúa mỳ bằng cách sử dụng phấn ngô làm dòng cho chỉnh sửa trong phép lai xa. Ở cây lúa mỳ, cây lúa nước, cây lúa mạch, cây thuộc họ cải bắp, và các cây trồng khác, con đường để kích tạo đơn bội là sử dụng thể cho phấn mà kích tạo thể đơn bội thông qua lai xa. Ví dụ, người ta có thể sử dụng phấn ngô trên cây lúa mỳ, phấn kê trên cây lúa mỳ, phấn lúa mạch trên các loài lúa mạch khác, hoặc phương pháp lai xa khác bất kỳ. Trong các trường hợp kích tạo đơn bội theo

dòng mẹ ưu tiên là dòng đực chứa bộ máy chỉnh sửa, vì nó là ADN đực (có nguồn gốc từ phấn) mà được loại bỏ trong quy trình kích tạo đơn bội. Trong trường hợp của sự kích tạo đơn bội đực, ví dụ như trong hệ thống *ig1* ở cây ngô hoặc thông qua CENH3 đã được biến đổi ở cây trồng bất kỳ (mà có thể hoạt động thông qua thể đực hoặc thể cái), bộ máy chỉnh tốt nhất là có mặt ở thể cái, vì nhiễm sắc thể thể cái bị loại bỏ trong quy trình kích tạo đơn bội.

Trong sự chỉnh sửa đồng thời cộng với sự kích tạo đơn bội, mục đích là chỉnh sửa nhanh chóng và tiết kiệm chi phí các cây trồng và dòng cao cấp ("các dòng dự định chỉnh sửa") mà không nuôi cấy mô. Dòng tiếp nhận việc chỉnh sửa có thể là chất mầm nguyên sinh cao cấp, và bản thân bộ máy chỉnh sửa sẽ bị loại trừ trong quy trình kích tạo đơn bội. Cùng lúc đó, dòng đơn bội kép đã chỉnh sửa được tạo ra.

Bản chất kỹ thuật của sáng chế

Tính chống lại việc nuôi cấy mô là thách thức chủ yếu đối với việc chỉnh sửa dòng cao cấp nhanh ở các cây trồng. Sử dụng dòng kích tạo đơn bội để chuyển bộ máy gây đột biến hướng đích vào dòng cao cấp và đồng thời kích tạo thể đơn bội là cách khắc phục trở ngại chính này. Chương trình gây giống thế hệ tiếp theo có thể phụ thuộc vào quy trình này.

Bộ máy chỉnh sửa được chuyển thông qua dòng kích tạo. Bộ máy chỉnh sửa thường gặp nhất là protein liên kết ADN trong một số trường hợp được kết hợp với ARN và trong một số trường hợp còn được kết hợp với ADN. ADN, ARN, và protein mà tạo thành bộ máy chỉnh sửa được mã hóa bởi và có mặt trong dòng kích tạo vì chúng đã được cài xen ổn định trong thể kích tạo, ví dụ, thông qua sự biến nạp qua trung gian sự bắn phá hoặc agrobacterium. Trong các ví dụ khác, bộ máy chỉnh sửa được đưa vào tạm thời (thông qua việc dùng bên ngoài) hoặc được biểu hiện tạm thời trong thể giao tử trước khi thụ tinh. Sau khi thụ tinh, việc chỉnh sửa được thực hiện bằng bộ máy chỉnh sửa trong gen đích không ở thể kích tạo trước hoặc trong quá trình loại bỏ của nhiễm sắc thể thể kích tạo. Kết quả là phôi hoặc cây trồng hoặc hạt đơn bội mà chứa bộ nhiễm sắc thể chỉ từ bố mẹ không phải thể kích tạo, trong đó bộ nhiễm sắc thể đó chứa các trình tự ADN mà đã được chỉnh sửa. Thể đơn bội đã được chỉnh sửa này có thể được xác định, sinh trưởng, và nhiễm sắc thể của chúng nhân đôi, tốt hơn là bằng colchixin

hoặc chất ức chế nguyên phân khác. Sau đó dòng này có thể được sử dụng trực tiếp trong các chương trình gây giống xuôi dòng.

Theo một phương án, sáng chế đề xuất phương pháp chỉnh sửa ADN hệ gen của cây trồng. Phương pháp này được thực hiện bằng cách lấy cây trồng thứ nhất, mà là cây trồng kích tạo đơn bội và cũng mã hóa vào ADN của nó bộ máy cần để thực hiện việc chỉnh sửa (ví dụ, enzym Cas9 và ARN dẫn) và sử dụng phấn của cây trồng thứ nhất đó để thụ phấn cây trồng thứ hai. Cây trồng thứ hai là cây trồng cần được chỉnh sửa. Từ việc thụ phấn đó, con cháu (ví dụ, phôi hoặc hạt) được tạo ra; ít nhất là một trong số chúng là hạt đơn bội. Hạt đơn bội này chỉ chứa các nhiễm sắc thể của cây trồng thứ hai; các nhiễm sắc thể của cây trồng thứ nhất đã không còn nữa (đã bị loại bỏ, bị mất hoặc đã bị thoái hóa), nhưng trước khi làm vậy, nhiễm sắc thể của cây trồng thứ nhất cho phép bộ máy chỉnh sửa gen được biểu hiện. Theo cách khác, và không muốn bị ràng buộc bởi lý thuyết, cây trồng thứ nhất chuyển bộ máy chỉnh sửa đã được biểu hiện khi thụ phấn thông qua ống phấn. Hoặc, trong trường hợp mà dòng kích tạo đơn bội là thể cái trong phép lai, tế bào giao tử cái của cây trồng kích tạo đơn bội chứa bộ máy chỉnh sửa mà có mặt và có lẽ đã được biểu hiện, khi thụ tinh với hạt phấn "kiểu dại" hoặc không kích tạo đơn bội. Thông qua con đường bất kỳ trong số các con đường này, cây con đơn bội thu được bằng phép lai này cũng có hệ gen được chỉnh sửa.

Theo một khía cạnh, bộ máy chỉnh sửa là enzym cải biến ADN bất kỳ, nhưng tốt hơn là nucleaza định hướng vị trí. Nucleaza định hướng vị trí tốt hơn là dựa trên CRISPR, nhưng cũng có thể là meganucleaza, nucleaza tác động giống chất kích hoạt phiên mã (transcription-activator like effector nuclease - TALEN), hoặc nucleaza ngón tay kẽm. Nucleaza dùng trong sáng chế có thể là Cas9, Cfp1, dCas9-FokI, FEN1-FokI khảm. Theo một khía cạnh, enzym cải biến ADN là enzym chỉnh sửa bazơ định hướng vị trí chẳng hạn như Cas9-xytiđin đeaminaza hoặc Cas9-ađenin đeaminaza, trong đó Cas9 có thể có một trong hai hoặc cả hai hoạt tính nucleaza của nó bị bất hoạt, tức là Cas9 nickaza khảm (nCas9) hoặc Cas9 mất hoạt tính (dCas9) dung hợp với xytiđin đeaminaza hoặc ađenin đeaminaza. ARN dẫn tùy chọn nhắm đích vào hệ gen ở vị trí đặc hiệu dự kiến được chỉnh sửa. Theo một khía cạnh, ARN dẫn tùy chọn có chứa trình tự có 18-21 nucleotit tương đồng với bất kỳ trong số các SEQ ID NO: 2, 4, 8, 21, 23, 25, 29, 32, và 33.

Ngay khi thu được cây con đơn bội đã được chỉnh sửa, một cách tùy ý nhiễm sắc thể của nó có thể được nhân đôi bằng tác nhân gây nhân đôi nhiễm sắc thể (ví dụ colchixin, pronamit, đithipyr, trifluralin, hoặc chất kháng vi ống đã biết khác).

Theo một phương án, cây trồng thứ nhất là cây một lá mầm hoặc cây hai lá mầm. Các khía cạnh của cây trồng thứ nhất bao gồm cây ngô, cây lúa mỳ, cây lúa, cây lúa mạch, cây yến mạch, cây cao lương tiểu hắc mạch, cây lúa miến, cây kê ngọc trai, cỏ teosinte, cây tre, cây mía, cây măng tây, cây hành, và cây tỏi. Theo một phương án, cây trồng thứ hai là cây một lá mầm hoặc cây hai lá mầm. Các khía cạnh của cây trồng thứ hai bao gồm cây ngô, cây lúa mỳ, cây lúa, cây lúa mạch, cây yến mạch, cây cao lương tiểu hắc mạch, cây lúa miến, cây kê ngọc trai, cỏ teosinte, cây tre, cây mía, cây măng tây, cây hành, và cây tỏi. Theo một phương án, cây trồng thứ nhất là cây một lá mầm hoặc cây hai lá mầm thuộc loài khác với cây trồng thứ hai. Ví dụ, theo một khía cạnh, cây trồng thứ nhất là cây ngô và cây trồng thứ hai là cây lúa mỳ. Theo khía cạnh khác, cây trồng thứ nhất là cây lúa mỳ và cây trồng thứ hai là cây ngô. Theo phương án khác, cây trồng thứ nhất là cây ngô được chọn và/hoặc có nguồn gốc từ dòng Stock 6, RWK, RWS, UH400, AX5707RS, NP2222-*matl*, hoặc dòng HI bất kỳ trong số một vài dòng HI đã biết khác. Theo phương án khác nữa, cây trồng thứ nhất có chứa đột biến ở gen *CENH3*, gen *ig1*, hoặc đột biến khác mang lại hệ thống kích tạo đơn bội theo dòng cha. Theo phương án khác, cây trồng thứ nhất là cây lúa có gen *MATL* được cải biến hoặc được làm bất hoạt mà làm nó trở thành dòng kích tạo đơn bội.

Theo phương án khác, cây trồng thứ nhất không nhất thiết là thể kích tạo đơn bội, mà cây trồng thứ nhất có chứa gen cần để mã hóa cho bộ máy chỉnh sửa gen. Theo phương án này, sự kích tạo đơn bội được thực hiện bằng cách dùng hợp chất trong quá trình, ngay trước khi, hoặc ngay sau khi thụ phấn. Theo một khía cạnh, hợp phần có chứa lipit hoặc chất ức chế phospholipaza. Theo khía cạnh khác, hợp phần có chứa metyl alpha-linolenoyl flophosphonat (MALFP), etyl este của axit linoleic (LLAEE), axit linoleic (LLA), dầu ngô, đistearoyl-phosphatiđyl cholin (DSPC), metyl arachiđonyl flophosphonat (MAFP), palmityl triflometylketon (PACOCF3), arachiđonyl triflometylketon (AACOCF3), manoalit, etyl este của axit linolenic (LNAEE), etyl este của axit linolenic (LNAEE), metyl este của axit oleic (OAME), etyl este của axit oleic (OAEE), etyl este của axit palmitic (PAEE), etyl este của axit palmitoleic (PLAEE), dầu

hạt lanh, dầu ngô, axit alpha-Linolenic (aLNA), axit gamma-linolenic (gLNA), axit oleic, axit arachiđonic, axit stearic, axit linoleic liên hợp 9(Z)-11(E), hoặc 2-oleoyl-1-palmitoyl-sn-glyxero-3-phospho-etanolamin.

Mô tả vắn tắt các trình tự trong danh mục trình tự

SEQ ID NO: 1 là trình tự nucleotit của véc-tơ 23396.

SEQ ID NO: 2 là trình tự nucleotit mã hóa cho trình tự gARN để chỉnh sửa VLHP1 ở cây ngô.

SEQ ID NO: 3 là trình tự nucleotit của véc-tơ 23399.

SEQ ID NO: 4 là trình tự gARN để chỉnh sửa GW2-2 ở cây ngô.

SEQ ID NO: 5 là trình tự nucleotit của véc-tơ 22808, có chứa cấu trúc TALEN.

SEQ ID NO: 6 là trình tự đích của TALEN của 22808.

SEQ ID NO: 7 là trình tự nucleotit của véc-tơ 23123 có chứa cấu trúc Cas9.

SEQ ID NO: 8 là gARN để chỉnh sửa *MATL* ở cây ngô.

SEQ ID NO: 9 là trình tự nucleotit của phần liên quan của *MATL* trong NP2222.

SEQ ID NO: 10 là trình tự nucleotit của phần liên quan của *MATL* trong Stock6.

SEQ ID NO: 11 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350333-3 Alen 1.

SEQ ID NO: 12 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350333-3 Alen 2.

SEQ ID NO: 13 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350344-2 Alen 1.

SEQ ID NO: 14 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350344-2 Alen 2.

SEQ ID NO: 15 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350343-1 Alen 1.

SEQ ID NO: 16 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350328-1 Alen 1.

SEQ ID NO: 17 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350337-2 Alen 1.

SEQ ID NO: 18 là trình tự nucleotit của phần liên quan của *MATL* trong USR01350337-2 Alen 2.

SEQ ID NO: 19 là trình tự nucleotit của cADN *MATL* kiểu dại.

SEQ ID NO: 20 là trình tự nucleotit của véc-tơ 23397.

SEQ ID NO: 21 là trình tự gARN để chỉnh sửa VLHP2 ở cây ngô.

SEQ ID NO: 22 là trình tự nucleotit của véc-tơ 23398.

SEQ ID NO: 23 là trình tự gARN để chỉnh sửa GW2-1 ở cây ngô.

SEQ ID NO: 24 là trình tự nucleotit của véc-tơ 23763.

SEQ ID NO: 25 là trình tự gARN của VLHP1 ở cây lúa mỳ.

SEQ ID NO: 26 là trình tự đích VLHP ở cây lúa mỳ của TaVLHP2.

SEQ ID NO: 27 là trình tự đích VLHP ở cây lúa mỳ của TaVLHP3.

SEQ ID NO: 28 là trình tự đích ở ZmVLHP2-03 để chỉnh sửa.

SEQ ID NO: 29 là trình tự được chỉnh sửa ở ZmVLHP2-03.

SEQ ID NO: 30 là trình tự khuôn thể cho sửa chữa để tạo ra đột biến E149L ở ZmPYL-D.

SEQ ID NO: 31 là trình tự nucleotit của véc-tơ 23136.

SEQ ID NO: 32 là gARN của véc-tơ 23136.

SEQ ID NO: 33 là trình tự nucleotit của gen PLA ở cây lúa Os03g27610.

SEQ ID NO: 34 là trình tự nucleotit của véc-tơ 24038.

SEQ ID NO: 35 là trình tự nucleotit của véc-tơ 24039.

SEQ ID NO: 36 là trình tự nucleotit của véc-tơ 24079.

SEQ ID NO: 37 là trình tự nucleotit của véc-tơ 24091.

SEQ ID NO: 38 là trình tự nucleotit của véc-tơ 24094.

Các SEQ ID NO: từ 39 đến 97 là các đoạn mồi và các đoạn dò dùng trong các thử nghiệm PCR Taqman được xác định.

SEQ ID NO: 98 là trình tự nucleotit của véc-tơ 24075.

Mô tả vắn tắt các hình vẽ

Hình 1 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350334-3: thể lưỡng bội (đỉnh chính ở 200, đỉnh phụ ở 400).

Hình 2 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350333-3: thể đơn bội (đỉnh chính ở 100, đỉnh phụ ở 200).

Hình 3 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350333-10: thể lưỡng bội (đỉnh chính ở 200, đỉnh phụ ở 400).

Hình 4 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350344-2: thể đơn bội (đỉnh chính ở 100, đỉnh phụ ở 200).

Hình 5 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350343-1: thể đơn bội (đỉnh chính ở 100, đỉnh phụ ở 200).

Hình 6 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350341-1: thể lưỡng bội (đỉnh chính ở 200, đỉnh phụ ở 400).

Hình 7 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350328-1: thể đơn bội (đỉnh chính ở 100, đỉnh phụ ở 200).

Hình 8 thể hiện dữ liệu phân tích mức bội thể (đếm tế bào theo dòng) đối với USR01350321-3: thể lưỡng bội (đỉnh chính ở 200, đỉnh phụ ở 400).

Hình 9 là hình vẽ dạng sơ đồ của véc-tơ 23396 (SEQ ID NO: 1) dùng cho sự biến nạp qua trung gian *Agrobacterium* của phôi chưa trưởng thành của ngô để tạo ra các đột biến được nhắm đích ở các gen ZmVLHP1. xZmVLHP-01: trình tự ARN dẫn (gRNA) (5'-GCAGGAGGCGTCGAGCAGCG-3', SEQ ID NO: 2); rsgRNAZmVLHP-01: ARN dẫn đơn lẻ (sgRNA) gồm có các trình tự kết thúc gARN, tracARN và PolIII. cPMI: Thành phần đánh dấu chọn lọc PMI; cCas9: Gen Cas9 nucleaza; RB: Biên phải T-ADN; LB: Biên trái T-ADN; tNOS: Trình tự kết thúc nopalin synthetaza. cSpec: Gen kháng spectinomyxin.

Hình 10 là hình vẽ dạng sơ đồ của véc-tơ 23399 (SEQ ID NO: 3) dùng cho sự biến nạp qua trung gian *Agrobacterium* của phôi chưa trưởng thành của ngô để tạo ra các đột biến được nhắm đích ở các gen ZmGW2. xZmGW2-02: trình tự ARN dẫn (gRNA) (5'-AAGCTCGCGCCCTGCTACCC-3', SEQ ID NO: 4); rsgRNAZmGW2-02: ARN dẫn đơn lẻ (sgRNA) gồm có các trình tự kết thúc gARN, tracrRNA và PolIII. cPMI-09: Thành phần đánh dấu chọn lọc PMI; cCas9-01: Gen Cas9 nucleaza; RB: Biên phải T-ADN; LB: Biên trái T-ADN; tNOS: Trình tự kết thúc nopalin synthetaza. cSpec: Gen kháng spectinomyxin.

Hình 11 thể hiện thử nghiệm mức bội thể của dòng ngô ngọt đơn bội được chỉnh sửa JSER82A056 và hình 12 thể hiện cùng thử nghiệm này đối với dòng ngô ngọt đơn bội được chỉnh sửa JSER82A063. Các dòng này thu được thông qua phép lai với dòng kích tạo đơn bội RWKS mang locut gen chuyển của véc-tơ biểu hiện CRISPR-Cas9 23399.

Hình 13 thể hiện sự xác nhận trình tự của việc chỉnh sửa vị trí đích GW2-02 ở dòng ngô ngọt đơn bội JSER82A063. Bazơ đơn lẻ C ở cạnh vị trí phân cắt Cas9 dự đoán được loại bỏ.

Hình 14 thể hiện phân tích mức bội thể của đối chứng kiểu dại, và hình 15 thể hiện phân tích mức bội thể của dòng lúa mỳ đơn bội đã chỉnh sửa JSWER30A22.

Hình 16 thể hiện sự xác nhận trình tự của việc chỉnh sửa vị trí đích TaVLHP1-4B ở dòng lúa mỳ đơn bội JSWER30A22. Ô dưới thể hiện 97 bp của trình tự TaVLHP1-4B được loại bỏ ngay lập tức xuôi dòng của vị trí phân cắt Cas9 dự đoán. các trình tự đã loại bỏ 97 bp được chỉ ra bằng 2 mũi tên.

Hình 17 thể hiện sự biểu hiện phấn như đo được bằng phấn được thu gom từ cây T0 ngô chuyển gen mang T-ADN của véc-tơ 24038, 24039, 24079, 24091, và 24094, mà được sử dụng để thụ phấn dòng lúa mỳ mùa xuân đã khử đực AC-Nanda. Sự biểu hiện là cao ở phấn, trung bình cao hơn khoảng 100 lần ở cây mang các véc-tơ T-ADN 24038, 24039, và 24079 so với vùng khởi động ubiquitin ở cây mía dùng trong nhiều ví dụ về ngô và lúa mỳ. Sự biểu hiện cũng cao hơn ở phấn từ cây chứa các véc-tơ 24038, 24039, và 24079 khi so với phấn mang các cấu trúc vùng khởi động *MATRILINEAL* (24094 và 24091).

Hình 18 thể hiện biểu đồ phân tích mức bội thể của đối chứng lưỡng bội (bố mẹ USR01424135). Đỉnh thứ nhất nằm ở vị trí "2" trên trục X lôgarit. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 19 thể hiện biểu đồ phân tích mức bội thể của đối chứng lưỡng bội (bố mẹ USR01431603). Đỉnh thứ nhất nằm ở vị trí "2" trên trục X lôgarit. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 20 thể hiện biểu đồ phân tích mức bội thể của đối chứng lưỡng bội (bố mẹ USR01431609). Đỉnh thứ nhất nằm ở vị trí "2" trên trục X lôgarit. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 21 thể hiện biểu đồ phân tích mức bội thể của thể đơn bội đã được chỉnh sửa từ đĩa 1033, giếng C3 (USR01424135 X Ler-427). Đỉnh thứ nhất nằm ở vị trí "1" trên trục X lôgarit, chỉ ra nó là thể đơn bội. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 22 thể hiện biểu đồ phân tích mức bội thể của thể đơn bội đã được chỉnh sửa từ đĩa 1033, giếng E4 (USR01424135 X Ler-437). Đỉnh thứ nhất nằm ở vị trí "1" trên trục X lôgarit, chỉ ra nó là thể đơn bội. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 23 thể hiện biểu đồ phân tích mức bội thể của thể đơn bội đã được chỉnh sửa từ đĩa 1046, giếng H12 (USR01431609 X Ler-123). Đỉnh thứ nhất nằm ở vị trí "1" trên trục X lôgarit, chỉ ra nó là thể đơn bội. Vì đây là mô lá từ *Arabidopsis*, chúng tôi thấy nhiều đỉnh chỉ ra mức bình thường của sự nội phân.

Hình 24 thể hiện các đột biến trình tự vị trí đích GL1 trong bố mẹ #USR01424135 và tất cả các thể đơn bội đã được chỉnh sửa trình tự từ phép lai xa bằng phấn hoa Landsberg erecta. Rõ ràng rằng sự chỉnh sửa chính xác được thực hiện khác nhau ở các thể đơn bội khác nhau.

Định nghĩa

Mặc dù các thuật ngữ sau đây được tin là được hiểu rõ bởi người có hiểu biết trung bình trong lĩnh vực, các định nghĩa sau đây được nêu để thuận tiện cho việc giải thích đối tượng được bộc lộ.

Tất cả các thuật ngữ khoa học và kỹ thuật dùng trong bản mô tả này, trừ khi được định nghĩa khác ở dưới đây, được dự định là có cùng nghĩa như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực. Việc đề cập đến các kỹ thuật dùng trong bản mô tả này được dự định để chỉ các kỹ thuật như thường được hiểu trong lĩnh vực, bao gồm các thay đổi trên các kỹ thuật này và/hoặc sự thay thế bằng kỹ thuật tương đương mà rõ ràng đối với người có hiểu biết trung bình trong lĩnh vực. Mặc dù các thuật ngữ sau đây được tin là được hiểu rõ bởi người có hiểu biết trung bình trong lĩnh vực, các định nghĩa sau đây được nêu để thuận tiện cho việc giải thích đối tượng được bộc lộ.

Theo thông lệ luật sáng chế đã có từ lâu, các thuật ngữ số ít dùng để chỉ "một hoặc nhiều" khi dùng trong đơn này, bao gồm các yêu cầu bảo hộ. Ví dụ, cụm từ "tế bào" dùng để chỉ một hoặc nhiều tế bào, và theo một số phương án có thể dùng để chỉ

mô và/hoặc cơ quan. Tương tự, cụm từ "ít nhất là một", khi dùng trong bản mô tả này để chỉ một thực thể, dùng để chỉ, ví dụ, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, hoặc nhiều thực thể, bao gồm nhưng không giới hạn ở tất cả các giá trị là số nguyên ở giữa 1 và 100 cũng như là các số nguyên lớn hơn 100.

Trừ khi có chỉ dẫn khác, tất cả các số biểu diễn khối lượng của các thành phần, điều kiện phản ứng, và v.v. được dùng trong bản mô tả và yêu cầu bảo hộ cần được hiểu là được sửa đổi trong tất cả các trường hợp bằng thuật ngữ "khoảng". Thuật ngữ "khoảng," như dùng trong bản mô tả này khi đề cập đến giá trị có thể đo được chẳng hạn như lượng của khối lượng, trọng lượng, thời gian, thể tích, nồng độ hoặc tỉ lệ phần trăm được dự định là bao hàm mức biến đổi theo một số phương án là ±20%, theo một số phương án ±10%, theo một số phương án ±5%, theo một số phương án ±1%, theo một số phương án ±0,5%, và theo một số phương án ±0,1 % so với lượng xác định, vì sự biến đổi này thích hợp để thực hiện phương pháp được bộc lộ và/hoặc sử dụng các hợp phần, axit nucleic, polypeptit, v.v. được bộc lộ. Theo đó, trừ khi chỉ ra điều ngược lại, các tham số bằng số nêu trong phần mô tả này và các yêu cầu bảo hộ kèm theo là con số xấp xỉ mà có thể thay đổi tùy thuộc vào các tính chất mong muốn muốn thu được bởi đối tượng được bộc lộ.

Như dùng trong bản mô tả này, thuật ngữ "alen" dùng để chỉ biến thể hoặc dạng trình tự thay thế ở locut di truyền. Trong thể lưỡng bội, alen đơn lẻ được di truyền bởi cây con khác biệt riêng rẽ từ mỗi bố mẹ ở mỗi locut. Hai alen của locut nhất định có mặt trong sinh vật lưỡng bội chiếm các vị trí tương ứng trên cặp nhiễm sắc thể tương đồng, mặc dù người có hiểu biết trung bình trong lĩnh vực hiểu rằng alen ở các thể cụ thể bất kỳ không nhất thiết đại diện cho tất cả các alen có trong loài.

Như dùng trong bản mô tả này, thuật ngữ "và/hoặc" khi dùng trong ngữ cảnh danh sách thực thể, dùng để chỉ các thực thể có mặt ở dạng đơn lẻ hoặc ở dạng kết hợp. Do đó, ví dụ, cụm từ "A, B, C, và/hoặc D" bao gồm A, B, C, và D riêng lẻ, nhưng cũng bao gồm dạng kết hợp và dạng kết hợp con bất kỳ và tất cả các dạng kết hợp và dạng kết hợp con của A, B, C, và D (ví dụ, AB, AC, AD, BC, BD, CD, ABC, ABD, và BCD). Theo một số phương án, một hoặc nhiều thành phần mà "và/hoặc" đề cập đến cũng có

thể có mặt riêng lẻ trong sự xuất hiện đơn lẻ hoặc nhiều lần xuất hiện trong (các) dạng kết hợp và/hoặc (các) dạng kết hợp con.

Như dùng trong bản mô tả này, cụm từ "kết hợp với" dùng để chỉ mối quan hệ có thể nhận ra và/hoặc có thể thử nghiệm giữa hai thực thể. Ví dụ, cụm từ "kết hợp với HI" dùng để chỉ tính trạng, locut, gen, alen, thành phần đánh dấu, kiểu hình, v.v., hoặc sự biểu hiện của chúng, sự có mặt hoặc không có mặt của chúng có thể ảnh hưởng đến phạm vi và/hoặc mức độ mà cây trồng hoặc cây con của nó thể hiện HI. Như vậy, thành phần đánh dấu "có liên quan đến" tính trạng khi nó được liên kết với và khi sự có mặt của thành phần đánh dấu là dấu hiệu chỉ thị của việc tính trạng hoặc dạng tính trạng mong muốn có ở cây/chất mầm nguyên sinh có chứa thành phần đánh dấu hay không và/hoặc có ở mức độ nào. Tương tự, thành phần đánh dấu "có liên quan đến" alen khi nó được liên kết với và khi sự có mặt của thành phần đánh dấu là dấu hiệu chỉ thị của việc alen có mặt ở cây/chất mầm nguyên sinh có chứa thành phần đánh dấu hay không. Ví dụ, "thành phần đánh dấu kết hợp với HI" dùng để chỉ thành phần đánh dấu mà sự có mặt hoặc không có mặt của nó có thể được dùng để dự đoán việc cây trồng có thể hiện sự kích tạo đơn bội hay không và/hoặc có ở mức độ nào.

Thuật ngữ "có chứa," mà đồng nghĩa với "bao gồm," "chứa," và "đặc trưng bởi,"có nghĩa bao gồm hoặc có nghĩa mở và không loại trừ các phần tử và/hoặc bước phương pháp khác không được chỉ ra. "Có chứa" là thuật ngữ trong lĩnh vực mà có nghĩa là các phần tử và/hoặc bước được kể tên có mặt, nhưng các phần tử và/hoặc bước khác có thể được bổ sung và vẫn nằm trong phạm vi của đối tượng liên quan.

Như dùng trong bản mô tả này, cụm từ "gồm có" không bao gồm phần tử, bước, hoặc thành phần không được chỉ ra cụ thể. Khi cụm từ "gồm có" xuất hiện trong mệnh đề của phần thân của yêu cầu bảo hộ, chứ không phải là ngay sau phần mở đầu, nó chỉ giới hạn phần tử nêu trong mệnh đề đó; các phần tử khác không bị loại trừ khỏi toàn bộ yêu cầu bảo hộ.

Như dùng trong bản mô tả này, cụm từ "về cơ bản gồm có" giới hạn phạm vi của phần bộc lộ hoặc yêu cầu bảo hộ liên quan ở các nguyên liệu và/hoặc bước được xác định, cộng với các nguyên liệu và/hoặc bước mà không ảnh hưởng thiết yếu đến (các) đặc điểm cơ bản và mới của đối tượng được bộc lộ và/hoặc được yêu cầu bảo hộ.

Đối với các thuật ngữ "có chứa," "về cơ bản gồm có," và "gồm có," khi mà một trong ba thuật ngữ này được sử dụng trong bản mô tả này, đối tượng hiện được bộc lộ và được yêu cầu bảo hộ có thể bao gồm theo một số phương án việc sử dụng một trong hai thuật ngữ kia. Ví dụ, nếu đối tượng theo một số phương án đề cập đến axit nucleic mà mã hóa cho polypeptit có chứa trình tự axit amin mà tương đồng ít nhất là 95% với SEQ ID NO: 9 hoặc 19. Cần hiểu rằng đối tượng được bộc lộ do đó cũng bao hàm axit nucleic mà mã hóa cho polypeptit mà theo một số phương án về cơ bản gồm có trình tự axit amin mà tương đồng ít nhất là 95% với SEQ ID NO: 9 hoặc 19 đó cũng như là axit nucleic mà mã hóa cho polypeptit mà theo một số phương án gồm có trình tự axit amin mà tương đồng ít nhất là 95% với SEQ ID NO: 9 hoặc 19 đó. Tương tự, cũng cần hiểu rằng theo một số phương án các phương pháp đối với đối tượng được bộc lộ có chứa các bước mà được bộc lộ trong bản mô tả này, theo một số phương án phương pháp đối với đối tượng được bộc lộ về cơ bản gồm có các bước mà được bộc lộ, và theo một số phương án phương pháp đối với đối tượng được bộc lộ gồm có các bước mà được bộc lộ trong bản mô tả này.

Như dùng trong bản mô tả này, thuật ngữ "sự kích tạo đơn bội mới" dùng để chỉ sự kích hoạt của sự kích tạo đơn bội bằng sự đưa vào của chất kích hoạt đơn bội tự động. Sự đưa vào này có thể đạt được bằng các phương pháp phun cục bộ, thụ phấn bằng tay, gây đột biến, hoặc chuyển gen. Các thuật ngữ "sự kích tạo đơn bội mới," "HI mới," và "sự kích tạo đơn bội mới" được sử dụng thay thế lẫn nhau trong toàn bộ bản mô tả này.

Như dùng trong bản mô tả này, thuật ngữ "gen" dùng để chỉ đơn vị di truyền bao gồm trình tự của ADN mà chiếm vị trí cụ thể trên nhiễm sắc thể và chứa chỉ thị di truyền đối với đặc điểm hoặc tính trạng cụ thể ở sinh vật.

"Bản đồ di truyền" là bản mô tả của mối quan hệ liên kết di truyền giữa các locut trên một hoặc nhiều nhiễm sắc thể ở loài nhất định, thường được minh họa ở dạng sơ đồ hoặc bảng.

Như dùng trong bản mô tả này, cây trồng được đề cập đến dưới dạng "đơn bội" có số lượng nhiễm sắc thể (n) giảm ở cây trồng đơn bội, và bộ nhiễm sắc thể của nó bằng bộ nhiễm sắc thể của giao tử. Ở sinh vật đơn bội, chỉ một nửa số lượng nhiễm sắc thể bình thường có mặt. Do đó thể đơn bội của sinh vật lưỡng bội (ví dụ, cây ngô) có

một bộ nhiễm sắc thể; thể đơn bội của sinh vật tứ bội (ví dụ, cỏ hoang) có hai bộ nhiễm sắc thể; thể đơn bội của sinh vật lục bội (ví dụ, cây lúa mỳ) có ba bộ nhiễm sắc thể; v.v. Như dùng trong bản mô tả này, cây trồng được đề cập đến dưới dạng "đơn bội kép" được phát triển bằng cách nhân đôi bộ nhiễm sắc thể đơn bội. Cây trồng hoặc hạt mà thu được từ cây trồng đơn bội kép mà được tự thụ phấn đến số lượng thế hệ bất kỳ vẫn có thể được xác định là cây trồng đơn bội kép. Cây trồng đơn bội kép được coi là cây đồng hợp tử. Cây trồng được coi là thể đơn bội kép nếu nó có khả năng sinh sản, ngay cả khi toàn bộ phần sinh dưỡng của cây trồng không có các tế bào có bộ nhiễm sắc thể kép; tức là, cây trồng sẽ được coi là thể đơn bội kép nếu nó chứa giao tử có thể phát triển độc lập được, ngay cả khi nó khảm ở mô sinh dưỡng.

Như dùng trong bản mô tả này, thuật ngữ "đột biến do con người tạo ra" dùng để chỉ đột biến bất kỳ mà xảy ra dưới sự tác động trực tiếp hoặc gián tiếp của con người. Thuật ngữ này bao gồm, nhưng không giới hạn ở, các đột biến thu được bằng phương pháp gây đột biến hướng đích bất kỳ.

Như dùng trong bản mô tả này, "được đưa vào" có nghĩa là được chuyển, được biểu hiện, được dùng, được vận chuyển, được chuyển, được thấm, hoặc thuật ngữ tương tự khác để chỉ việc chuyển, axit nucleic hoặc protein hoặc dạng kết hợp của chúng, của đối tượng mong muốn vào đối tượng. Ví dụ, axit nucleic mã hóa cho nucleaza định hướng vị trí và tùy ý ít nhất là một ARN dẫn có thể được đưa vào phôi đơn bội khi kích tạo đơn bội. Tương tự, bộ máy chỉnh sửa hiện có (có chứa protein nucleaza định hướng vị trí và tùy ý ít nhất là một ARN dẫn) có thể được đưa vào phôi đơn bội khi dùng peptit thấm vào tế bào thích hợp.

Như dùng trong bản mô tả này, các thuật ngữ "đoạn dò đánh dấu" và "đoạn dò" dùng để chỉ trình tự nucleotit hoặc phân tử axit nucleic mà có thể được dùng để phát hiện sự có mặt hoặc không có mặt của trình tự ở trong trình tự lớn hơn, ví dụ, đoạn dò axit nucleic mà bổ sung với tất cả hoặc một phần của thành phần đánh dấu hoặc locut thành phần đánh dấu, thông qua sự lai hóa axit nucleic. Các đoạn dò thành phần đánh dấu có chứa khoảng 8, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 hoặc nhiều nucleotit liên tiếp có thể được dùng cho sự lai hóa axit nucleic.

Như dùng trong bản mô tả này, thuật ngữ "thành phần đánh dấu phân tử" có thể được dùng để chỉ thành phần đánh dấu di truyền, như xác định ở trên, hoặc sản phẩm được mã hóa của chúng (ví dụ, protein) dùng làm điểm tham chiếu khi nhận dạng sự có mặt/không có mặt của locut liên kết HI. Thành phần đánh dấu phân tử có thể có nguồn gốc từ trình tự nucleotit hệ gen hoặc từ trình tự nucleotit được biểu hiện (ví dụ, từ ARN, cADN, v.v.). Thuật ngữ này cũng dùng để chỉ trình tự nucleotit bổ sung với hoặc củng cố biên trình tự thành phần đánh dấu, chẳng hạn như trình tự nucleotit được sử dụng làm đoạn dò và/hoặc đoạn mồi có khả năng khuếch đại trình tự thành phần đánh dấu. Trình tự nucleotit là "bổ sung" khi chúng lai hóa đặc hiệu trong dung dịch (ví dụ, theo quy tắc bắt cặp bazơ Watson-Crick). Thuật ngữ này cũng dùng để chỉ thành phần đánh dấu di truyền mà biểu thị tính trạng bởi sự không có mặt của trình tự nucleotit bổ sung với hoặc củng cố biên trình tự thành phần đánh dấu, chẳng hạn như trình tự nucleotit được sử dụng làm đoạn dò và/hoặc đoạn mồi có khả năng khuếch đại trình tự thành phần đánh dấu.

Như dùng trong bản mô tả này, các thuật ngữ "trình tự nucleotit," "polynucleotit," "trình tự axit nucleic," "phân tử axit nucleic," và "mảnh axit nucleic" dùng để chỉ polyme của ARN hoặc ADN mà là sợi đơn hoặc sợi kép, tùy ý chứa bazơ nucleotit tổng hợp, không tự nhiên, và/hoặc được làm thay đổi. "Nucleotit" là đơn vị monome mà từ đó các polyme ADN hoặc ARN được tạo cấu trúc và có chứa bazơ purin hoặc pyrimiđin, pentoza, và nhóm axit phosphoric. Nucleotit (thường được tìm thấy ở dạng 5'-monophosphat của chúng) được đề cập đến bằng ký hiệu chữ cái đơn lẻ của chúng như sau: "A" cho ađenylat hoặc đeoxyađenylat (lần lượt đối với ARN hoặc ADN), "C" cho xytiđylat hoặc đeoxyxytiđylat, "G" cho guanylat hoặc đeoxyguanylat , "U" cho uriđylat, "T" cho đeoxytymiđylat, "R" cho purin (A hoặc G), "Y" cho pyrimiđin (C hoặc T), "K" cho G hoặc T, "H" cho A hoặc C hoặc T, "I" cho inosin, và "N" cho nucleotit bất kỳ.

Như dùng trong bản mô tả này, thuật ngữ "sự tương đồng trình tự nucleotit" dùng để chỉ sự có mặt của nucleotit tương đồng ở vị trí tương ứng của hai polynucleotit. Các polynucleotit có trình tự "tương đồng" nếu trình tự của nucleotit trong hai polynucleotit là giống nhau khi được sắp xếp thẳng hàng để có độ tương ứng lớn nhất (ví dụ, trong cửa sổ so sánh). Sự so sánh trình tự giữa hai hoặc nhiều polynucleotit thường được thực hiện bằng cách so sánh các phần của hai trình tự qua cửa sổ so sánh để xác định và so

sánh các vùng cục bộ của sự tương tự trình tự. Cửa sổ so sánh thường có từ khoảng 20 đến 200 nucleotit liên tiếp. "Tỉ lệ phần trăm của độ tương đồng trình tự" đối với polynucleotit, chẳng hạn như độ tương đồng trình tự khoảng 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 hoặc 100 phần trăm, có thể được xác định bằng cách so sánh hai trình tự được sắp xếp thẳng hàng tối ưu qua cửa sổ so sánh, trong đó phần của trình tự polynucleotit trong cửa sổ so sánh có thể bao gồm đột biến thêm hoặc đột biến mất (tức là, chỗ cách quãng) khi so với trình tự tham chiếu đối với sự sắp xếp thẳng hàng tối ưu của hai trình tự. Theo một số phương án, tỉ lệ phần trăm được tính bằng cách: (a) xác định số lượng vị trí mà tại đó bazơ axit nucleic tương đồng có mặt ở cả hai trình tự; (b) chia số lượng vị trí bắt cặp cho tổng số vị trí trong cửa sổ so sánh; và (c) nhân kết quả với 100. Sự sắp xếp thẳng hàng tối ưu của các trình tự để so sánh cũng có thể được thực hiện bằng cách thực hiện vi tính hóa các thuật toán đã biết, hoặc bằng cách quan sát bằng mắt. Các thuật toán sắp xếp thẳng hàng so sánh trình tự và sắp xếp thẳng hàng nhiều trình tự có sẵn, lần lượt, là công cụ tra cứu sắp xếp thẳng hàng cục bộ cơ bản (Basic Local Alignment Search Tool - BLAST) và chương trình ClustalW/ClustalW2/Clustal Omega có sẵn trên Internet (ví dụ, trang web của EMBL-EBI). Các chương trình thích hợp khác bao gồm, nhưng không giới hạn ở, GAP, BestFit, Plot Similarity, và FASTA, mà là một phần của Gói Accelrys GCG có sẵn từ Accelrys, Inc. ở San Diego, California, Hoa Kỳ. Cũng xem Smith & Waterman, 1981; Needleman & Wunsch, 1970; Pearson & Lipman, 1988; Ausubel *et al.*, 1988; và Sambrook & Russell, 2001.

Một ví dụ về thuật toán mà là thích hợp để xác định tỉ lệ phần trăm độ tương đồng trình tự và độ tương tự trình tự là thuật toán BLAST, mà được mô tả trong Altschul *et al.*, 1990. Theo một số phương án, tỉ lệ phần trăm của độ tương đồng trình tự dùng để chỉ độ tương đồng trình tự trên chiều dài đầy đủ của một trong số gADN, cADN, hoặc trình tự protein được dự đoán trong ORF lớn nhất của SEQ ID No: 1 được so sánh. Theo một số phương án, phép tính để xác định tỉ lệ phần trăm của độ tương đồng trình tự axit nucleic không bao gồm trong phép tính vị trí nucleotit bất kỳ trong đó một trong hai axit nucleic được so sánh bao gồm "N" (tức là, trong đó nucleotit bất kỳ có thể có mặt tại vị trí đó).

Thuật ngữ "khung đọc mở" (ORF) dùng để chỉ trình tự axit nucleic mà mã hóa cho polypeptit. Theo một số phương án, ORF có chứa bộ ba mở đầu dịch mã (tức là, bộ ba mở đầu), sự kết thúc dịch mã (tức là, bộ ba kết thúc), và trình tự axit nucleic ở giữa chúng mà mã hóa cho các axit amin có mặt trong polypeptit. Các thuật ngữ "bộ ba mở đầu" và "bộ ba kết thúc" dùng để chỉ đơn vị gồm ba nucleotit liền kề (tức là, bộ ba mã hóa) trong trình tự mã hóa mà lần lượt chỉ ra sự mở đầu và sự kết thúc chuỗi, của sự tổng hợp protein (sự dịch mã mARN).

Phospholipaza giống patatin A2α cũng có thể được gọi là PLA, pPLA, pPLAIIA pPLAIIα, PLA2alpha, hoặc PLA2, hoặc biến thể tương tự khác. Phospholipaza giống patatin AIIα còn được đề cập đến dưới dạng *MATRILINEAL* (*MATL*). Các thuật ngữ này được dùng thay thế lẫn nhau trong toàn bộ bản mô tả này. Gen *MATRILINEAL* có chứa đột biến dịch khung bốn cặp bazơ được đề cập đến dưới dạng *matrilineal* (*matl*).

Như dùng trong bản mô tả này, các thuật ngữ "kiểu hình," "tính trạng kiểu hình tính trạng" hoặc "tính trạng" dùng để chỉ một hoặc nhiều tính trạng của cây trồng hoặc tế bào cây trồng. Kiểu hình có thể quan sát được bằng mắt thường, hoặc bằng phương tiện đánh giá khác bất kỳ đã biết trong lĩnh vực, ví dụ, kính hiển vi, phân tích hóa sinh, hoặc thử nghiệm điện cơ. Trong một số trường hợp, kiểu hình được kiểm soát trực tiếp bởi gen hoặc locut di truyền đơn lẻ (tức là, tương ứng với "tính trạng đơn gen"). Trong trường hợp kích tạo đơn bội việc sử dụng thành phần đánh dấu màu sắc, chẳng hạn như R Navajo, và các thành phần đánh dấu khác bao gồm gen chuyển được trực quan hóa bởi sự có mặt hoặc không có mặt của màu sắc trong hạt chứng tỏ hạt có phải là hạt đơn bội được cảm ứng hay không. Việc sử dụng R Navajo làm thành phần đánh dấu màu sắc và việc sử dụng gen chuyển được biết rõ trong lĩnh vực là phương tiện để phát hiện sự gây tạo hạt đơn bội trên cây cái. Trong các trường hợp khác, kiểu hình là kết quả của sự tương tác giữa một vài gen, mà theo một số phương án cũng bắt nguồn từ sự tương tác của cây trồng và/hoặc tế bào cây trồng với môi trường của nó.

Như dùng trong bản mô tả này, thuật ngữ "cây trồng" có thể dùng để chỉ cây trồng nguyên vẹn, phần bất kỳ của chúng, hoặc môi trường nuôi cấy tế bào hoặc mô có nguồn gốc từ cây trồng. Do đó, thuật ngữ "cây trồng" có thể dùng để chỉ thành phần bất

kỳ trong số: cây trồng nguyên vẹn, thành phần cây trồng hoặc cơ quan (ví dụ, lá, thân, rễ, v.v.), mô cây trồng, hạt và/hoặc tế bào cây trồng.

Tế bào cây trồng là tế bào của cây trồng, được lấy từ cây trồng, hoặc được tạo ra thông qua nuôi cấy từ tế bào được lấy từ cây trồng. Do đó, thuật ngữ "tế bào cây trồng" bao gồm nhưng không giới hạn ở các tế bào trong hạt, môi trường nuôi cấy huyền phù, phôi, vùng mô phân sinh, mô sẹo, lá, cành non, thể giao tử, thể bào tử, phấn hoa và tiểu bào tử. Cụm từ "phần cây trồng" dùng để chỉ phần của cây trồng, bao gồm tế bào đơn lẻ và mô tế bào chẳng hạn như tế bào cây trồng mà nguyên vẹn trong cây trồng, cụm tế bào, và môi trường nuôi cấy mô mà cây trồng có thể được tái tạo từ đó. Ví dụ về phần cây trồng bao gồm, nhưng không giới hạn ở, tế bào đơn lẻ và mô từ phấn, noãn, lá, phôi, rễ, đỉnh rễ, bao phấn, hoa, quả, thân, chồi, và hạt; cũng như là mầm, thân rễ, thể nguyên sinh, thể sần, và dạng tương tự.

Như dùng trong bản mô tả này, thuật ngữ "đoạn mồi" dùng để chỉ oligonucleotit mà có khả năng ủ để gắn vào đích axit nucleic (theo một số phương án, việc ủ để gắn đặc hiệu vào đích axit nucleic) cho phép ADN polymeraza và/hoặc transcriptaza ngược gắn vào đó, bằng cách đó đóng vai trò làm điểm mở đầu của sự tổng hợp ADN khi đặt trong điều kiện trong đó sự tổng hợp của sản phẩm kéo dài đoạn mồi được gây ra (ví dụ, trong sự có mặt của nucleotit và tác nhân để polyme hóa chẳng hạn như ADN polymeraza và ở nhiệt độ và độ pH thích hợp). Theo một số phương án, một hoặc nhiều đoạn mồi được sử dụng để khuếch đại axit nucleic cây trồng (ví dụ, bằng cách sử dụng phản ứng chuỗi polymeraza; PCR).

Như dùng trong bản mô tả này, thuật ngữ "đoạn dò" dùng để chỉ axit nucleic (ví dụ, axit nucleic sợi đơn hoặc sợi của axit nucleic sợi kép hoặc bậc cao hơn, hoặc trình tự con của chúng) mà có thể tạo thành bộ đôi được liên kết bằng hyđro với trình tự bổ sung trong trình tự axit nucleic đích. Thông thường, đoạn dò có chiều dài đủ để tạo thành phân tử bộ đôi đặc hiệu trình tự và ổn định với phần bổ sung của nó, và vì thế có thể được sử dụng theo một số phương án để phát hiện trình tự quan tâm có mặt trong một số lượng các axit nucleic.

Như dùng trong bản mô tả này, các thuật ngữ "con cháu" và "cây trồng con cháu" dùng để chỉ cây trồng được tạo ra từ sự sinh sản sinh dưỡng hoặc hữu tính từ một hoặc

nhiều cây trồng bố mẹ. Trong sự kích tạo đơn bội qua trung gian sự sinh sản đơn tính cái, phôi đơn bội trên thể mẹ có chứa nhiễm sắc thể cái để loại trừ nhiễm sắc thể đực, do đó nó không phải là con cháu của dòng kích tạo đơn bội đực. Hạt ngô đơn bội thường vẫn có nội nhũ tam bội bình thường mà chứa hệ gen đực. Cây con đơn bội đã được chỉnh sửa và cây đơn bội kép đã được chỉnh sửa tiếp và hạt tiếp theo không phải là con cháu mong muốn duy nhất. Cũng có hạt từ bản thân dòng kích tạo đơn bội, thường mang gen chuyển Cas9, và cây con cháu và hạt con cháu tiếp theo của cây trồng tạo thể đơn bội. Cả hạt đơn bội và hạt kích tạo đơn bội (có nguồn gốc từ tự thụ phấn) có thể là con cháu. Cây trồng con có thể thu được bằng cách tách dòng hoặc tự thụ phấn cây bố mẹ đơn lẻ, hoặc bằng cách lai hai hoặc hơn hai cây bố mẹ. Ví dụ như, cây trồng con có thể thu được bằng cách tách dòng hoặc tự thụ phấn của cây trồng bố mẹ hoặc bằng cách lai chéo hai cây trồng bố mẹ và bao gồm tự thụ phấn cũng như là F_1 hoặc F_2 hoặc các thế hệ tiếp theo nữa. F_1 là cây con thế hệ thứ nhất được tạo ra từ bố mẹ ít nhất là một trong số chúng được sử dụng trong lần thứ nhất làm thể cho của tính trạng, trong khi cây con của thế hệ thứ hai (F_2) hoặc các thế hệ tiếp theo (F_3, F_4, và dạng tương tự) là các vật mẫu được tạo ra từ tự thụ phấn, lai chéo, lai ngược, và/hoặc các phép lai khác của F_1, F_2, và dạng tương tự. Do đó F_1 có thể là (và theo một số phương án là) thể lai bắt nguồn từ phép lai giữa hai bố mẹ dòng thuần (tức là, bố mẹ mà là dòng thuần là mỗi đồng hợp tử đối với tính trạng quan tâm hoặc alen của chúng), trong khi F_2 có thể là (và theo một số phương án là) cây con bắt nguồn từ sự tự thụ phấn của thể lai F_1.

Như dùng trong bản mô tả này, cụm từ "tái tổ hợp" dùng để chỉ sự trao đổi của các mảnh ADN giữa hai phân tử ADN hoặc nhiễm sắc tử của các nhiễm sắc thể ghép cặp ("trao đổi chéo") trong khu vực của trình tự nucleotit tương tự hoặc tương đồng. "Biến cố tái tổ hợp" trong bản mô tả này được hiểu theo một số phương án là dùng để chỉ trao đổi chéo trong giảm phân.

Như được dùng trong bản mô tả này, thuật ngữ "trình tự tham chiếu" dùng để chỉ trình tự nucleotit xác định được sử dụng làm cơ sở để so sánh trình tự nucleotit. Theo một số phương án, trình tự bất kỳ trong số các SEQ ID NO: 2, 4, 8, 21, 23, 25, 29, 32, và 33 có thể đóng vai trò làm trình tự tham chiếu để so sánh với trình tự khác thu được từ cây trồng.

Như dùng trong bản mô tả này, thuật ngữ "tái tạo," và các biến thể ngữ pháp của chúng, dùng để chỉ sự sản xuất cây trồng từ nuôi cấy mô.

Như dùng trong bản mô tả này, cụm từ "điều kiện lai hóa nghiêm ngặt" dùng để chỉ điều kiện mà tại đó polynucleotit lai hóa với trình tự con đích của nó, thường là trong hỗn hợp phức của các axit nucleic, nhưng về cơ bản không lai hóa với các trình tự khác. Điều kiện nghiêm ngặt phụ thuộc trình tự và có thể khác nhau trong các trường hợp khác nhau.

Trình tự dài hơn thường lai hóa đặc hiệu ở nhiệt độ cao hơn. Hướng dẫn tổng quát về sự lai hóa của axit nucleic được tìm thấy trong Sambrook & Russell, 2001. Thông thường, điều kiện nghiêm ngặt được chọn là nhiệt độ thấp hơn khoảng 5-10°C so với điểm nóng chảy nhiệt (T_m) đối với trình tự cụ thể ở độ mạnh ion độ pH xác định. Tm là nhiệt độ (ở độ mạnh ion, độ pH, và nồng độ axit nucleic xác định) mà tại đó 50% của đoạn dò bổ sung với đích lai hóa với trình tự đích ở trạng thái cân bằng (khi trình tự đích có mặt dư, ở Tm, 50% đoạn dò bị chiếm giữ ở trạng thái cân bằng). Điều kiện nghiêm ngặt ví dụ là điều kiện trong đó nồng độ muối nhỏ hơn khoảng 1,0M ion natri, thường là nồng độ ion natri nằm trong khoảng từ khoảng 0,01 đến 1,0M (hoặc muối khác) ở độ pH từ 7,0 đến 8,3 và nhiệt độ bằng ít nhất là khoảng 30°C đối với đoạn dò ngắn (ví dụ, từ 10 đến 50 nucleotit) và ít nhất là khoảng 60°C đối với đoạn dò dài (ví dụ, lớn hơn 50 nucleotit).

Điều kiện nghiêm ngặt cũng có thể đạt được với sự bổ sung của chất khử ổn định chẳng hạn như formamit. Điều kiện lai hóa nghiêm ngặt ví dụ khác bao gồm 50% formamit, 5x SSC, và 1% SDS ủ ở nhiệt độ 42°C; hoặc SSC, 1% SDS, ủ ở nhiệt độ 65°C; với một hoặc nhiều lần rửa trong 0,2x SSC và 0,1% SDS ở nhiệt độ 65°C. Để PCR, nhiệt độ khoảng 36°C là thông thường đối với sự khuếch đại có độ nghiêm ngặt thấp, mặc dù nhiệt độ ủ có thể thay đổi trong khoảng từ khoảng 32°C đến 48°C (hoặc cao hơn) tùy thuộc vào chiều dài đoạn mồi. Các hướng dẫn khác để xác định thông số lai hóa được nêu trong nhiều tài liệu tham khảo (xem tài liệu ví dụ, Ausubel *et al.*, 1999).

Như dùng trong bản mô tả này, thuật ngữ "tính trạng" dùng để chỉ kiểu hình quan tâm, gen mà đóng góp vào kiểu hình quan tâm, cũng như là trình tự axit nucleic kết hợp với gen mà đóng góp vào kiểu hình quan tâm. Ví dụ, "tính trạng HI" dùng để chỉ kiểu

hình kích tạo đơn bội cũng như là gen (ví dụ, *matl* ở cây ngô hoặc Os03g27610 ở cây lúa) mà đóng góp vào sự kích tạo đơn bội và trình tự axit nucleic (ví dụ, sản phẩm gen liên quan đến HI) tức là liên quan đến sự có mặt hoặc không có mặt của kiểu hình kích tạo đơn bội.

Như dùng trong bản mô tả này, thuật ngữ "gen chuyển" dùng để chỉ phân tử axit nucleic được đưa vào sinh vật hoặc một hoặc nhiều tổ tiên của nó bằng một số dạng của kỹ thuật chuyển nhân tạo. Do đó kỹ thuật chuyển nhân tạo tạo ra "sinh vật chuyển gen" hoặc "tế bào chuyển gen". Cần hiểu rằng kỹ thuật chuyển nhân tạo có thể xảy ra ở sinh vật tổ tiên (hoặc tế bào trong đó và/hoặc có thể phát triển thành sinh vật tổ tiên) và còn cá thể cây con bất kỳ mà có phân tử axit nucleic đã được chuyển nhân tạo hoặc mảnh của chúng vẫn được coi là chuyển gen ngay cả khi một hoặc nhiều sự gây giống tự nhiên và/hoặc được hỗ trợ dẫn đến phân tử axit nucleic đã được chuyển nhân tạo có mặt trong cá thể cây con.

Như dùng trong bản mô tả này, thuật ngữ "sự gây đột biến hướng đích" hoặc "chiến lược gây đột biến" dùng để chỉ phương pháp gây đột biến bất kỳ mà dẫn đến sự gây đột biến có chủ đích của gen được chọn. Sự gây đột biến hướng đích bao gồm các phương pháp CRISPR, TILLING, TALEN, và các phương pháp khác chưa được phát hiện nhưng mà có thể được dùng để đạt được cùng kết quả.

Như dùng trong bản mô tả này, tỉ lệ kích tạo đơn bội ("HIR") có nghĩa là số lượng nhân đơn bội sống sót trong tổng số nhân sau khi bông được thụ phấn với phấn hoa kích tạo đơn bội.

Các vấn đề cụ thể gây hại cho sự kích tạo đơn bội: tỉ lệ chết phôi tăng và tỉ lệ thụ tinh thất bại tăng (tỉ lệ kết hạt giảm). Vì các lí do này, có nhu cầu về việc xác định thành công nguyên nhân của HI, và sử dụng hiểu biết đó để xác định phương pháp tạo ra một cách ổn định hoặc gia tăng cây đơn bội trong khi đồng thời làm giảm sự thụ tinh thất bại và chết phôi.

Dự định cụ thể là người ta có thể gây đột biến vùng khởi động để có khả năng cải thiện tính hữu dụng của các thành phần để biểu hiện gen chuyển ở cây trồng. Sự gây đột biến của các thành phần này có thể được thực hiện một cách ngẫu nhiên và trình tự vùng

khởi động đã được gây đột biến được sàng lọc về hoạt tính trong quy trình thử và sai. Theo cách khác, trình tự cụ thể mà cung cấp vùng khởi động với các đặc điểm biểu hiện mong muốn, hoặc vùng khởi động với hoạt tính tăng cường biểu hiện, có thể được xác định và các trình tự này hoặc các trình tự tương tự được đưa vào vùng khởi động thông qua đột biến. Dự tính thêm rằng người ta có thể gây đột biến các trình tự này để tăng cường sự biểu hiện của chúng của gen chuyển ở các loài cụ thể. Phương thức để gây đột biến đoạn ADN mã hóa cho trình tự vùng khởi động theo sáng chế được biết rõ đối với người có hiểu biết trung bình trong lĩnh vực. Như đã chỉ ra, các cải biến đối với vùng khởi động hoặc thành phần điều hòa khác có thể được tạo ra bằng quy trình gây đột biến ngẫu nhiên, hoặc đặc hiệu vị trí. Vùng khởi động và thành phần điều hòa khác có thể được cải biến bằng cách làm thay đổi cấu trúc của chúng thông qua đột biến thêm hoặc đột biến mất của một hoặc nhiều nucleotit từ trình tự mà mã hóa cho trình tự chưa được cải biến tương ứng.

Sự gây đột biến có thể được thực hiện theo kỹ thuật bất kỳ đã biết trong lĩnh vực, chẳng hạn như, và không giới hạn ở, tổng hợp oligonucleotit có một hoặc nhiều đột biến trong trình tự của trình tự điều hòa cụ thể. Cụ thể là, sự gây đột biến đặc hiệu vị trí là kỹ thuật hữu dụng trong việc điều chế thể đột biến vùng khởi động, thông qua sự gây đột biến đặc hiệu của ADN ẩn dưới. Enđonucleaza được dẫn bằng ARN ("RGEN," ví dụ, CRISPR/Cas9) cũng có thể được dùng. Kỹ thuật này còn mang lại khả năng sẵn sàng để điều chế và thử nghiệm biến thể trình tự, ví dụ, kết hợp một hoặc nhiều suy xét trên đây, bằng cách đưa một hoặc nhiều sự thay đổi trình tự nucleotit vào ADN. Sự gây đột biến đặc hiệu vị trí cho phép sản xuất thể đột biến thông qua việc sử dụng trình tự oligonucleotit đặc hiệu mà mã hóa cho trình tự ADN của đột biến mong muốn, cũng như là số lượng đủ của nucleotit liền kề, để cung cấp trình tự đoạn mồi có kích thước và độ phức tạp trình tự đủ để tạo thành bộ đôi ổn định trên cả hai phía của chỗ tiếp giáp đột biến mất được cắt ngang. Thông thường, đoạn mồi có chiều dài từ khoảng 17 đến khoảng 75 nucleotit hoặc nhiều hơn được ưu tiên, với từ khoảng 10 đến khoảng 25 hoặc nhiều hơn các gốc trên cả hai phía của chỗ tiếp giáp của trình tự được làm thay đổi.

Khi dòng có chứa vùng khởi động được phân lập theo sáng chế, người ta có thể muốn phân định các khu vực vùng khởi động thiết yếu trong dòng này. Một phương thức nhắm đích, hữu hiệu để điều chế vùng khởi động đã được gây đột biến dựa trên

việc xác định các thành phần điều hòa giả định trong trình tự vùng khởi động. Điều này có thể được bắt đầu bằng cách so sánh với các trình tự vùng khởi động đã biết được biểu hiện ở kiểu mẫu duy nhất về mặt phát triển hoặc đặc hiệu mô tương tự. Trình tự mà được dùng chung giữa các vùng khởi động có kiểu mẫu biểu hiện tương tự có vẻ phù hợp làm ứng viên cho sự liên kết của các yếu tố phiên mã và do đó có vẻ phù hợp làm các thành phần mà mang lại kiểu mẫu biểu hiện. Việc xác nhận các thành phần điều hòa giả định này có thể đạt được bằng phân tích đột biến mất của mỗi trình tự điều hòa giả định sau đó là phân tích chức năng của mỗi cấu trúc mất bằng thử nghiệm của gen báo cáo mà được gắn về mặt chức năng với mỗi cấu trúc. Như vậy, khi trình tự vùng khởi động bắt đầu được cung cấp, số lượng bất kỳ của thể đột biến mất khác nhau của vùng khởi động bắt đầu có thể dễ dàng được điều chế.

Sáng chế như được bộc lộ trong bản mô tả này đề xuất phân tử polynucleotit có chứa mảnh thành phần điều hòa mà có thể được dùng trong việc tạo cấu trúc các thành phần điều hòa khảm mới. Các dạng kết hợp mới có chứa mảnh của các phân tử polynucleotit này và ít nhất là một thành phần điều hòa khác hoặc mảnh có thể được tạo cấu trúc và được thử nghiệm ở cây trồng và được coi là nằm trong phạm vi của sáng chế. Do đó việc thiết kế, tạo cấu trúc, và sử dụng của các thành phần điều hòa khảm là một phương án của sáng chế. Vùng khởi động theo sáng chế bao gồm thể tương đồng của các thành phần cis đã biết là ảnh hưởng đến sự điều hòa gen mà thể hiện sự tương đồng với trình tự vùng khởi động theo sáng chế.

Mảnh tương đương chức năng của một trong số các axit nucleic điều hòa phiên mã được mô tả trong bản mô tả này có chứa ít nhất là 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, hoặc 1000 cặp bazơ của axit nucleic điều hòa phiên mã. Mảnh tương đương của axit nucleic điều hòa phiên mã, mà thu được bằng cách làm mất vùng mã hóa cho vùng không dịch mã 5' của mARN, thì sẽ chỉ cung cấp vùng khởi động (không phiên mã). Vùng không dịch mã 5' có thể được xác định dễ dàng bằng các phương pháp đã biết trong lĩnh vực (chẳng hạn như phân tích 5'-RACE). Theo đó, một số axit nucleic điều hòa phiên mã, được mô tả trong bản mô tả này, là mảnh tương đương của trình tự khác.

Như đã chỉ ra ở trên, thể đột biến mất của vùng khởi động theo sáng chế cũng có thể được điều chế theo cách ngẫu nhiên và sau đó được thử nghiệm. Theo chiến lược này, một loạt cấu trúc được điều chế, mỗi cấu trúc chứa phần khác nhau của vùng khởi động (dòng phụ), và sau đó các cấu trúc này được sàng lọc về hoạt tính. Phương thức thích hợp để sàng lọc về hoạt tính là gắn vùng khởi động hoặc cấu trúc intron đã được làm mất mà chứa đoạn đã được làm mất vào thành phần đánh dấu có thể chọn lọc được hoặc có thể sàng lọc được, và phân lập chỉ những tế bào biểu hiện gen đánh dấu. Theo cách này, một số lượng của các cấu trúc vùng khởi động đã được làm mất, khác nhau được xác định mà vẫn giữ lại hoạt tính mong muốn, hoặc thậm chí được tăng cường. Đoạn nhỏ nhất mà cần thiết đối với hoạt tính bằng cách đó được xác định thông qua việc so sánh các cấu trúc được chọn. Sau đó đoạn này có thể được dùng để tạo cấu trúc véc-tơ để biểu hiện gen ngoại sinh.

Cát xét biểu hiện như được mô tả trong bản mô tả này có thể có chứa các thành phần điều hòa khác. Thuật ngữ này trong ngữ cảnh của sáng chế cần được hiểu theo nghĩa rộng có chứa tất cả các trình tự mà có thể ảnh hưởng đến sự tạo cấu trúc hoặc chức năng của cát xét biểu hiện. Các thành phần điều hòa có thể, ví dụ, làm biến đổi sự phiên mã và/hoặc sự dịch mã ở sinh vật nhân sơ hoặc sinh vật nhân chuẩn. Cát xét biểu hiện được mô tả trong bản mô tả này có thể nằm xuôi dòng (theo chiều 3') của trình tự axit nucleic được biểu hiện và tùy ý chứa các thành phần điều hòa khác, chẳng hạn như vùng tăng cường phiên mã hoặc dịch mã. Mỗi thành phần điều hòa khác này có thể được liên kết có điều khiển với trình tự axit nucleic được biểu hiện (hoặc trình tự nucleotit điều hòa phiên mã). Các thành phần điều hòa khác có thể có chứa vùng khởi động khác, vùng khởi động tối thiểu, các thành phần khởi động, các thành phần gen nhảy mà có thể làm biến đổi hoặc tăng cường tính chất điều hòa biểu hiện. Cát xét biểu hiện cũng có thể chứa một hoặc nhiều intron, một hoặc nhiều exon và một hoặc nhiều vùng kết thúc.

Hơn nữa, dự định rằng vùng khởi động kết hợp các thành phần từ hơn một vùng khởi động có thể hữu dụng. Ví dụ, bằng sáng chế Mỹ Số 5,491,288 bộc lộ việc kết hợp vùng khởi động virut khảm súp lơ với vùng khởi động histon. Do đó, các thành phần từ vùng khởi động được bộc lộ trong bản mô tả này có thể được kết hợp với các thành phần từ vùng khởi động khác. Vùng khởi động mà hữu dụng đối với sự biểu hiện gen chuyển ở cây trồng bao gồm các vùng khởi động mà là vùng khởi động cảm ứng, vùng khởi

động virut, vùng khởi động tổng hợp, vùng khởi động cấu trúc (Odell Nature 313: 810–812 (1985)), được điều hòa về mặt thời gian, được điều hòa về mặt không gian, đặc hiệu mô, và được điều hòa về mặt không thời gian. Bằng cách sử dụng các thành phần điều hòa được mô tả trong bản mô tả này, nhiều gen nông nghiệp có thể được biểu hiện ở cây trồng được biến nạp. Cụ thể hơn là, cây trồng có thể được thiết kế di truyền để biểu hiện các kiểu hình khác nhau được quan tâm trong nông nghiệp.

Mô tả chi tiết sáng chế

Một phương án của sáng chế đề xuất phương pháp chỉnh sửa ADN hệ gen cây trồng, có chứa các bước: (i) thu lấy cây trồng thứ nhất, trong đó cây trồng thứ nhất này là dòng kích tạo đơn bội của cây trồng, và trong đó cây trồng thứ nhất này có khả năng biểu hiện enzym cải biến ADN và tùy ý axit nucleic dẫn; (ii) thu lấy cây trồng thứ hai, trong đó cây trồng thứ hai này có chứa ADN hệ gen cây trồng mà cần được chỉnh sửa; (iii) thụ phấn cây trồng thứ hai với phấn từ cây trồng thứ nhất; và (iv) chọn lọc ít nhất là một cây con đơn bội được tạo ra bởi sự thụ phấn của bước (c) trong đó cây con đơn bội có chứa hệ gen của cây trồng thứ hai chứ không phải cây trồng thứ nhất, và hệ gen của cây con đơn bội đã được cải biến bằng enzym cải biến ADN và axit nucleic dẫn tùy chọn được chuyển bởi cây trồng thứ nhất.

Theo một khía cạnh của phương pháp này, enzym cải biến ADN là nucleaza định hướng vị trí được chọn từ nhóm gồm có meganucleaza (MN), nucleaza ngón tay kẽm (ZFN), nucleaza tác động giống chất kích hoạt phiên mã (TALEN), Cas9 nucleaza, Cfp1 nucleaza, dCas9-FokI, dCpf1-FokI, Cas9-xytiđin đeaminaza khảm, Cas9-ađenin đeaminaza khảm, FEN1-FokI khảm, và Mega-TAL, nickaza Cas9 (nCas9), dCas9 không phải FokI nucleaza khảm và dCpf1 không phải FokI nucleaza; và hơn nữa trong đó axit nucleic dẫn là ARN dẫn.

Theo khía cạnh khác của phương pháp này, cây con đơn bội đã được chỉnh sửa được xử lý bằng tác nhân gây nhân đôi nhiễm sắc thể, bằng cách đó tạo ra cây con đơn bội kép đã được chỉnh sửa. Ví dụ, tác nhân gây nhân đôi nhiễm sắc thể là colchixin, pronamit, đithipyr, trifluralin, hoặc chất kháng vi ống đã biết khác.

Theo khía cạnh khác của phương pháp này, cây trồng thứ nhất là cây một lá mầm hoặc cây hai lá mầm. Ví dụ, cây trồng thứ nhất là cây một lá mầm được chọn từ nhóm gồm có cây ngô, cây lúa mỳ, cây lúa, cây lúa mạch, cây yến mạch, cây cao lương tiểu hắc mạch, cây lúa miến, cây kê ngọc trai, cỏ teosinte, cây tre, cây mía, cây măng tây, cây hành, và cây tỏi. Theo khía cạnh khác, cây trồng thứ hai là cây một lá mầm hoặc cây hai lá mầm. Ví dụ cây trồng thứ hai là cây một lá mầm được chọn từ nhóm gồm có cây ngô, cây lúa mỳ, cây lúa, cây lúa mạch, cây yến mạch, cây cao lương tiểu hắc mạch, cây lúa miến, cây kê ngọc trai, cỏ teosinte, cây tre, cây mía, cây măng tây, cây hành, và cây tỏi.

Theo khía cạnh khác của phương pháp này, ARN dẫn tùy chọn là trình tự có 18–21 nucleotit và tương đồng với trình tự được chọn từ nhóm gồm có các SEQ ID NO: 2, 4, 8, 21, 23, 25, 29, 32, và 33. Theo khía cạnh khác, cây trồng thứ nhất biểu hiện gen đánh dấu. Ví dụ, gen đánh dấu được chọn từ nhóm gồm có GUS, PMI, PAT, GFP, RFP, CFP, B1, C1, R-nj, sắc tố anthoxyanin, và gen đánh dấu khác bất kỳ.

Theo khía cạnh khác của phương pháp này, cây trồng thứ nhất là cây ngô được chọn và/hoặc có nguồn gốc từ dòng Stock 6, RWK, RWS, UH400, AX5707RS, NP2222-*matl*, hoặc dòng HI bất kỳ trong số một vài dòng HI đã biết khác.

Theo một phương án, cây trồng thứ nhất và cây trồng thứ hai là các loài khác nhau. Theo một khía cạnh, cây trồng thứ nhất là cây lúa mỳ và cây trồng thứ hai là cây ngô. Theo khía cạnh khác, cây trồng thứ nhất là cây ngô và cây trồng thứ hai là cây lúa mỳ.

Một mục đích của sáng chế là cây trồng chỉnh sửa gen được tạo ra bằng phương pháp này được đề xuất.

Theo phương án khác, sáng chế đề xuất phương pháp chỉnh sửa ADN hệ gen cây trồng, có chứa các bước: (i) thu lấy cây trồng thứ nhất, trong đó cây trồng thứ nhất này có khả năng biểu hiện enzym cải biến ADN và tùy ý axit nucleic dẫn; (ii) thu lấy cây trồng thứ hai, trong đó cây trồng thứ hai này có chứa ADN hệ gen cây trồng mà cần được chỉnh sửa; (iii) thụ phấn cây trồng thứ hai với phấn từ cây trồng thứ nhất; (iv) dùng hợp phần có chứa lipit hoặc chất ức chế phospholipaza ngay lập tức trước, trong, hoặc

sau sự thụ phấn của bước (iii); và (v) chọn lọc ít nhất là một cây con đơn bội được tạo ra bởi sự thụ phấn của bước (iii) trong đó cây con đơn bội có chứa hệ gen của cây trồng thứ hai chứ không phải cây trồng thứ nhất, và hệ gen của cây con đơn bội đã được cải biến bằng enzym cải biến ADN và axit nucleic dẫn tùy chọn được chuyển bởi cây trồng thứ nhất. Theo một khía cạnh, hợp phần có chứa metyl alpha-linolenoyl flophosphonat (MALFP), etyl este của axit linoleic (LLAEE), axit linoleic (LLA), dầu ngô, đistearoyl-phosphatiđyl cholin (DSPC), metyl arachiđonyl flophosphonat (MAFP), palmityl triflometylketon (PACOCF3), arachiđonyl triflometylketon (AACOCF3), manoalit, etyl este của axit linolenic (LNAEE), etyl este của axit linolenic (LNAEE), metyl este của axit oleic (OAME), etyl este của axit oleic (OAEE), etyl este của axit palmitic (PAEE), etyl este của axit palmitoleic (PLAEE), dầu hạt lanh, dầu ngô, axit alpha-linolenic (aLNA), axit gamma-Linolenic (gLNA), axit oleic, axit arachiđonic, axit stearic, axit linoleic liên hợp 9(Z)-11(E), hoặc 2-oleoyl-1-palmitoyl-sn-glyxero-3-phospho-etanolamin.

Theo phương án khác, sáng chế đề xuất phương pháp chỉnh sửa ADN hệ gen cây trồng, có chứa các bước: (i) thu lấy cây trồng thứ nhất, trong đó cây trồng thứ nhất này có khả năng biểu hiện enzym cải biến ADN và tùy ý axit nucleic dẫn; (ii) thu lấy cây trồng thứ hai, trong đó cây trồng thứ hai này có chứa ADN hệ gen cây trồng mà cần được chỉnh sửa; (iii) lai cây trồng thứ nhất với cây trồng thứ hai; và (iv) chọn lọc ít nhất là một cây con đơn bội được tạo ra bằng phép lai của bước (iii) trong đó cây con đơn bội có chứa hệ gen của cây trồng thứ hai chứ không phải cây trồng thứ nhất, và hệ gen của cây con đơn bội đã được cải biến bằng enzym cải biến ADN và axit nucleic dẫn tùy chọn được chuyển bởi cây trồng thứ nhất. Theo một khía cạnh, cây trồng thứ nhất đóng vai trò làm thể mẹ trong phép lai của bước (iii). Theo khía cạnh khác, cây trồng thứ nhất có chứa đột biến ở gen CENH3, gen ig1, hoặc đột biến khác mang lại hệ thống kích tạo đơn bội theo dòng cha.

Ví dụ thực hiện sáng chế

I. Sản xuất dòng kích tạo đơn bội mới có chứa bộ máy chỉnh sửa.

Chúng tôi biến nạp dòng có thể biến nạp của cây ngô gọi là NP2222 với cấu trúc TALEN, và biến nạp một cách riêng rẽ dòng này với cấu trúc Cas9 và ARN dẫn. Cấu

trúc TALEN (pBSC22808 (SEQ ID NO: 5), với sự phân cắt nhắm đích TALEN trong trình tự đích, 5'-TCCAGGGTCAACGTGGAGACAGGGAGGTACGAACCGGTGACTGGCGAAGGAAGCA- 3', SEQ ID NO: 6; trình tự nhận diện TALEN được gạch chân) và cấu trúc Cas9 (pBSC23123 (SEQ ID NO: 7) với trình tự ARN dẫn của xZmPLAIIA, 5'-GGGTCAACGTGGAGACAGGG-3', SEQ ID NO: 8) được thiết kế để nhắm đích các đột biến vào exon thứ tư của gen ngô gọi là *MATRILINEAL* (*MATL*; GRAMENE ID: GRMZM2G471240). Gen này, khi được gây đột biến ở vị trí đích bằng TALEN hoặc bằng Cas9 và ARN dẫn, được làm bất hoạt, dẫn đến sự mất chức năng của sản phẩm protein. Chúng tôi đã xác minh rằng dòng mà là đồng hợp tử đối với các đột biến mất chức năng ở *MATL* là dòng kích tạo đơn bội, có nghĩa là khi chúng được sử dụng làm thể cho phấn trong phép lai, chúng gây ra sự tạo thành của thể đơn bội trên các bông thu được (xem tài liệu đơn sáng chế PCT số PCT/US2016/62548, nộp ngày 17 tháng 11 năm 2016, được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó).

Chúng tôi đã tạo ra một vài biến cố và tự thụ phấn chúng để tạo ra hạt T1. Chúng tôi trồng các cá thể T1 từ biến cố MZET152408A042A. Chúng tôi thu lấy năm cây con T1 mà giữ lại được hai bản sao của bộ máy chỉnh sửa Cas9 và ARN dẫn đã được biến nạp ổn định, và cũng là thể đột biến đồng hợp tử đối với gen *MATL*. Xem bảng 1.

Bảng 1. Dòng HI mới có chứa bộ máy chỉnh sửa hệ gen.

ID cá thể dòng HI mới	Sự có mặt của wt *MATL*	Sự có mặt của Cas9	Đột biến ở *MATL*
USR01283349	–	+	mất 13 bp, đồng hợp tử
USR01283378	–	+	mất 13 bp, đồng hợp tử
USR01283388	–	+	mất 8 bp, đồng hợp tử
USR01283391	–	+	mất 8 bp, đồng hợp tử

USR01283398	–	+	mất 13 bp, đồng hợp tử

Các đột biến *MATL* được phát hiện bằng cách sử dụng thử nghiệm TaqMan, mà khuếch đại bản sao kiểu dại của *MATL* (được đề cập đến trong bản mô tả này dưới dạng *MATL* hoặc wt-*MATL*; các thuật ngữ này được sử dụng thay thế lẫn nhau trong toàn bộ bản mô tả). Khi cả hai bản sao của *MATL* bị đột biến, thử nghiệm này đọc là âm (tức là, "–"). Bộ máy chỉnh sửa Cas9 và ARN dẫn được cài xen ổn định thông qua cấu trúc 23123 (SEQ ID NO: 7). Chúng tôi đã xác định trình tự của các đột biến ở *MATL* thông qua PCR và tách dòng phụ. Bốn khuẩn lạc của mỗi sản phẩm PCR được xác định trình tự, và tất cả các khuẩn lạc đối với cá thể nhất định có cùng trình tự, chỉ ra rằng các cây trồng này đều là thể đột biến đồng hợp tử đối với alen *MATL* (cũng được đề cập đến trong bản mô tả này dưới dạng *matl* khi tham chiếu đến đột biến cài xen 4 cặp bazơ trong *MATRILINEAL* phát hiện thấy trong Stock6 và các dòng có nguồn gốc từ Stock6 khác, hoặc *μMATL* khi tham chiếu đến đột biến do con người tạo ra khác bất kỳ ở *MATRILINEAL*;). Có hai cây trồng mà có đột biến mất 8 bp, và ba cây trồng mà có đột biến mất 13 bp.

II. Sử dụng dòng HI mới làm thể đực và phân tích con cháu.

Chúng tôi đã lai cây trồng HI mới ở trên làm thể cho phấn đực với dòng thử cái, mà chứa thành phần đánh dấu màu sắc lặn nhưng là kiểu dại đối với gen *MATL*. Dòng kích tạo đơn bội thể đực là kiểu dại đồng hợp tử đối với cùng thành phần đánh dấu màu sắc này. Do đó dòng cái này là thể kích tạo không-đơn bội và là kiểu dại đồng hợp tử đối với gen *MATL* nhưng là thể đột biến đồng hợp tử đối với thành phần đánh dấu màu sắc. Chúng tôi thu lấy hạt từ phép lai, và cho nảy mầm thành cây con từ đó.

Cây con con cháu được đưa vào một vài thử nghiệm. Cây con con cháu được tính điểm là thể lưỡng bội nếu chúng không thể hiện thành phần đánh dấu màu sắc (vì thành phần đánh dấu lặn được bù bởi ADN thể kích tạo đực). Cây con con cháu được tính điểm là thể đơn bội giả định nếu chúng thể hiện thành phần đánh dấu màu sắc vì thành phần đánh dấu lặn không được bù. Trong số 2656 hạt được trồng, chúng tôi sử dụng thử nghiệm màu sắc và xác định 90 cây con là thể đơn bội giả định.

Chúng tôi phân tích thêm 90 thể đơn bội giả định này về sự có mặt của gen *MATL* kiểu dại bằng cách sử dụng thử nghiệm thành phần đánh dấu Taqman. Trong số đó, 82 thể đơn bội dương tính đối với *MATL*, có nghĩa là chúng không được chỉnh sửa bằng bộ máy chỉnh sửa cung cấp bởi thể đực. 8 cây con đơn bội giả định còn lại âm tính đối với *MATL* kiểu dại sử dụng thành phần đánh dấu Taqman, chỉ ra rằng chúng có thể đã được chỉnh sửa bằng bộ máy chỉnh sửa cung cấp bởi thể đực.

Chúng tôi thực hiện phân tích mức bội thể thông qua đếm tế bào theo dòng trên 8 cây con đơn bội đã được chỉnh sửa, giả định này bằng cách sử dụng mô lá trong thiết bị phân tích mức bội. Xem các hình 1–8. Chúng tôi phát hiện thấy rằng bốn trong số chúng là thể đơn bội thực, trong khi các cây con khác thực sự là thể lưỡng bội. Như thảo luận dưới đây, chúng tôi chạy PCR và xác định trình tự các đột biến trong gen *MATL* trong bốn thể đơn bội thực này cũng như là đối với cây trồng USR01350-337-2 mà, theo thử nghiệm Taqman *MATL*, không được chỉnh sửa bằng bộ máy chỉnh sửa hệ gen.

Phát hiện về việc có bốn thể lưỡng bội trong số 90 thể đơn bội giả định không phải là bất ngờ, thử nghiệm cây con không hoàn hảo và đôi khi có dương tính giả. Chúng tôi thử nghiệm 90 thể đơn bội này về sự có mặt của cấu trúc Cas9 (cấu trúc 23123), và phát hiện thấy nó bị mất đi ở 86 trong số 90 thể đơn bội này, bao gồm bốn thể đơn bội thực nêu trên. Ngược lại, bốn thể lưỡng bội đã được chỉnh sửa mà chúng tôi phát hiện thấy trong quá trình phân tích mức bội thể đều có mặt cấu trúc Cas9, xác nhận tình trạng của chúng là thể lưỡng bội lai mà được xác định nhầm bởi thử nghiệm cây con đơn bội là thể đơn bội.

Sau đó chúng tôi sử dụng mô lá để phân lập ADN hệ gen và chạy phản ứng PCR để xác định trình tự gen *MATL* trong bốn cá thể đã được chỉnh sửa giả định, đơn bội thực này, cụ thể là tập trung vào trình tự ở bên cạnh vị trí gây đột biến đích ARN dẫn. Điều này là để xác định bản chất của sự chỉnh sửa mà có thể xảy ra hoặc không xảy ra ở đó. Chúng tôi tách dòng phụ mảnh PCR bằng cách sử dụng bộ kit TOPO Blunt IV có sẵn trên thị trường, và xác định trình tự ít nhất là bốn khuẩn lạc mỗi lần (xác định trình tự xuôi và ngược). Xem bảng 2, dưới đây, về việc so sánh alen đã được chỉnh sửa và alen wt-MATL tham chiếu.

Bảng 2. So sánh các alen đã được chỉnh sửa với wt-*MATL*.

ID cá thể	Kiểu alen	Trình tự (tương ứng với 1126-1166 của SEQ ID No: 19)
NP2222	wt-*MATL*	AGGGTCAACGTGGAGACAGGGAGGTACGAACCGGTGACTGG
Stock6	*matl*	AGGGTCAACGTGGAGACAGGCGAGGAGGTACGAACCGGTGACTGG
USR01350333-3 Alen 1	đã chỉnh sửa	AGGGTCAACGTGGAGACAAGGGAGGTACGAACCGGTGACTGG
USR01350333-3 Alen 2	tạp nhiễm PCR	AGGGTCAACGTGGA::::::::::::GAACCGGTGACTGG
USR01350344-2 Alen 1	đã chỉnh sửa	AGGGTCAACGTGGAGAC:GGGAGGTACGAACCGGTGACTGG
USR01350344-2 Alen 2	tạp nhiễm PCR	AGGGTCAACGTGGA::::::::::::GAACCGGTGACTGG
USR01350343-1 Alen 1	đã chỉnh sửa	AGGGTCAACGTGGAGACAAGGGAGGTACGAACCGGTG
USR01350328-1 Alen 1	đã chỉnh sửa	AGGGTCAACGTGGAGAC:GGGAGGTACGAACCGGTGACTGG

USR01350337-2 Alen 1	chưa chỉnh sửa	AGGGTCAACGTGGAGACAGGGAGGTACGA ACCGGTGACTGG
USR01350337-2 Alen 2	tạp nhiễm PCR	AGGGTCAACGTGGA::::::::::::GAACCGGTGACTGG

Cá thể USR01350333-3 tạo ra alen *MATL* đã chỉnh sửa bằng đột biến cài xen của alanin ở cặp bazơ 1143 của trình tự cADN (được gạch chân trong bảng 2). Điều này là đủ để gây ra sự dịch khung trong trình tự mã hóa, mà sẽ tạo ra bộ ba kết thúc sớm. Điều chúng tôi đã nghĩ tới là alen đã được chỉnh sửa #2 của USR01350333-3 (đột biến mất 13 cặp bazơ của GACAAGGGAGGTAC) thực sự là kết quả của sự tạp nhiễm PCR. Sau khi xác định trình tự lại, chúng tôi xác nhận rằng cây trồng này chỉ có một alen đã được chỉnh sửa, và nó được phát hiện thấy ở 6 trong số 6 khuẩn lạc.

Alen này là mới trong đó nó không phải là ở cây bố hoặc cây mẹ của cá thể này. ID bố đối với cá thể này là USR01283391, và cây trồng được phát hiện là đồng hợp tử đối với đột biến mất 8 bp.

Cá thể USR01350344-2 cung cấp đột biến mất của A (đột biến mất của cặp bazơ 1143 của trình tự cADN kiểu dại). Đột biến này là đủ để gây ra sự dịch khung trong trình tự mã hóa, và tạo ra bộ ba kết thúc sớm. Sau khi xác định trình tự lại và tìm ra sự tạp nhiễm PCR, chúng tôi xác nhận nó được phát hiện ở 6 trong số 6 khuẩn lạc. Như đã xác định từ trước là alen đã được chỉnh sửa #2 của USR01350344-2, nó được xác định là sự tạp nhiễm PCR.

Cá thể USR01350343-1 cung cấp đột biến cài xen của A ở cặp bazơ 1143 của trình tự cADN. Điều này là đủ để gây ra sự dịch khung trong trình tự mã hóa, và tạo ra bộ ba kết thúc sớm. Nó được phát hiện ở 4 trong số 4 khuẩn lạc.

Cá thể USR01350328-1 cung cấp đột biến mất của A (đột biến mất của cặp bazơ 1143 từ trình tự cADN kiểu dại). Đột biến này là đủ để gây ra sự dịch khung trong trình tự mã hóa, và tạo ra bộ ba kết thúc sớm. Nó được phát hiện ở 4 trong số 4 khuẩn lạc.

Cá thể USR01350337-2 không có sự thay đổi nào: trình tự của nó tương đồng 100% với trình tự của wt-*MATL*.

Tóm lại, chúng tôi phát hiện thấy rằng 4 trong số 86 thể đơn bội đã được xác nhận có các đột biến trong gen *MATL*. Chúng tôi đã xác nhận rằng các cây trồng này là thể đơn bội và không chứa bất kỳ ADN Cas9 nào. Rõ ràng rằng gen chuyển Cas9 đã bị loại bỏ cùng với phần còn lại của ADN có nguồn gốc từ bố trong quá trình tạo phôi, và sự chỉnh sửa xảy ra đối với hệ gen mẹ (có nguồn gốc từ tế bào giao tử cái) trong quá trình tạo phôi.

Chúng tôi biết rằng sự chỉnh sửa này là mới và xảy ra ở hệ gen mẹ trong quá trình tạo phôi vì dòng kích tạo đơn bội thường tạo ra thể đơn bội theo dòng mẹ và chúng tôi đã xác nhận rằng chúng quả thật là thể đơn bội. Người ta có thể tranh luận rằng có khả năng chúng thực sự là thể đơn bội theo dòng cha, và sự chỉnh sửa chúng tôi nhìn thấy thực sự là sự chỉnh sửa mà đã có mặt trong ADN theo dòng cha. Tuy nhiên, chúng tôi có thể chứng minh rằng đó không phải là trường hợp này. Thứ nhất, các đột biến này không phù hợp với các đột biến của bố mẹ dòng bố. Điều này có thể thấy rõ ràng trong bảng 3 và bảng 4 (được thể hiện dưới đây). Cây trồng đơn bội đã chỉnh sửa USR01350343-1 là đồng hợp tử đối với đột biến cài xen của nucleotit đơn lẻ ("A"), nhưng cây trồng bố có đột biến mất của 13 nucleotit. Tương tự, cây trồng USR01350328-1 là đồng hợp tử đối với đột biến mất của A, nhưng bố có đột biến mất của 13 nucleotit. Các ví dụ này, cùng với nhau, chứng minh rằng trong quy trình kích tạo đơn bội, có thể có sự chỉnh sửa của hệ gen dòng mẹ xảy ra, dẫn đến sự tạo thành của thể đơn bội theo dòng mẹ đã chỉnh sửa. Theo đó và dựa trên thử nghiệm phát hiện sự có mặt của *MATL* và sự xác nhận thông qua phân tích mức bội thể, và bằng cách sử dụng gen chuyển Cas9 ở phía thể đực chịu sự kiểm soát của vùng khởi động ubiquitin của cây ngô, tỉ lệ của sự chỉnh sửa trong quy trình kích tạo đơn bội là khoảng 4/86, hay 4,65%.

Hơn nữa, tỉ lệ của sự chỉnh sửa trong quá trình kích tạo đơn bội có thể rất khác nhau khi sử dụng các dòng kích tạo đơn bội khác nhau hoặc sử dụng phép lai xa. Dường như là cả sự kích tạo đơn bội ở cây ngô bằng cách sử dụng dòng đột biến *MATL* và phép lai xa ở cây lúa mạch, cây lúa mỳ, hoặc các cây trồng khác đều hoạt động thông qua cơ chế tương tự: thụ tinh sau đó là loại bỏ hệ gen. Cũng có vẻ là khoảng thời gian giữa thụ

tinh và loại bỏ hệ gen đủ dài để bộ máy chỉnh sửa chỉnh sửa gen đích trong hệ gen của dòng mà dòng kích tạo được lai hóa với (chất mầm nguyên sinh đích). Lưu ý rằng việc lựa chọn vùng khởi động gây ra sự biểu hiện của hệ thống protein chỉnh sửa đã được biến nạp ổn định có thể có ảnh hưởng lớn đến tỉ lệ chỉnh sửa ở thể đơn bội. Chúng tôi sử dụng vùng khởi động cấu trúc của mía (prSoUbi4) nhưng vùng khởi động khác gây ra sự biểu hiện cao hoặc đặc hiệu trong túi phôi, tế bào giao tử cái, trong phấn, hoặc trong tế bào giao tử đực có thể hiệu quả hơn, cụ thể là trong trường hợp của phép lai xa, trong đó ADN thể đực bị loại bỏ theo kiểu nhanh và mạnh hơn rất nhiều so với ở hệ thống thể kích tạo đơn bội cùng loài như hệ thống thể kích tạo đơn bội cây ngô hoặc hệ thống thể kích tạo đơn bội kiểu *CENH3* . Nói cách khác, trong quá trình lai xa, ví dụ như khi lai phấn ngô với bông lúa mỳ, mà được thực hiện để tạo ra thể đơn bội theo dòng mẹ của cây lúa mỳ, nó có thể hoạt động tốt nhất để có bộ máy chỉnh sửa trong phấn ngô được điều khiển bởi vùng khởi động mà có sự biểu hiện phấn hoặc tế bào giao tử đực mạnh, có lẽ là ngoài sự biểu hiện hợp tử, sao cho bộ máy chỉnh sửa (ARN và protein) dồi dào được chuyển và có mặt trong tế bào hợp tử và trong giai đoạn hai, bốn, hoặc tám tế bào phôi sau đó, ngay cả khi ADN thể đực bị loại bỏ hoặc bị mất đi rất nhanh.

Bảng 3. Con cháu đơn bội được tạo ra

Mã ID con cháu cá thể	Sự có mặt của wt *MATL*	Phân tích mức bội thể	Sự có mặt của Cas9	Alen 1
USR01350333-3	–	đơn bội	–	đột biến cài xen của A
USR01350344-2	–	đơn bội	–	đột biến mất của A

USR01350343-1	–	đơn bội	–	đột biến cài xen của A
USR01350328-1	–	đơn bội	–	đột biến mất của A
USR01350337-2	+	đơn bội	–	không đột biến
USR01350334-3	–	lưỡng bội	+	
USR01350333-10	–	lưỡng bội	+	
USR01350341-1	–	lưỡng bội	+	
USR01350321-3	–	lưỡng bội	+	

Bảng 4. Thông tin về bố và con cháu của chúng

ID Bố	Sự có mặt của wt *MATL*	Xác định trình tự (# khuẩn lạc)	Sự có mặt của Cas9	ID con cháu
USR01283391	–	đột biến mất của 8 nt (4)	+	USR01350333-3 và USR01350333-10

USR01283349	–	đột biến mất của 13 nt (4)	+	USR01350344-2, USR01350328-1 và USR01350321-3
USR01283378	–	đột biến mất của 13 nt (4)	+	USR01350343-1 và USR01350341-1
USR01283398	–	đột biến mất của 13 nt (4)	+	USR01350337-2
USR01283388	–	đột biến mất của 8 nt (4)	+	USR01350334-3

III. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa ở các dòng lai cùng dòng của cây ngô cao cấp.

Dòng kích tạo đơn bội có thể biến nạp, NP2222-HI, RWK, RWS, hoặc UH400 hoặc Stock6 hoặc dòng kích tạo đơn bội khác bất kỳ, tất cả chúng đã có các phiên bản đột biến của *MATL*, được biến nạp ổn định với cấu trúc biểu hiện hệ thống cải biến hệ gen chẳng hạn như Cas9 + ARN dẫn (Cong, L. *et al. 2013.* Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823), dCas9-FokI+ ARN dẫn (Tsai, S. Q. *et al.* 2014, Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature Biotechnol.* 32, 569–576), TALEN (Li et al., 2012, High-efficiency TALEN-based gene editing produces disease-resistant rice. Nature Biotech. 30, 390–392), meganucleaza được thiết kế (Gao et al., 2010, Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant Journal.* 61:176–187), nucleaza ngón tay kẽm (Shukla et al. 2009. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature 459, 437–441), dCas9-xytiđin đeaminaza (Komor et al. 2016, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature doi:10,1038/nature17946) hoặc hệ

thống cải biến hệ gen khác bất kỳ. Dòng kích tạo đơn bội chuyển gen cũng biểu hiện bộ máy chỉnh sửa sau đó được sử dụng làm thể cho phấn để tạo ra các đột biến và thể đơn bội ở dòng đích thông qua lai xa. Sau đó phôi hoặc hạt đơn bội được thu hồi, được xác định là thể đơn bội, và được thử nghiệm đối với sự chỉnh sửa ở vị trí đích (bất kể vị trí đích được chọn theo thiết kế cấu trúc TALEN hay thiết kế Cas9 ARN dẫn). Thể đơn bội chứa sự chỉnh sửa mong muốn được nhân đôi theo nhiễm sắc thể bằng cách sử dụng các quy trình tiêu chuẩn bằng cách sử dụng phương tiện tiêu chuẩn chẳng hạn như colchixin, trifluralin hoặc tác nhân gây nhân đôi nhiễm sắc thể khác. Việc xác định thể đơn bội được tạo ra có thể được đơn giản hóa bằng cách sử dụng thành phần đánh dấu màu sắc như thường được thực hiện trong sự sản xuất thể đơn bội kép ở ngô -thành phần đánh dấu màu sắc này có thể hiển thị trong phôi, hạt, cây con, hoặc cây trưởng thành thu được. Sự có mặt của các đột biến ở vị trí đích có thể được kiểm tra bằng cách phân tích trình tự (xác định trình tự ADN), bằng cách phân tích thành phần đánh dấu, hoặc bằng kiểu hình. Vì chỉ có một bản sao của ADN để gây đột biến ở cây đơn bội, kiểu hình lặn sẽ hiển thị do đó có thể là cách khác để xác định thể đơn bội mà đã được chỉnh sửa.

A. Gây đột biến đích VLHP ở dòng lai cùng dòng ngô cao cấp với locut chỉnh sửa chuyển gen được tạo ra trực tiếp ở dòng kích tạo đơn bội.

VLHP1 và VLHP2 là các gen homeodomain-khóa kéo lơxin lớp-I homeobox và thành viên của lớp protein mà là duy nhất đối với cây trồng. Miền HD tham gia vào việc liên kết ADN trong khi miền Zip tham gia vào sự đime hóa đồng loại và dị loại của protein. Protein HD-Zip I thường tham gia vào các đáp ứng liên quan đến áp lực không sống, axit abscisic (ABA), ánh sáng xanh, sự khử úa vàng và sự tạo phôi (Elhiti and Stasolla, 2009. Structure and function of homodomain-leucine zipper (HD-Zip) proteins. *Plant Signal Behav.* 4: 86-88). VLHP1 và VLHP2 ở trong cùng họ gen với Grassy Tillers1 (GT1). GT1 thúc đẩy trạng thái tiềm sinh chồi bên và kìm hãm sự kéo dài của nhánh bông bên ở cây ngô.

Trong ví dụ này, véc-tơ 23396 (SEQ ID NO: 1; cũng xem hình 9) để biểu hiện Cas9 và ARN dẫn đơn lẻ (sgRNA) được tạo ra để nhắm đích các gen VLHP1 ngô (GRMZM2G104204) và thể tương đồng của nó VLHP2 (GRMZM2G062244). Véc-tơ 23396 biểu hiện sgRNA với trình tự nhắm đích có 20-nucleotit xZmVLHP-01 (5'-

GCAGGAGGCGTCGAGCAGCG-3', SEQ ID NO: 2). xZmVLHP-01 nhắm đích cả hai gen VLHP1 và VLHP2 ở exon thứ hai. Véc-tơ 23396 được đưa vào dòng kích tạo đơn bội có thể biến nạp NP2222-HI bằng cách sử dụng sự biến nạp qua trung gian *Agrobacterium* với sự chọn lọc manoza. NP2222-HI được tạo ra từ phép lai của dòng lai cùng dòng ngô có thể biến nạp NP2222 với dòng có nguồn gốc từ Stock 6 RWKS để thẩm nhập locut kích tạo đơn bội (HI). NP2222-HI có tỉ lệ kích tạo đơn bội trung bình bằng khoảng 9,2%.

Các thể biến nạp NP2222-HI từ véc-tơ 23396 được thử nghiệm đối với sự cải biến của các trình tự đích VLHP hệ gen (5'-GCAGGAGGCGTCGAGCA/GCG-3'; SEQ ID NO: 2). Dấu gạch chéo ("/") thể hiện vị trí phân cắt Cas9. Hoạt tính chỉnh sửa locut đích được xác định bằng cách sử dụng phương pháp PCR Taqman định lượng như đã mô tả trước đây (tài liệu công bố đơn WO2016106121, được kết hợp trong bản mô tả này để tham khảo). Dòng chuyển gen với hoạt tính cải biến vị trí đích cao, tức là, cả hai gen VLHP1 và VLHP2 được cải biến, và tốt hơn là chứa bản sao gen chuyển đơn lẻ được chọn để nghiên cứu thêm và dùng để lai hoặc sản xuất cây con.

Phấn từ các thể biến nạp T0 của 23396 được sử dụng trực tiếp để thụ phấn bông của dòng lai cùng dòng cao cấp ID5829 hoặc dòng ngô khác bao gồm dòng ngô ngọt để gây ra sự sản xuất của phôi đơn bội. Theo cách khác, các thể biến nạp T0 của 23396 trong nền tảng NP2222-HI được tự thụ phấn để tạo ra dòng con cháu mang gen chuyển đồng hợp tử và phấn từ cây trồng con được sử dụng để thụ phấn dòng ngô khác để gây ra sự tạo thành phôi đơn bội. Phôi đơn bội đã được tạo ra được chiết xuất từ nhân và được đặt trên môi trường cứu phôi để nảy mầm trực tiếp hoặc cho phép trưởng thành để tạo thành hạt. Mô từ phôi đơn bội đã được tạo ra và cây trồng thu được được thử nghiệm để xác định liệu sự chỉnh sửa đã xảy ra trong trình tự đích VLHP hay chưa. Nếu phôi đơn bội đã được tạo ra hoặc cây trồng chứa các đột biến mong muốn, sự xử lý nhân đôi nhiễm sắc thể được dùng để tạo ra dòng đơn bội kép từ chúng. Ví dụ, bằng cách sử dụng phương pháp cứu phôi, phôi được chiết xuất từ bông dòng cao cấp ID5829 được thụ phấn với dòng kích tạo đơn bội chuyển gen mang locut chỉnh sửa 23396 vào 18–22 ngày sau thụ phấn ("DAP," sự chiết xuất giữa 10–25 DAP là có thể được về mặt lý thuyết). ADN được phân lập từ cây con đơn bội đã nảy mầm và dùng để thử nghiệm. Việc xử lý colchixin được áp dụng cho cây con để nhân đôi nhiễm sắc thể. Theo cách khác, tác

nhân gây nhân đôi nhiễm sắc thể có thể được dùng cho phôi đã được phân lập trong quá trình nảy mầm. ADN được chiết xuất từ cây con đã được nảy mầm và dùng để xác định rằng đột biến có xảy ra ở trình tự đích xZmVLHP-01 hay không.

Các phương pháp thay thế là có sẵn. Người ta có thể cho phép hạt chín và chọn lọc thể đơn bội sau bằng kiểu hình khác. Người ta có thể để cho hạt khô đi và vào ngày hôm sau làm nảy mầm hạt để xác định thể đơn bội không có thành phần đánh dấu (ví dụ, bằng cách sử dụng kích thước cây trồng chứ không phải là gen mang lại thành phần đánh dấu màu sắc), tại thời điểm đó người ta sẽ thử nghiệm về sự chỉnh sửa và dùng tác nhân gây nhân đôi nhiễm sắc thể khi thích hợp. Phương pháp này có thể có ưu điểm của nó trong đó tránh được việc sàng lọc và/hoặc cứu phôi.

B. Gây đột biến đích GW2 ở dòng lai cùng dòng ngô cao cấp với locut chỉnh sửa chuyển gen được đưa vào trực tiếp ở dòng kích tạo đơn bội.

Đột biến ở DA2, gen E3-ubiquitin ligaza, ở cây lúa dẫn đến hạt lớn hơn (Song et al., 2007). DA2 lúa có 2 thể tương đồng ngô, GW2-1 (GRMZM2G170088) và GW2-2 (GRMZM2G007288). Các gen ngô tương đồng 94% ở cấp độ protein và tương đồng 90% ở cấp độ ADN. GRMZM2G170088 có đoạn cài xen 177 bp lớn (59 aa) khi so với GRMZM2G007288.

Trong ví dụ này, véc-tơ 23399 (SEQ ID NO: 3, cũng xem hình 10) được tạo ra để biểu hiện Cas9 và sgRNA để nhắm đích cả hai gen GW2-1 ngô (GRMZM2G170088) và thể tương đồng của nó GW2-2 (GRMZM2G007288). Cả hai gen GW2-1 và GW2-2 cùng chứa trình tự đích xZmGW2-02 (5'-AAGCTCGCGCCCTGCTACCC-3', SEQ ID NO: 4) trong exon 1 và trình tự này được sử dụng để thiết kế sgRNA được biểu hiện từ véc-tơ 23399. Véc-tơ hai thành phần 23399 biểu hiện ARN dẫn đơn lẻ (sgRNA) với trình tự nhắm đích có 20-nucleotit xZmGW2-02 dung hợp với giá đỡ ARN dẫn đơn lẻ gồm có cả crRNA và tracrRNA. Véc-tơ 23399 được đưa vào dòng kích tạo đơn bội có thể biến nạp NP2222-HI bằng cách sử dụng sự biến nạp qua trung gian *Agrobacterium* với sự chọn lọc manoza. NP2222-HI được tạo ra từ phép lai của dòng lai cùng dòng ngô có thể biến nạp NP2222 với dòng có nguồn gốc từ Stock 6 RWKS để thẩm nhập locut kích tạo đơn bội (HI).

Các thể biến nạp NP2222-HI của véc-tơ 23399 được thử nghiệm đối với sự cải biến của trình tự đích GW2-2 hệ gen (5'-AAGCTCGCGCCCTGCTA/CCC-3', SEQ ID NO: 4; dấu gạch chéo ("/") chỉ ra vị trí phân cắt Cas9). Hoạt tính chỉnh sửa trình tự đích được xác định bằng cách sử dụng phương pháp PCR Taqman định lượng như đã mô tả trước đây (WO2016106121). Dòng chuyển gen với hoạt tính cải biến vị trí đích cao, tức là, cả hai gen GW2-1 và GW2-2 được cải biến, và tốt hơn là chứa bản sao gen chuyển đơn lẻ được chọn để nghiên cứu thêm và dùng để lai hoặc sản xuất con cháu.

Phấn từ các thể biến nạp T0 của 23399 được sử dụng trực tiếp để thụ phấn bông của dòng lai cùng dòng cao cấp ID5829 hoặc dòng ngô khác bao gồm dòng ngô ngọt để gây ra sự sản xuất của phôi đơn bội. Theo cách khác, các thể biến nạp T0 của 23399 trong nền tảng NP2222-HI được tự thụ phấn để tạo ra dòng con cháu mang gen chuyển đồng hợp tử và phấn từ cây trồng con được sử dụng để thụ phấn dòng ngô khác để gây ra sự tạo thành phôi đơn bội. Phôi đơn bội đã được tạo ra được chiết xuất từ nhân và được đặt trên môi trường cứu phôi để nảy mầm trực tiếp hoặc cho phép trưởng thành để tạo thành hạt. Mô từ phôi đơn bội đã được tạo ra và cây trồng thu được được thử nghiệm để xác định liệu sự chỉnh sửa đã xảy ra trong trình tự đích GW2 ở ngô hay chưa. Nếu phôi đơn bội đã được tạo ra hoặc cây trồng chứa các đột biến mong muốn, sự xử lý nhân đôi nhiễm sắc thể được dùng để tạo ra dòng đơn bội kép từ chúng. Ví dụ, bằng cách sử dụng phương pháp cứu phôi, phôi được chiết xuất từ bông dòng cao cấp ID5829 được thụ phấn với dòng kích tạo đơn bội chuyển gen mang locut chỉnh sửa 23396 vào 18-22 ngày sau thụ phấn. ADN được phân lập từ cây con đơn bội đã nảy mầm và dùng để thử nghiệm. Việc xử lý colchixin được áp dụng cho cây con để nhân đôi nhiễm sắc thể. Theo cách khác, tác nhân gây nhân đôi nhiễm sắc thể có thể được dùng cho phôi đã được phân lập trong quá trình nảy mầm. ADN được chiết xuất từ cây con đã được nảy mầm và dùng để xác định liệu đột biến có xảy ra ở trình tự đích xZmGW2-02 hay không. Theo cách khác, người ta có thể cho phép hạt chín và chọn lọc thể đơn bội sau đó bằng kiểu hình khác. Người ta thậm chí có thể để cho hạt khô đi và vào ngày hôm sau làm nảy mầm hạt để xác định thể đơn bội không có thành phần đánh dấu (ví dụ, bằng cách sử dụng kích thước cây trồng chứ không phải là gen mang lại thành phần đánh dấu màu sắc), tại thời điểm đó người ta sẽ thử nghiệm về sự chỉnh sửa và dùng tác nhân gây nhân đôi nhiễm sắc thể khi thích hợp. Phương pháp này có thể có ưu điểm của nó trong đó tránh được việc sàng lọc và/hoặc cứu phôi.

IV. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa ở cây ngô, cây lúa, cây hướng dương, hoặc cây trồng khác bất kỳ thông qua sự kích tạo đơn bội dựa trên hóa chất

Dòng bất kỳ của cây ngô, cây lúa, cây lúa mỳ, cây cà chua, cây hướng dương, cây lúa mạch, hoặc cây trồng khác bất kỳ có thể biến nạp được bằng cấu trúc chỉnh sửa (Cas9 cộng với ARN dẫn được thiết kế để gây đột biến vị trí đích cụ thể) và sau đó tùy ý tạo ra cấu trúc chỉnh sửa dị hợp tử hoặc đồng hợp tử (thông qua sự tự thụ phấn của biến cố đã được biến nạp), và sau đó sử dụng lipit hoặc sử dụng dầu trong quá trình lai xa (thụ phấn trên các dòng đích) để tạo ra thể đơn bội mới và đồng thời chỉnh sửa các vị trí đích trong hệ gen đích. Việc dùng lipit này có khả năng tạo ra thể đơn bội khi được dùng cho phấn, râu, hoa, hoặc cờ của cây trồng bất kỳ, bất kể bố thế nào. Cụ thể là, thể đực không cần phải có đột biến bất kỳ trong gen *MATL* (tức là, nó có thể là kiểu dại đồng hợp tử đối với gen *MATRILINEAL*). Việc dùng lipit này tạo ra thể đơn bội mới, không có yêu cầu di truyền bất kỳ nhân danh bố hoặc mẹ. Xem đơn sáng chế PCT số PCT/US2016/62548, được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó. Cơ chế của sự kích tạo đơn bội mới thông qua việc phun dịch lỏng dường như hoạt động theo cùng một cách như khi nó hoạt động ở các dòng đột biến *matl* (thể kích tạo đơn bội di truyền): thông qua sự loại bỏ nhiễm sắc thể sau thụ tinh. Con cháu đơn bội được phân lập và kiểm tra đối với các đột biến được tạo ra (được tạo ra thông qua quy trình chỉnh sửa) và sau đó nhân đôi để tạo ra cây đơn bội đã được nhân đôi, đã chỉnh sửa.

V. Gây đột biến trình tự đích ở các dòng ngô làm thức ăn gia súc cao cấp và các dòng lai cùng dòng ngô ngọt với locut chỉnh sửa chuyển gen được thẩm nhập vào dòng kích tạo đơn bội.

Locut chuyển gen biểu hiện bộ máy chỉnh sửa hệ gen cũng có thể được tạo ra ở dòng ngô có thể biến nạp thông thường không có hoạt tính tạo thể đơn bội chẳng hạn như A188, Hi-II hoặc NP2222 và sau đó được thẩm nhập vào dòng kích tạo đơn bội chẳng hạn như NP2222-HI, RWK, RWKS, RWS, hoặc UH400 hoặc Stock6 hoặc dòng kích tạo đơn bội khác bất kỳ.

Trong ví dụ này, dòng lai cùng dòng ngô NP2222 được biến nạp với véc-tơ VLHP Cas9-sgRNA (23396 và 23397) và véc-tơ GW2 Cas9-sgRNA (23398 và 23399). Các véc-tơ 23396 và 23399 đã được mô tả trong các ví dụ trước (ví dụ IIIA và ví dụ IIIB).

Véc-tơ 23397 (SEQ ID NO: 20) tương đồng với 23396 ngoại trừ trình tự mã hóa gARN xZmVLHP-01 (5'-GCAGGAGGCGTCGAGCAGCG-3', SEQ ID NO: 2) được thay thế bằng xZmVLHP-02 (5'-GCTGGAGCTGAGCTTCCGGG-3', SEQ ID NO: 21). Véc-tơ 23398 (SEQ ID NO: 23) tương đồng với 23399 ngoại trừ trình tự mã hóa gARN xZmGW2-02 (5'-AAGCTCGCGCCCTGCTACCC-3', SEQ ID NO: 4) trong 23399 được thay thế bằng xZmGW2-01 (5'-GAGCGGTTCACGCGGCCGCA-3', SEQ ID NO: 23). Các véc-tơ này được đưa vào chủng *Agrobacterium* LBA4404 (pVGW7). Chủng *Agrobacterium* thu được chứa véc-tơ 23396, 23397, 23398, hoặc 23399 được sử dụng để biến nạp phôi chưa trưởng thành của dòng lai cùng dòng cao cấp có thể biến nạp NP2222. Thể sần được tạo ra từ phôi chưa trưởng thành đã được lây nhiễm và được chọn trên môi trường manoza để thu hồi thể sần chuyển gen. Thể sần chuyển gen được đặt trên môi trường tái tạo và ra rễ để thu hồi cây trồng chuyển gen biểu hiện bộ máy chỉnh sửa CRISPR-Cas9. Cây trồng chuyển gen được thử nghiệm về số lượng bản sao gen chuyển và chuyển vào nhà kính để sản xuất hạt.

Các thể biến nạp có bản sao đơn lẻ của véc-tơ 23396 (MZET154902A004A, MZET154902B006A), 23397 (MZET154903B009A, MZET154903B012A), 23398 (MZET154904B005A, MZET154904B014A) và 23399 (MZET154905A002A, MZET154905A010A) được xác định và lai ngược với NP2222 không chuyển gen. Bông của cây trồng con chuyển gen chứa đột biến cài xen T-ADN của mỗi véc-tơ nêu trên được thụ phấn với phấn của dòng kích tạo đơn bội RWKS để tạo ra cây con F1. Cây con F1 chứa locut chuyển gen và locut kích tạo đơn bội được xác định bằng thử nghiệm xác định kiểu gen và tự thụ phấn để tạo ra hạt con cháu F2. Hạt con cháu F2 được trồng và cây trồng con được thử nghiệm để xác định cây trồng đồng hợp tử đối với locut Cas9-sgRNA chuyển gen (thử nghiệm #2540) và locut kích tạo đơn bội (thử nghiệm #2827) bằng thử nghiệm qPCR Taqman.

Các dòng đồng hợp tử đối với locut kích tạo đơn bội và tốt hơn là đồng hợp tử locut chỉnh sửa 23396, 23397, 23398, và 23399 Cas9-sgRNA chuyển gen được sử dụng để thụ phấn bông từ dòng ngô làm thức ăn gia súc cao cấp đích ID5829 và dòng ngô ngọt (SWC726 hoặc SWC412F) để kích tạo đơn bội. Phôi đơn bội đã được tạo ra được phân lập từ bông ID5829, SWC412F, SWC726 đã thụ phấn và nảy mầm trên môi trường cứu phôi. Theo cách khác, bông đã thụ phấn được để cho trưởng thành và nhân có phôi

đơn bội được cho nảy mầm. Mẫu lá được thu gom và phân tích bằng thử nghiệm Taqman để xác định cây trồng chứa các đột biến ở các gen VLHP và GW2 nhưng không có mặt các thành phần di truyền từ dòng cảm ứng chẳng hạn như Cas9-sgRNA chuyển gen hoặc trình tự gen đánh dấu không chuyển gen khác. Cây đơn bội đã được xác định với các đột biến gen GW2 hoặc VLHP đã được nhắm đích được xử lý bằng colchixin để nhân đôi nhiễm sắc thể để thu hồi cây đơn bội đã được nhân đôi để sản xuất hạt. Theo cách khác, phôi đơn bội đã được chiết xuất có thể được xử lý bằng tác nhân gây nhân đôi nhiễm sắc thể chẳng hạn như colchixin và cây trồng thu được được phân tích về mức bội thể và sự có mặt của các đột biến được nhắm đích ở các gen GW2 hoặc VLHP. Cây trồng có các đột biến gen GW2 và VLHP đã được nhắm đích được cho sinh trưởng đến khi trưởng thành để sản xuất hạt và sau đó đánh giá con cháu.

Ví dụ, dòng đơn bội đã chỉnh sửa (JSER82A056 và JSER82A063) được xác định từ phép lai giữa bông dòng ngô ngọt SWC412F được thụ phấn với thể kích tạo đơn bội chứa gen chuyển 23399 Cas9-sgRNA. Dòng JSER82A056 có cả hai gen đích GW2-01 và GW2-02 bị đột biến, trong khi dòng JSER82A063 chỉ có gen GW2-02 bị đột biến (xem bảng 5). Không dòng nào trong số các dòng đó chứa gen chuyển Cas9 (thử nghiệm #2540 đối với Cas9 hoặc #1750 đối với thành phần đánh dấu chọn lọc PMI) hoặc gen thể kích tạo đơn bội (thử nghiệm #2827) khi hệ gen đực đã bị loại bỏ khỏi thể đơn bội. Phân tích mức đa bội thể xác nhận rằng cả hai dòng là thể đơn bội (hình 11 và hình 12). Lưu ý rằng gen kiểu dại ("WT") trong thể đơn bội có số lượng bản sao của "2" và thể đột biến sẽ là "0" vì tên gọi bản sao tương quan với số lượng bản sao gen ADH nội sinh. Do đó, dòng đơn bội mang các gen GW2-01 hoặc GW2-02 chưa chỉnh sửa WT sẽ có tên gọi bản sao bằng "2." Locut thể kích tạo đơn bội WT sẽ có tên gọi bản sao bằng "2" đối với thử nghiệm #2826 và "0" đối với thử nghiệm #2827 (biến thể thể kích tạo đơn bội). Nếu dòng cây ngô là thể lưỡng bội giữa ngô ngọt và thể kích tạo chuyển gen, nó sẽ là dị hợp tử đối với gen thể kích tạo đơn bội và do đó có tên gọi bản sao bằng "1" đối với cả thử nghiệm #2826 và thử nghiệm #2827.

Bảng 5. Phân tích tình trạng tiếp hợp con cháu từ phép lai. Kết quả phân tích Taqman cho thấy các dòng không chứa gen chuyển hoặc locut thể kích tạo đơn bội từ thể cho phấn, nhưng có sự chỉnh sửa ở các đích GW2-01 và/hoặc GW2-02.

	Alen:	cCas9-01	cPMI-09	Đích CRISPR ở GW2-01 (23399)	Đích CRISPR ở GW2-02 (23399)	Alen pPLAIIa WT	Alen RWK (Thể kích tạo đơn bội) của pPLAIIa
	ID Thử nghiệm:	2540	1750	3065	3095	2826	2827
ID cây trồng	ID Cấu trúc	Mức # bản sao	Mức # bản sao	Mức # bản sao	Mức # bản sao	Mức # bản sao	Mức # bản sao
đối chứng 1 bản sao		+	1	Không phát hiện	Không phát hiện	1	1
đối chứng kiểu dại		0	0	2	2	2	0
JSER82A056	23399	0	0	0	0	2	0
JSER82A063	23399	0	0	1 hoặc 2	0	2	0
JSER85A021	23399	0	0	0	>2	2	0
JSER85A022	23399	0	0	0	>2	2	0
JSER85A024	23399	0	0	0	>2	2	0
JSER85A027	23399	0	0	0	>2	2	0

JSER85A037	23399	0	0	0	>2	2	0
JSER85A039	23399	0	0	0	>2	2	0
JSER85A044	23399	0	0	0	>2	2	0
JSER85A055	23399	0	0	0	>2	2	0

Để xác nhận thêm sự chỉnh sửa đặc hiệu đích ở các dòng đơn bội này, vùng đích GW2-02 được khuếch đại từ JSER82A063 bằng PCR và sản phẩm PCR được xác định trình tự. Bazơ đơn lẻ C được loại bỏ trong JSER82A063 khi so với trình tự WT chính xác ở vị trí phân cắt Cas9 (Hình 13). Kết quả này chứng minh rõ ràng rằng bộ máy chỉnh sửa mà được đưa vào trong tế bào giao tử cái từ thể giao tử đực có thể chỉnh sửa hệ gen thể cái trước khi hệ gen đực bị loại bỏ sau khi thụ tinh kép để tạo thành phôi đơn bội. Các dòng đơn bội đã chỉnh sửa ứng viên không có gen chuyển được xử lý bằng cách phun colchixin 0,125% trong DMSO 0,5% hoặc làm ướt cây con trong dung dịch colchixin 0,06% (Eder and Chalyk, 2002, In vivo haploid induction in maize. Theor. Appl. Genetics 104:703-708). Các dòng đã xử lý được trồng trong đất và sinh trưởng trong nhà kính để sản xuất hạt con cháu.

VI. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa ở cây lúa mỳ và các cây một lá mầm khác thông qua lai xa.

Sự kích tạo đơn bội cũng đạt được bằng cách sử dụng phép lai xa khác loài hoặc khác chi (Kasha and Kao, 1970, High frequency haploid production in barley (*Hordeum vulgare L.).* Nature 225:874-886). Ví dụ, thể đơn bội lúa mỳ có thể thu được bằng cách thụ phấn bằng các phép lai khác chi khác nhau với cây ngô (Suenaga and Nakajima 1989), cây kê ngọc trai (Inagaki and Mujeeb-Kazi 1995), cỏ teosinte (Ushiyama et al. 1991), *H. bulbosum* (Barclay 1975), và cây lúa miến (Ohkawa et al. 1992). Thể đơn bội cây lúa mạch thu được bằng cách thụ phấn bằng phấn *Hordeum bulbosum*. Thể đơn bội cây thuốc lá có thể thu được bằng cách lai với phấn *N. africana*. Nhiều ví dụ khác tồn tại ở các cây trồng khác.

Tương tự như các ví dụ nêu trên trong việc đưa locut chỉnh sửa chuyển gen vào dòng cảm ứng Stock6, locut chỉnh sửa chuyển gen có thể được đưa vào các dòng này dùng cho phép lai xa để gây ra sự kích tạo đơn bội và trình tự đột biến được nhắm đích. Dòng chuyển gen biểu hiện bộ máy chỉnh sửa có thể được tạo ra ở dòng bất kỳ của cây ngô, cây lúa mỳ, cây lúa mạch, cây lúa mạch đen, cây kê ngọc trai, cây lúa, cây thuộc họ cải bắp, cây rau diếp, cây cà chua, hoặc cây trồng khác bất kỳ bằng sự biến nạp trực tiếp hoặc lai xa. Tốt hơn là locut chuyển gen được tạo ra là đồng hợp tử và sau đó dòng này được sử dụng làm thể cho phấn trong phép lai xa với các cây trồng nhận tương thích khác để tạo ra thể đơn bội để tạo ra sự chỉnh sửa mong muốn. Quy trình loại bỏ hệ gen sau thụ tinh trong phép lai xa về cơ bản giống như quy trình trong hệ thống đột biến *MATL* ở cây ngô, mặc dù trong một số trường hợp ADN và bộ máy chỉnh sửa có nguồn gốc từ phấn lạ có thể bị loại bỏ sớm hơn một chút trong sự phát triển phôi, đó là lý do tại sao phương pháp này tốt hơn là được thực hành bằng cách sử dụng vùng khởi động mà gây ra sự biểu hiện của bộ máy chỉnh sửa trong phấn, tế bào giao tử đực, và/hoặc tế bào hợp tử, sao cho ARN và protein chỉnh sửa có mặt và có thể chỉnh sửa hệ gen đích dù ADN thể đực bị loại bỏ khá nhanh sau khi thụ tinh.

Để chứng minh tính khả thi của sự kích tạo đơn bội đồng thời và sự chỉnh sửa thông qua phép lai xa, dòng chuyển gen ngô biểu hiện Cas9 và sgRNA nhắm đích trình tự gen VLHP ở cây lúa mỳ được tạo ra. Véc-tơ 23763 (SEQ ID NO: 24) chứa các cát xét biểu hiện đối với Cas9 và sgRNA chứa trình tự tiền đoạn đệm xTaVLHP1 (5'-GACGAGCAGGCGCAGTTCC-3', SEQ ID NO: 25) để dẫn sự phân cắt qua trung gian Cas9 của các vị trí đích TaVLHP1 ở cây lúa mỳ. Hệ gen cây lúa mỳ có tổng số ba đích xTaVLHP1 (TaVLHP1-4A, TaVLHP1-4B và TaVLHP1-4D), với mỗi một đích trong ba hệ gen con của nó. Trình tự dẫn trong 23397 (SEQ ID NO: 20), xZmVLHP (5'-GCTGGAGCTGAGCTTCCGGG-3', SEQ ID NO: 21) cũng sẽ định hướng sự phân cắt của trình tự đích VLHP ở cây lúa mỳ, xTaVLHP2-1A (5'-GCTGGAGCTGAGCTTCCGGG-3', SEQ ID NO: 26) hoặc xTaVLHP2-1B (5'-TCTGGAGCTGAGCTTCCGGG -3', SEQ ID NO: 27). Có ba gen VLHP2A chứa xTaVLHP2-1A và 3 gen VLHP2B chứa trình tự xTaVLHP2-1B trong hệ gen cây lúa mỳ Mùa Xuân Trung Quốc. Các véc-tơ 23397 và 23763 được biến nạp vào dòng lai cùng dòng ngô NP2222 bằng cách sử dụng sự biến nạp qua trung gian *Agrobacterium*

để tạo ra dòng chuyển gen biểu hiện Cas9 và sgRNA. Dòng ngô chuyển gen được trồng trong nhà kính và tự thụ phấn để tạo ra cây T1.

Phấn được thu gom từ cây ngô chuyển gen T0 hoặc cây T1 con mang T-ADN của các véc-tơ 23397 hoặc 23763 được sử dụng để thụ phấn dòng lúa mỳ mùa xuân đã khử đực AC-Nanda. Vào từ một đến hai ngày trước khi nở hoa, các bông hoa nhỏ của cây lúa mỳ bị cắt đi và hai ngày sau được thụ phấn với phấn ngô mới mang bộ máy chỉnh sửa. Để cho thuận tiện, bông con từ dòng cây lúa mỳ bất thụ đực bào tương cao cấp Syngenta ("CMS") (16A300292) cũng được sử dụng trực tiếp làm thể cho cái để gây ra sự tạo thành phôi đơn bội với phấn ngô chuyển gen biểu hiện 23397 hoặc 23763 Cas9-sgRNA. Phôi được chiết xuất từ hoa nhỏ đã thụ phấn vào 14-20 ngày sau thụ phấn để cứu phôi để thu hồi cây con đơn bội từ hệ thống kích tạo đơn bội cây lúa mỳ x cây ngô. Phôi đã cắt ra được nuôi cấy trên MS độ mạnh đầy đủ (Murashige and Skoog 1962) hoặc ½ MS hoặc môi trường cơ bản B5 chứa các cải biến khác nhau của chất bổ sung hữu cơ và sinh trưởng in vitro trong thời gian 3–5 tuần ở nhiệt độ 20–25°C và độ dài ngày 16 giờ.

Ví dụ, phấn của cây con T1 từ dòng cây ngô chuyển gen MZET164902A044A chứa véc-tơ 23763 được sử dụng để thụ phấn cụm hoa của dòng CMS 16A300292 để tạo ra thể đơn bội lúa mỳ. Phôi đơn bội được cứu và cây con đơn bội lúa mỳ thu được được lấy mẫu để phân tích qPCR để xác định số lượng bản sao của trình tự đích VLHP (xem bảng 6). Một trong các dòng đơn bội (JSWER30A22) được phát hiện là chứa đột biến ở gen TaVLHP1-4B, nhưng không có ở các dạng đồng đẳng của nó TaVLHP1-4A và TaVLHP1-4D trong các hệ gen con A và D. Phân tích mức đa bội thể xác nhận rằng JSWER30A22 là thể đơn bội thực (xem hình 14 và hình 15). Đột biến trong vùng đích TaVLHP1-4B được xác định đặc điểm thêm bằng cách xác định trình tự và phát hiện ra là chứa đột biến mất 97 bp bắt đầu từ vị trí phân cắt Cas9 dự đoán (Hình 16). Chúng tôi cũng đã xác định dòng khác JSW16A07 với "0" bản sao ở gen TaVLHP1-4A (thử nghiệm #3252), gợi ý sự chỉnh sửa được nhắm đích trong trình tự đích. Tuy nhiên, đột biến mất trong gen đích này có khả năng là khá lớn trong việc làm mất (các) vị trí liên kết đoạn mồi vì chúng tôi không thể thu hồi được sản phẩm PCR để xác định trình tự. Cây con đơn bội với vị trí đích đã chỉnh sửa được trồng lại ra đất sau 3–5 tuần nuôi cấy in vitro. Cây con đã được trồng lại được để cho cứng vững trong một tuần trong buồng

sinh trưởng trong cùng điều kiện môi trường. Colchixin được bổ sung sau khi hình thành chồi. Tuy nhiên, sự xử lý nhân đôi nhiễm sắc thể có thể được thực hiện sớm hơn ở giai đoạn nuôi cấy in vitro cứu phôi hoặc muộn hơn sau khi trồng lại. Khi toàn bộ cây lúa mỳ con được xử lý để nhân đôi, rễ của cây con đơn bội được xén đi để lại khu vực 2–3 cm và sau đó nhúng chìm trong dung dịch colchixin 0,1% với đimetyl sulfoxit (DMSO) 2% và Tween-20 xấp xỉ 0,05% ở nhiệt độ 20°C trong thời gian 5 giờ. Sau việc xử lý này, rễ được rửa để loại bỏ colchixin dư và trồng trong chậu đất bùn. Mẫu mô cây trồng có thể được lấy từ cây con đơn bội để phát hiện đột biến để xác định cây trồng chứa các đột biến ở trình tự gen đích TaVLHP nhưng với nhiễm sắc thể ngô bao gồm trình tự mã hóa cho bộ máy chỉnh sửa chuyển gen được loại bỏ hoàn toàn. Vì JSWER30A22 là từ dòng CMS, cây trồng được thụ phấn với cây phục hồi để tạo ra hạt con cháu.

Bảng 6. Phân tích Taqman đối với con cháu cây lúa mỳ từ phép lai xa. Dòng JSW30A22 được chỉnh sửa.

	Alen:	TAV_4A	TAV_4B	TAV_4D	PMI	CAS9
	ID Thử nghiệm:	3252	3253	3254	1750	2540
ID Cây trồng	ID Cấu trúc	Mức # bản sao	Mức # bản sao	Mức # bản sao	Mức # bản sao	Mức # bản sao
WT, AC-Nanda	Không có	>2	2	>2	0	0
WT, AC-Nanda	Không có	2	2	2	0	0
WT, CMS	Không có	2	2	2	0	0
WT, CMS	Không có	2	2	2	0	0
JSW29A01	23763	2	2	2	0	0

JSW29A02	23763	2	2	2	0	0
JSW29A03	23763	2	2	2	0	0
JSW29A04	23763	2	2	2	0	0
JSW29A05	23763	1 hoặc 2	2	2	0	0
JSW29A06	23763	2	2	2	0	0
JSW29A07	23763	2	2	2	0	0
JSW29A08	23763	2	2	2	0	0
JSW29A09	23763	2	2	2	0	0
JSW29A10	23763	2	2	2	0	0
JSW29A11	23763	2	2	2	0	0
JSW29A12	23763	2	2	2	0	0
JSW29A13	23763	2	2	2	0	0
JSW29A14	23763	2	2	2	0	0
JSW29A15	23763	2	2	2	0	0
JSW29A16	23763	2	2	2	0	0
JSW30A01	23763	2	2	2	0	0
JSW30A02	23763	2	1 hoặc 2	2	0	0
JSW30A03	23763	2	2	2	0	0
JSW30A04	23763	2	2	2	0	0

JSW30A05	23763	2	2	2	0	0
JSW30A06	23763	2	2	2	0	0
JSW30A07	23763	2	2	2	0	0
JSW30A08	23763	2	2	2	0	0
JSW30A09	23763	2	2	2	0	0
JSW30A10	23763	2	2	2	0	0
JSW30A11	23763	2	2	2	0	0
JSW30A12	23763	2	2	2	0	0
JSW30A13	23763	2	2	2	0	0
JSW30A14	23763	2	2	2	0	0
JSW30A15	23763	2	2	2	0	0
JSW30A16	23763	2	2	2	0	0
JSW30A17	23763	2	2	2	0	0
JSW30A18	23763	2	2	2	0	0
JSW30A19	23763	>2	2	2	0	0
JSW30A20	23763	2	2	2	0	0
JSW30A21	23763	2	2	2	0	0
JSW30A22	23763	2	0	2	0	0
JSW30A23	23763	2	2	1 hoặc 2	0	0

JSW30A24	23763	2	2	2	0	0
JSW30A25	23763	2	2	2	0	0
JSW30A26	23763	2	2	2	0	0
JSW30A27	23763	2	2	2	0	0
JSW30A28	23763	2	2	2	0	0
JSW30A29	23763	2	2	2	0	0
JSW30A30	23763	2	1 hoặc 2	1 hoặc 2	0	0
JSW30A31	23763	2	2	2	0	0

Để chứng minh tính khả thi của sự kích tạo đơn bội đồng thời và sự chỉnh sửa thông qua phép lai xa, dòng chuyển gen ngô biểu hiện Cas9 từ năm vùng khởi động mà có sự biểu hiện cao và/hoặc đặc hiệu ở phấn, cùng với sgRNA nhắm đích trình tự gen VLHP ở cây lúa mỳ được tạo ra. Năm véc-tơ này là 24038 (SEQ ID NO: 34), 24039 (SEQ ID NO: 35), 24079 (SEQ ID NO: 36), 24091 (SEQ ID NO: 37), và 24094 (SEQ ID NO: 38). Cả năm trong số các véc-tơ này sử dụng cùng sgRNA chứa trình tự tiền đoạn đệm xTaVLHP2 (5'- GCTGGAGCTGAGCTTCCGGG -3', SEQ ID NO: 21) để dẫn sự phân cắt qua trung gian Cas9 của các vị trí đích TaVLHP2 ở cây lúa mỳ. Hệ gen cây lúa mỳ có tổng số ba đích xTaVLHP2 (TaVLHP2-2A, TaVLHP2-2B và TaVLHP2-2D), với mỗi đích trong ba hệ gen con của nó. Trình tự dẫn trong năm cấu trúc này cũng định hướng sự phân cắt của trình tự đích VLHP ở cây lúa mỳ, xTaVLHP2 (5'-GCTGGAGCTGAGCTTCCGGG-3', SEQ ID NO: 26) hoặc xTaVLHP3 (5'-TCTGGAGCTGAGCTTCCGGG -3', SEQ ID NO: 27). Có ba gen TaVLHP2 chứa các gen xTaVLHP2 và 3 TaVLHP3 chứa trình tự xTaVLHP2-1B trong hệ gen cây lúa mỳ Mùa Xuân Trung Quốc.

Véc-tơ 24038 (SEQ ID NO: 34) chứa các cát xét biểu hiện đối với Cas9 chịu sự kiểm soát của vùng khởi động biểu hiện cao ưu tiên ở phấn prZmGRMZM5G876285 và vùng kết thúc tZmGRMZM5G876285 từ gen prf3 (profilin homolog3) ngô, mà có sự

biểu hiện tự nhiên cực cao ở cấp độ ARN và protein ở phấn và mà có bằng chứng về sự biểu hiện ở tế bào giao tử đực cao.

Véc-tơ 24039 (SEQ ID NO: 35) chứa các cát xét biểu hiện đối với Cas9 chịu sự kiểm soát của vùng khởi động biểu hiện cao ưu tiên ở phấn prZmGRMZM2G020852 và vùng kết thúc tZmGRMZM2G020852 từ gen EXPB2 (BETA EXPANSIN2) ngô, mà có sự biểu hiện tự nhiên cực cao ở cấp độ ARN và protein ở phấn và có bằng chứng của sự biểu hiện ở tế bào giao tử đực.

Véc-tơ 24079 (SEQ ID NO: 36) chứa các cát xét biểu hiện đối với Cas9 chịu sự kiểm soát của vùng khởi động biểu hiện cao ưu tiên ở phấn prZmGRMZM2G146551 và vùng kết thúc tZmGRMZM2G146551 từ gen EXPB1 (BETA EXPANSIN1) ngô, mà có sự biểu hiện tự nhiên cực cao ở cấp độ ARN và protein ở phấn và có bằng chứng của sự biểu hiện ở tế bào giao tử đực.

Véc-tơ 24091 (SEQ ID NO: 37) chứa các cát xét biểu hiện đối với Cas9 chịu sự kiểm soát của vùng khởi động ưu tiên ở phấn prZmGRMZM2G471240 và vùng kết thúc tZmGMRMZM2G471240 từ gen *MATL* (*MATRILINEAL*) ngô, mà thể hiện bằng chứng về sự biểu hiện ở phấn và có thể có ở tế bào giao tử đực ở cấp độ ARN và protein.

Véc-tơ 24094 (SEQ ID NO: 38) chứa các cát xét biểu hiện đối với Cas9 chịu sự kiểm soát của vùng khởi động ưu tiên ở phấn prZmGRMZM2G471240 và vùng kết thúc tZmGMRMZM2G471240 từ gen *MATL* (*MATRILINEAL*) ngô, mà thể hiện bằng chứng về sự biểu hiện ở phấn và có thể có ở tế bào giao tử đực ở cấp độ ARN và protein. Cấu trúc này có thêm sự dung hợp đầu tận cùng N của protein huỳnh quang AmCyan trên phân tử Cas9 để chụp ảnh và trực quan hóa sự định vị Cas9 trong phấn.

Năm véc-tơ này (24038, 24039, 24079, 24091, và 24094) được biến nạp vào dòng lai cùng dòng ngô NP2222 bằng cách sử dụng sự biến nạp qua trung gian *Agrobacterium* để tạo ra dòng chuyển gen biểu hiện Cas9 và sgRNA.

Dòng ngô chuyển gen được trồng trong nhà kính và các cây trồng chuyển gen một bản sao và hai bản sao được lai xa với cây lúa mỳ mùa xuân và dòng cây lúa mỳ CMS. Phấn được thu gom từ cây T0 ngô chuyển gen mang T-ADN của một trong số

các véc-tơ 24038, 24039, 24079, 24091, và 24094 được sử dụng để thụ phấn dòng lúa mỳ mùa xuân đã khử đực AC-Nanda. Phấn cũng được dùng cho thí nghiệm qRT, trong đó sự biểu hiện của Cas9 được đo ở cấp độ ARN và so với sự biểu hiện Cas9 trong mẫu lá khi Cas9 được điều khiển bởi vùng khởi động ubiquitin cây mía dùng trong nhiều ví dụ về ngô và lúa mỳ nêu trên. Như bạn có thể thấy trên hình 17, sự biểu hiện cao ở phấn, trung bình cao hơn khoảng 100 lần ở cây trồng mang các véc-tơ T-ADN 24038, 24039, và 24079 so với vùng khởi động ubiquitin. Sự biểu hiện cũng cao hơn ở phấn từ cây trồng chứa véc-tơ 24038, 24039, và 24079 khi so với phấn mang các cấu trúc vùng khởi động *MATRILINEAL* (24094 và 24091), mà đã biết là có sự biểu hiện gen tự nhiên thấp hơn. Cả năm trong số các vùng khởi động này có kiểu mẫu biểu hiện mà bị hạn chế ở phấn. Dấu hiệu của việc vùng khởi động đang hoạt động một cách phù hợp là, chúng tôi quan sát thấy không có sự biểu hiện T0 của Cas9 ở lá cây con thể chai, và không có sự chỉnh sửa của các vị trí đích VLHP ở lá cây ngô T0 (không muốn bị ràng buộc bởi lý thuyết, sự chỉnh sửa có thể xảy ra ở các vị trí đích ở cây ngô, trong tất cả các khả năng có thể xảy ra, ở giai đoạn phấn chín, khi Cas9 được biểu hiện trong lần thứ nhất).

Vào lúc một đến hai ngày trước khi nở hoa, các bông hoa nhỏ của cây lúa mỳ bị cắt đi từ dòng CMS và dòng AC Nanda. Hai ngày sau các bông hoa nhỏ được thụ phấn với phấn ngô mới mang bộ máy chỉnh sửa, Cas9-sgRNA, từ một trong các cấu trúc 24038, 24039, 24091, hoặc 24094 (cây T0 đã biến nạp với cấu trúc 24079 được làm chậm, và không lai với cây lúa mỳ theo phương thức này). Phôi lúa mỳ được chiết xuất từ hoa nhỏ đã thụ phấn vào 14-20 ngày sau thụ phấn để cứu phôi để thu hồi cây con đơn bội từ hệ thống kích tạo đơn bội cây lúa mỳ x cây ngô. Phôi đã cắt ra được nuôi cấy trên MS độ mạnh đầy đủ (Murashige and Skoog 1962) hoặc ½ MS hoặc môi trường cơ bản B5 chứa các cải biến khác nhau của chất bổ sung hữu cơ và sinh trưởng in vitro trong thời gian 1-5 tuần ở nhiệt độ 20-25°C và độ dài ngày 16 giờ. Ví dụ, phấn của cây con T0 từ dòng cây ngô chuyển gen MZKE172601A100A chứa véc-tơ 24039 được sử dụng để thụ phấn cụm hoa của dòng CMS 16A300292 để tạo ra thể đơn bội lúa mỳ. Phôi đơn bội được cứu và cây con đơn bội lúa mỳ thu được được lấy mẫu để phân tích qPCR để xác định số lượng bản sao của trình tự đích VLHP (Bảng 7). Trong phân tích này, chúng tôi thử nghiệm về gen chuyển Cas9 bằng cách sử dụng thử nghiệm #2540. Tất cả các phôi lúa mỳ được cứu và được thử nghiệm thiếu gen chuyển này và mang lại điểm bằng "0" đối với Cas9, vì chúng không có bất kỳ ADN ngô nào trong phôi đang phát triển và

do đó không có gen chuyển. ADN ngô bị loại bỏ hoàn toàn, bị đánh bật ra hoặc không được chuyển đầy đủ ở nơi đầu tiên trong quy trình kích tạo đơn bội, diễn ra trong và/hoặc sau khi thụ tinh). Ngoài Cas9, chúng tôi kiểm tra đối với các thử nghiệm #3332 và #3333, mà tạo ra sự khuếch đại không đặc hiệu của cả hai alen VLHP2-2A và -2D. Các thử nghiệm này thường đọc dưới dạng "2" hoặc ">2" ở cây lúa mỳ đơn bội, và phần lớn thể đơn bội chúng tôi đã tạo ra bằng cách sử dụng phấn ngô chuyển gen có điểm 2 hoặc >2 đối với các thử nghiệm này. Chúng tôi sử dụng các thử nghiệm này để tìm kiếm thể đơn bội đã chỉnh sửa giả định, bằng cách tìm kiếm điểm bằng 0 hoặc 1. Tên gọi là "1" có thể chỉ ra rằng một trong hai alen, VLHP2-2A, hoặc -2D, đã được chỉnh sửa. Cuối cùng, chúng tôi kiểm tra đối với thử nghiệm 3255 ở thể đơn bội AC Nanda, mà phát hiện ra VLHP2-2B một cách đặc hiệu. Dòng CMS không khuếch đại thử nghiệm này, ngay cả khi nó là kiểu dại, do vậy chúng tôi không sử dụng nó cho thể đơn bội CMS. Thể đơn bội chưa chỉnh sửa cho điểm bằng "2," trong khi thể đơn bội đã chỉnh sửa giả định được phát hiện vì chúng có điểm bằng "0." Điểm bằng "1" có thể chỉ ra sự đọc sai hoặc mẫu đã được chỉnh sửa một phần, khảm.

Một ví dụ là, một trong số các cây đơn bội AC Nanda 440-A5 được phát hiện là chứa đột biến ở gen TaVLHP2-2B, nhưng không ở các dạng đồng đẳng của nó TaVLHP2-2A và TaVLHP2-2D trong các hệ gen con A và D (bảng 7). Dữ liệu Taqman cũng cho thấy rằng nó thiếu gen chuyển Cas9. Đột biến trong vùng đích TaVLHP2-2B được xác định đặc điểm thêm bằng cách xác định trình tự, nhưng mặc dù chúng tôi có thể khuếch đại các alen A và D, chúng tôi không còn có thể khuếch đại alen B, gợi ý rằng có mặt sự chỉnh sửa lớn hơn, dường như là đột biến mất lớn, mà dẫn đến sản phẩm PCR không còn khuếch đại nữa.

Một ví dụ khác là, một trong số các cây đơn bội CMS 450-D11 được phát hiện là chứa đột biến ở thể tương đồng TaVLHP2-2D hoặc -2A, theo điểm bằng "1" đối với cả hai thử nghiệm 3332 và 3333 (bảng 7). Dữ liệu taqman cho thấy rằng nó thiếu gen chuyển Cas9. Các vùng đích TaVLHP2-2A, 2B và 2D được xác định đặc điểm thêm bằng cách xác định trình tự, nhưng mặc dù chúng tôi có thể khuếch đại các alen A và B, chúng tôi không thể khuếch đại alen D nữa, gợi ý rằng có mặt sự chỉnh sửa lớn hơn mà dẫn đến thất bại PCR.

Xem xét thể đơn bội lúa mỳ 2295 được tạo ra từ phép lai với phấn ngô mang một trong số năm cấu trúc biểu hiện ưu tiên ở phấn sau đây (24038, 24039, 24091, và 24094), chúng tôi phát hiện thấy 15 thể đơn bội mà cho dữ liệu thử nghiệm Taqman mà chỉ ra sự chỉnh sửa có thể có ở một trong số các vị trí đích VLHP2-2A, VLHP2-2D, hoặc VLHP2-2B. Sau khi xác định trình tự, bảy trong số các thể đơn bội này được phát hiện là có trình tự kiểu dại ở các vị trí đích, và được gọi là dương tính giả do lỗi Taqman. Các lỗi này được cho là do thực tế rằng các thử nghiệm #3332 và #3333 mang lại sự khuếch đại không đặc hiệu của cả hai alen VLHP-2A và -2D, dẫn đến một số tên gọi bị bỏ sót, hoặc do chất lượng ADN thấp.

Trong số 8 thể đơn bội đã chỉnh sửa giả định còn lại, sáu là AC Nanda (440-B3, 440-D3, 440-A5, 447-G8, 456-G9, 459-A2) trong đó gen chuyển chỉnh sửa là từ cấu trúc 24038. Bốn trong số đó (440-B3, 440-D3, 440-A5, và 456-G9) chứa sự chỉnh sửa ở VLHP2-2B. Chúng được phát hiện thấy vì chúng có điểm Taqman bằng "0" đối với thử nghiệm 3255. Các cây trồng này thiếu Cas9 (điểm bằng "0") nhưng có điểm "2" kiểu dại đối với VLHP2-2A hoặc VLHP2-2D (thử nghiệm #3332 và #3333) chỉ ra rằng chúng không được chỉnh sửa các vị trí đó. Sáu cây trồng này được xác nhận là thể đơn bội bằng phân tích mức bội thể. Chúng tôi cố gắng xác định trình tự alen đã được chỉnh sửa, nhưng mặc dù các phản ứng PCR và xác định trình tự hoạt động tốt đối với 2A và 2D, chúng tôi không thể thu được sản phẩm PCR đối với 2B. Chúng tôi lặp lại PCR một vài lần bằng cách sử dụng khoảng giá trị của điều kiện phản ứng nhưng không thể khuếch đại các chất tương đồng 2B từ các cây đơn bội này. Điều này có thể chỉ ra rằng việc chỉnh sửa gây ra sự thay đổi lớn trong gen 2B ở các cây trồng này mà có thể gây ra việc làm mất vị trí ủ đoạn mồi. Chúng tôi kỳ vọng nhiều cây trồng CMS cũng có sự chỉnh sửa ở vị trí đích VLHP2-2B, nhưng chúng tôi không có thử nghiệm để phát hiện alen VLHP2-2B từ dòng CMS.

Xem xét một mình AC Nanda, chúng tôi tính tỉ lệ chỉnh sửa tổng thể tại alen đó bằng 0,7% đối với tất cả các cấu trúc, nhưng tỉ lệ chỉnh sửa đặc biệt cao bằng 1,4% đối với cấu trúc 24038.

Ngoài bốn thể đơn bội đã chỉnh sửa này với điểm bằng "0" đối với 3255, một vài cây trồng khác mang lại điểm bằng "0 hoặc 1" hoặc "1" đối với 3255, mà chỉ ra hiện

tượng khảm có thể có (sự chỉnh sửa một phần ở các dòng tế bào nhất định của phôi hoặc cây con), nhưng chúng tôi không tiếp tục trên các cây trồng này. Đối với thể tương đồng AC Nanda VLHP2-2A, cây trồng 447-G8 chứa sự chỉnh sửa mà chúng tôi cũng không thể xác định trình tự vì phản ứng PCR thất bại, mặc dù vậy 2B và 2D đã khuếch đại và chứa trình tự kiểu dại. Chúng tôi lặp lại PCR một vài lần bằng cách sử dụng khoảng giá trị của điều kiện phản ứng nhưng không thể khuếch đại thể tương đồng 2A. Tương tự, đối với VLHP2-2D, cây trồng 459-A2 chứa sự chỉnh sửa mà chúng tôi không thể xác định trình tự vì phản ứng PCR thất bại. Chúng tôi lặp lại PCR một vài lần bằng cách sử dụng khoảng giá trị của điều kiện phản ứng nhưng không thể khuếch đại thể tương đồng 2D. Chúng tôi cũng phát hiện thấy sự chỉnh sửa giả định ở 447-H12 và 440-G6, nhưng khi xác định trình tự chúng tôi phát hiện thấy rằng chúng là dương tính giả.

Đối với thể đơn bội CMS, cây trồng 450-D11 mang lại điểm bằng "1" đối với cả thử nghiệm #3332 và 3333 (Bảng 7). Khi xác định trình tự, chúng tôi phát hiện thấy rằng thể tương đồng 2A có trình tự kiểu dại, nhưng chúng tôi không thể khuếch đại bằng PCR thể tương đồng 2D, gợi ý rằng sự chỉnh sửa lớn đã xảy ra. Chúng tôi lặp lại PCR một vài lần bằng cách sử dụng khoảng giá trị của điều kiện phản ứng nhưng không thể khuếch đại thể tương đồng 2D. Đối với cây trồng 452-B11, điểm Taqman là "0" đối với #3332 (VLHP2-2A), và chúng tôi không thể khuếch đại alen đó để xác định trình tự, mặc dù vậy sản phẩm PCR và trình tự của 2D và 2B là bình thường. Chúng tôi lặp lại PCR một vài lần bằng cách sử dụng khoảng giá trị của điều kiện phản ứng nhưng không thể khuếch đại thể tương đồng 2A. Chúng tôi cũng phát hiện thấy năm cây trồng mà có sự chỉnh sửa giả định theo dữ liệu Taqman đối với các thử nghiệm 3332 và 3333, nhưng xác định trình tự bằng PCR cho thấy chúng là dương tính giả; trình tự là kiểu dại (chưa chỉnh sửa).

Tổng cộng, chúng tôi phát hiện thấy hai thể đơn bội CMS đã chỉnh sửa và sáu thể đơn bội AC Nanda đã chỉnh sửa. Có thể có nhiều thể đơn bội đã chỉnh sửa hơn mà chúng tôi không thể phát hiện vì chúng tôi không có thử nghiệm đối với gen 2B đối với cây trồng CMS, hay là đối với các vị trí đích gen VLHP3 của ARN dẫn trong năm cấu trúc này.

Dữ liệu xác định trình tự từ thể đơn bội đã được chỉnh sửa thống nhất với khái niệm của đột biến mất, đột biến cài xen hoặc sự tái sắp xếp lớn xung quanh vị trí đích ARN dẫn, và kéo dài ra đủ xa để có thể bao gồm sự loại bỏ của một trong số các vị trí liên kết đoạn mồi. Kiểu thay đổi lớn này không phải là hiếm gặp trong quá trình chỉnh sửa bằng Cas9, đặc biệt là ở mô trong đó sự sửa chữa ADN thông qua sự kết hợp đầu không tương đồng chậm hơn hoặc bị ức chế – mà có thể là trường hợp trong hợp tử vừa mới được thụ tinh hoặc sớm ở phôi đơn bội lúa mỳ.

Bảng 7. Dữ liệu xác định trình tự từ thể đơn bội lúa mỳ đã chỉnh sửa.

CMS			TAV_2A		TAV_2D		TAV_2B		PMI		Cas9		
			3332		3333		3255		1750		2540		
ID Cây trồng	ID Cấu trúc	# Bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	Kết quả xác định trình tự
427-A2	WT	Không có	2,44	>2	2,38	2	Không được thử nghiệm		0,00	0	0,00	0	không được xác định trình tự
427-B2	WT	Không có	1,99	2	1,99	2	Không được thử nghiệm		0,00	0	0,00	0	không được xác định trình tự
427-C2	WT	Không có	2,02	2	2,07	2	Không được thử nghiệm		0,00	0	0,00	0	không được xác định trình tự
427-D2	WT	Không có	2,31	2	2,16	2	Không được thử nghiệm		0,00	0	0,00	0	không được xác

											định trình tự
427-A10	24091	2	2,07	2	1,66	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-B10	24091	2	1,95	2	1,84	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-C10	24091	2	1,93	2	2,28	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-D10	24091	2	2,59	>2	2,48	>2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-E10	24091	2	1,90	2	1,78	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-F10	24091	2	2,03	2	1,96	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-G10	24091	2	2,08	2	2,25	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác

							định trình tự					
427-H10	24091	2	0,58	1	0,81	1	Không được thử nghiệm	0,00	0	0,00	0	A và D đều là WT
427-A11	24091	2	1,57	1 hoặc 2	1,93	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-B11	24091	2	1,41	1 hoặc 2	1,63	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-C11	24091	2	1,06	1	1,21	1	Không được thử nghiệm	0,01	0	0,01	0	không được xác định trình tự
427-D11	24091	2	1,98	2	2,03	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-E11	24091	2	1,94	2	1,94	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-F11	24091	2	1,84	2	1,84	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự

427-G11	24091	2	1,54	1 hoặc 2	1,76	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-H11	24091	2	1,75	2	1,76	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-A12	24091	2	1,99	2	2,15	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-B12	24091	2	0,72	1	1,26	1	Không được thử nghiệm	0,00	0	0,00	0	A và D đều là WT
427-C12	24091	2	1,69	2	1,50	1 hoặc 2	Không được thử nghiệm	0,00	0	0,01	0	không được xác định trình tự
427-D12	24091	1	2,34	2	2,03	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-E12	24091	1	1,98	2	2,04	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-F12	24091	1	1,89	2	1,97	2	Không được thử nghiệm	0,00	0	0,00	0	không được

											xác định trình tự
427-G12	24091	1	1,56	1 hoặc 2	1,77	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
427-H12	24091	1	1,57	1 hoặc 2	2,36	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-A3	24091	1	2,12	2	1,75	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-B3	24091	1	2,69	>2	1,89	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-C3	24091	1	2,09	2	2,44	>2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-D3	24091	1	2,05	2	2,39	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-E3	24091	1	2,48	>2	2,87	>2	Không được thử nghiệm	0,00	0	0,00	0	không được xác

								định trình tự				
428-F3	24091	1	2,33	2	2,76	>2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
428-G3	24091	1	2,84	>2	0,22	0	Không được thử nghiệm	0,00	0	0,00	0	A và D đều là WT
428-H3	24091	1	2,83	>2	2,60	>2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-A11	24094	1	1,97	2	2,24	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-B11	24094	1	2,13	2	2,04	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-C11	24094	1	2,15	2	2,18	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-D11	24094	1	1,04	1	0,99	1	Không được thử nghiệm	0,00	0	0,00	0	A & B là WT; D thất bại

450-E11	24094	1	2,35	2	2,01	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-F11	24094	1	2,02	2	1,90	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-G11	24039	1	1,76	2	1,72	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
450-H11	24039	1	2,07	2	2,04	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
452-H4	24038	2	2,62	>2	0,01	0	Không được thử nghiệm	0,00	0	0,00	0	A và D đều là WT
452-A11	24038	2	2,24	2	2,28	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
452-B11	24038	2	0,00	0	2,22	2	Không được thử nghiệm	0,00	0	0,00	0	B & D là WT; A thất bại
452-C11	24038	2	2,55	>2	2,22	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác

												định trình tự
452-D11	24038	2	0,82	1	1,26	1	Không được thử nghiệm	0,00	0	0,00	0	A và D đều là WT
452-E11	24038	2	2,43	>2	2,36	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
452-F11	24038	2	2,12	2	2,21	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
452-G11	24038	2	2,38	2	1,99	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự
452-H11	24038	2	1,82	2	1,83	2	Không được thử nghiệm	0,00	0	0,00	0	không được xác định trình tự

NANDA			TAV_2A		TAV_2D		TAV_2B		PMI		Cas9		
			3332		3333		3255		1750		2540		
ID Cây trồng	ID Cấu trúc	# Bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	Kết quả xác định trình tự

425-A2	WT	Không có	2,30	2	2,62	>2	1,908	2		0,00	0	0,00	0	không được xác định trình tự
425-B2	WT	Không có	2,28	2	2,41	>2	2,274	2		0,00	0	0,00	0	không được xác định trình tự
425-C2	WT	Không có	2,47	>2	1,92	2	1,962	2		0,00	0	0,00	0	không được xác định trình tự
425-D2	WT	Không có	2,10	2	2,11	2	1,772	2		0,00	0	0,00	0	không được xác định trình tự
447-A12	24038	2	1,72	2	1,90	2	2,02	2		0,00	0	0,00	0	không được xác định trình tự
447-B12	24039	2	2,18	2	1,62	2	1,47	1	hoặc 2	0,00	0	0,00	0	không được xác định trình tự
447-C12	24039	2	1,78	2	2,40	2	1,90	2		0,00	0	0,00	0	không được xác định trình tự

447-D12	24039	2	1,58	1 hoặc 2	1,70	2	2,18	2	0,00	0	0,00	0	không được xác định trình tự
447-E12	24039	2	2,13	2	1,82	2	2,14	2	0,00	0	0,00	0	không được xác định trình tự
447-F12	24039	2	2,25	2	1,78	2	2,15	2	0,00	0	0,00	0	không được xác định trình tự
447-G12	24039	2	1,90	2	2,30	2	2,23	2	0,00	0	0,00	0	không được xác định trình tự
447-H12	24039	1	2,34	2	1,95	2	0,89	1	0,00	0	0,00	0	A, B, và D đều là WT
440-A2	24039	1	1,72	2	1,71	2	1,24	1	0,00	0	0,00	0	không được xác định trình tự
440-B2	24039	1	2,30	2	2,56	>2	1,77	2	0,00	0	0,00	0	không được xác định trình tự

440-C2	24039	1	3,05	>2	1,85	2	2,01	2	0,00	0	0,00	0	không được xác định trình tự
440-D2	24039	1	1,66	2	1,70	2	1,44	1 hoặc 2	0,00	0	0,00	0	không được xác định trình tự
440-E2	24039	1	2,23	2	1,91	2	2,04	2	0,00	0	0,00	0	không được xác định trình tự
440-F2	24039	1	2,63	>2	2,07	2	2,28	2	0,00	0	0,00	0	không được xác định trình tự
440-G2	24038	11	1,91	2	1,87	2	2,10	2	0,00	0	0,00	0	không được xác định trình tự
440-H2	24038	1	1,85	2	1,80	2	1,97	2	0,00	0	0,00	0	không được xác định trình tự
440-A3	24038	1	2,52	>2	2,05	2	1,99	2	0,00	0	0,00	0	không được xác định trình tự

440-B3	24038	1	2,16	2	2,19	2	0,00	0	0,00	0	0,00	0	A & D là WT; B thất bại
440-C3	24038	1	2,58	>2	2,02	2	2,78	>2	0,00	0	0,00	0	không được xác định trình tự
440-D3	24038	1	2,34	2	2,32	2	0,00	0	0,00	0	0,00	0	A & D là WT; B thất bại
440-E3	24038	1	2,49	>2	2,23	2	2,47	>2	0,00	0	0,00	0	không được xác định trình tự
440-F3	24038	1	2,08	2	2,10	2	2,17	2	0,00	0	0,00	0	không được xác định trình tự
440-F4	24038	1	1,73	2	1,47	1 hoặc 2	1,41	1 hoặc 2	0,00	0	0,00	0	không được xác định trình tự
440-G4	24038	1	1,53	1 hoặc 2	2,02	2	1,99	2	0,00	0	0,00	0	không được xác định trình tự
440-H4	24038	1	2,22	2	1,90	2	0,00	0	0,00	0	0,00	0	không được

												xác định trình tự
440-A5	24038	1	2,22	2	1,90	2	0,00	0	0,00	0	0,00	0	A & D là WT; B thất bại
440-A6	24039	2	2,49	>2	2,32	2	1,84	2	0,00	0	0,00	0	không được xác định trình tự
440-B6	24039	2	2,12	2	2,03	2	2,21	2	0,00	0	0,00	0	không được xác định trình tự
440-C6	24039	2	2,63	>2	2,07	2	2,28	2	0,00	0	0,00	0	không được xác định trình tự
440-D6	24039	2	2,49	>2	2,23	2	2,47	>2	0,00	0	0,00	0	không được xác định trình tự
440-E6	24039	2	2,45	>2	2,20	2	2,32	2	0,00	0	0,00	0	không được xác định trình tự
440-F6	24039	2	2,10	2	1,92	2	1,91	2	0,00	0	0,00	0	không được xác

												định trình tự
440-G6	24039	2	0,57	1	0,66	1	0,53	1	0,00	0	0,00	0	A, B & D đều là WT
440-H6	24039	2	1,81	2	1,96	2	2,51	>2	0,00	0	0,00	0	không được xác định trình tự
447-A8	24038	1	2,42	>2	2,21	2	2,10	2	0,00	0	0,00	0	không được xác định trình tự
447-B8	24038	1	2,46	>2	2,32	2	2,09	2	0,00	0	0,00	0	không được xác định trình tự
447-C8	24038	1	2,09	2	2,08	2	2,29	2	0,00	0	0,00	0	không được xác định trình tự
447-D8	24038	1	2,13	2	2,14	2	2,34	2	0,00	0	0,00	0	không được xác định trình tự
447-E8	24038	11	2,36	2	2,31	2	2,44	>2	0,00	0	0,00	0	không được xác định trình tự

447-F8	24038	1	2,72	>2	2,28	2	2,00	2	0,00	0	0,00	0	không được xác định trình tự
447-G8	24038	1	0,71	1	1,34	1 hoặc 2	2,33	2	0,00	0	0,00	0	B & D là WT; A thất bại
447-H8	24038	1	2,25	2	2,29	2	2,01	2	0,00	0	0,00	0	không được xác định trình tự
456-A9	24038	2	2,19	2	1,59	1 hoặc 2	2,03	2	0,00	0	0,00	0	không được xác định trình tự
456-B9	24038	2	2,13	2	2,11	2	2,02	2	0,00	0	0,00	0	không được xác định trình tự
456-C9	24038	2	2,16	2	1,85	2	1,45	1 hoặc 2	0,00	0	0,00	0	không được xác định trình tự
456-D9	24038	2	2,56	>2	2,18	2	1,76	2	0,00	0	0,00	0	không được xác định trình tự

456-E9	24038	2	2,29	2	2,03	2	1,65	2	0,00	0	0,00	0	không được xác định trình tự
456-F9	24038	2	2,24	2	2,02	2	2,05	2	0,00	0	0,00	0	không được xác định trình tự
456-G9	24038	2	2,49	>2	2,03	2	0,00	0	0,00	0	0,00	0	A & D là WT; B thất bại
456-H9	24038	2	1,78	2	1,62	2	1,38	1 hoặc 2	0,00	0	0,00	0	không được xác định trình tự
459-A2	24038	2	1,38	1 hoặc 2	1,11	1	0,94	1	0,00	0	0,00	0	A & B là WT; D thất bại
459-B2	24038	2	1,86	2	1,91	2	2,12	2	0,00	0	0,00	0	không được xác định trình tự
459-C2	24038	2	1,94	2	2,09	2	1,42	1 hoặc 2	0,00	0	0,00	0	không được xác định trình tự
459-D2	24038	2	2,09	2	2,05	2	1,91	2	0,00	0	0,00	0	không được

													xác định trình tự
459-E2	24038	2	2,18	2	2,12	2	2,12	2	0,00	0	0,00	0	không được xác định trình tự

Tóm lại, chúng tôi phát hiện thấy rằng tần suất chỉnh sửa (số lượng thể đơn bội đã chỉnh sửa xác định được chia cho tổng số thể đơn bội) đối với cấu trúc 24038 là 0,79%; đối với cấu trúc 24039 nó là 0%; đối với cấu trúc 24091 nó là 0%, và đối với cấu trúc 24094 nó là 0,75%. Tuy nhiên, tỉ lệ chỉnh sửa này dĩ nhiên là sự đánh giá thấp vì chúng tôi không có thử nghiệm để phát hiện sự chỉnh sửa ở nhiều vị trí đích ARN dẫn. Ngoài ra, vì chúng tôi sử dụng phấn T0 mà có 1 hoặc 2 bản sao, chúng tôi biết rằng với phấn 1 bản sao, chỉ 50% hạt phấn thụ tinh chứa Cas9, và do vậy chỉ một nửa số phôi có cơ hội được chỉnh sửa; tương tự, đối với bố mẹ 2 bản sao, giả thiết sự chia tách ngẫu nhiên của gen chuyển trong giảm phân ở thể đực, chúng tôi kỳ vọng khoảng 75% phấn chứa Cas9, do vậy 25% phôi không thể được chỉnh sửa. Có lý khi kết luận rằng, khi người ta cố gắng sử dụng sự chỉnh sửa đồng thời này cộng với công nghệ kích tạo đơn bội với bộ máy chỉnh sửa mang bởi phấn, trong một số trường hợp có thể tối ưu hơn khi sử dụng vùng khởi động mà biểu hiện đặc hiệu hoặc cao ở phấn và ở tế bào giao tử đực, sao cho Cas9 có thể được biểu hiện ở mức độ cao hơn. Trong các trường hợp mà đích gen có thể ảnh hưởng đến sự phát triển của cây trồng kích tạo đơn bội, có vùng khởi động được ưu tiên ở phấn hoặc giao tử đực mà không biểu hiện ở lá có thể hữu dụng vì nó sẽ tránh được sự chỉnh sửa gen đích ở cây trồng kích tạo đơn bội trong quá trình sự phát triển, có lẽ là chỉnh sửa nó trong lần thứ nhất ở phấn.

Vì tế bào giao tử đực thụ tinh giao tử cái, chúng có khả năng để chuyển ARN và protein Cas9 (cũng như là bản thân ADN gen chuyển, tích hợp vào một trong các nhiễm sắc thể của thể đực mà sẽ bị loại bỏ). Như chúng tôi đã chứng minh trong phép lai diện rộng trong ví dụ này, nó có thể hoạt động tốt để có Cas9 và/hoặc ARN dẫn chịu sự kiểm soát của vùng khởi động mà biểu hiện đặc hiệu hoặc cao ở phấn, và cụ thể là trong tế

bào giao tử đực, khi sử dụng thể kích tạo đơn bội làm thể đực để chỉnh sửa dòng cao cấp. Chúng tôi không biết chính xác liệu *MATRILINEAL*, *EXPB1*, *EXPB2*, và *PRF3* có biểu hiện trong nhân sinh dưỡng, tế bào giao tử đực, hoặc cả hai, và liệu có thể có sự biểu hiện bất kỳ nào trong loại tế bào hợp tử hay không, nhưng chúng được chọn vì chúng được cho là được biểu hiện cao và/hoặc đặc hiệu ở phấn. Vùng khởi động *PRF3* có motif liên kết *DUO1* trong vùng khởi động, mà có thể cho thấy nó biểu hiện trong tế bào giao tử đực. Điều này thống nhất với việc vùng khởi động có tần suất chỉnh sửa cao hơn. Thực tế là chúng tôi phát hiện thấy nhiều thể đơn bội lúa mỳ đã chỉnh sửa sau khi lai xa cho thấy rõ ràng là khi có sự biểu hiện cao của Cas9 ở phấn, bằng cách sử dụng các vùng khởi động này hoặc vùng khởi động khác bất kỳ, mà sự biểu hiện có thể dẫn đến sự chỉnh sửa trong phôi lúa mỳ sau khi lai xa. Có nhiều khả năng là các vùng khởi động này, cũng như là các vùng khởi động khác mà gây ra sự biểu hiện ở phấn, hoặc cụ thể là trong tế bào giao tử đực, có thể làm tăng hiệu quả của quy trình chỉnh sửa trong quá trình kích tạo đơn bội ở ngô, hoặc kích tạo đơn bội ở lúa.

Tương tự, trong ví dụ tiếp theo dưới đây, chúng tôi thể hiện sự chỉnh sửa đơn bội ở cây hai lá mầm bằng cách sử dụng dòng kích tạo đơn bội được cải biến *CENH3*, và chúng tôi sử dụng cấu trúc vùng khởi động để định hướng Cas9. Nhưng trong nỗ lực làm tăng hiệu quả của sự chỉnh sửa đơn bội, chúng tôi có thể lựa chọn sử dụng vùng khởi động mà gây ra sự biểu hiện cao và/hoặc đặc hiệu trong tế bào giao tử cái, chẳng hạn như vùng khởi động của gen EGG APPARATUS1 ("prEA1") (xem, ví dụ, Gray-Mitsumune, M. and Matton, D.P., *The* Egg apparatus 1 *gene from maize is a member of a large gene family found in both monocots and dicots*, PLANTA 223(3):618–625 (Feb. 2006)) hoặc EGG CELL1 (EC1) (xem, ví dụ, Sprunck S, et al., *Egg cell-secreted EC1 triggers sperm cell activation during double fertilization*. Science 2012; 338:1093-97; PMID:23180860; http://dx.doi.org/ 10.1126/science.1223944).

Một ví dụ của nó là, người ta có thể sử dụng vùng khởi động được biểu hiện ở tế bào giao tử đực, chẳng hạn như vùng khởi động DUO1 đặc hiệu giao tử đực *Arabidopsis* (*xem, ví dụ*, Engel, et al., *Green Sperm. Identification of Male Gamete Promoters in Arabidopsis*, PLANT PHYSIOLOGY Aug 2005, 138 (4) 2124-2133; DOI: 10.1104/pp.104.054213), hoặc các thể tương đồng của DUO1 từ các loài khác (ví dụ như, gen ngô GRMZM2G105137 và GRMZM2G046443 đều là thể tương đồng DUO1

mà có chung kiểu biểu hiện đặc hiệu phấn tương tự). Nếu sử dụng vùng khởi động bất kỳ trong số đó để gây ra sự biểu hiện Cas9 trong tế bào giao tử đực của dòng kích tạo đơn bội như RWK, NP2222-HI, hoặc thể đột biến *matl*, có thể tạo ra dòng thể chỉnh sửa đơn bội hiệu quả cao để sử dụng trong việc chỉnh sửa nhiều loại chất mầm nguyên sinh cây ngô hoặc cây lúa mỳ cao cấp, lần lượt thông qua lai cùng loài hoặc lai xa.

Vùng khởi động được biểu hiện ở giao tử đực thích hợp khác cho ý tưởng gây ra sự biểu hiện cao Cas9 trong tế bào giao tử đực sẽ bao gồm thể tương đồng DUO1 ở cây lúa mỳ, cây lúa, cây lúa mạch, cây cà chua, cây hướng dương, hoặc các cây một lá mầm hoặc cây hai lá mầm khác. Vùng khởi động thích hợp khác cho ý tưởng này được thể hiện trong bảng 8 dưới đây. Các vùng khởi động này, hoặc các thể tương đồng của chúng trong các loài cây trồng có thể rất hữu dụng cho ý tưởng này. Nguyên tắc hoạt động là sự biểu hiện ở tế bào giao tử của bộ máy chỉnh sửa có thể làm tăng tỉ lệ hoặc hiệu quả của sáng chế vì nó có nghĩa là sẽ có nhiều chỉnh sửa protein hoặc ARN có mặt hoặc chuyển cho phôi trong quá trình thụ tinh sao cho sự chỉnh sửa có thể diễn ra nhanh chóng.

Bảng 8. Danh mục vùng khởi động: vùng khởi động có thể sử dụng trong gen chuyển để gây ra sự biểu hiện cao ở tế bào giao tử đực của bộ máy chỉnh sửa để làm tăng mạnh hiệu quả của sự chỉnh sửa đồng thời và sự kích tạo đơn bội được nhân đôi ("SEDHI").

Tên Gen	ID Gen	Ortholog Cây ngô	Ortholog Cây lúa
DUO1	At3G60460	GRMZM2G105137, GRMZM2G046443	LOC_Os04g46384
MGH3	At1G19890	không có	không có
GEX1	At5G55490	GRMZM2G388045	LOC_Os09g27040, LOC_Os07g47194
GEX2	At5G49150	GRMZM2G036832	LOC_Os09g25650
GEX3	At5G16020	GRMZM2G458159	LOC_Os01g42060

HAP2 / GSC1	At4G11720	GRMZM2G412911	LOC_Os05g18730
CycB1	At4G37490	không có	không có
DAZ1	At2G17180	GRMZM2G132057	không có
DAZ2	At4G35280	không có	LOC_Os02g19180
DAZ3	At4G35700	không có	không có
PCR11	At1G68610	không có	không có
DAN1	At3G04620	không có	không có
TIP1	AT3G47440	không có	LOC_Os04g46490
MKKK20	AT3G50310	không có	không có
DAF1	At3G62230	không có	không có
DAW1	At4G35560	GRMZM2G176647	không có
DAU2 / DMP9	At5G39650	không có	không có

VII. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa ở cây hai lá mầm thông qua lai xa hoặc thông qua phép lai với dòng được làm thay đổi bằng CENH3 hoặc dòng kích tạo đơn bội khác.

Sự kích tạo đơn bội in vivo cũng có thể đạt được bằng cách sử dụng phép lai xa khác loài hoặc khác chi trên các loài cây hai lá mầm, ví dụ, ở cây bông (Turcotte et al. 1969, Semigametic production of haploids in pima cotton. Crop Sci. 9:653-655) và cây thuốc lá (Burke et al, 1979, Maternal haploids of *Nicotiana tabacum* L. Science 206:585; Wernsman et al. 1989, Androgenetic *vs.* gynogenetic doubled haploids of tobacco. Crop

Sci. 29:1151-1155). Cây *Arabidopsis* đơn bội có thể thu được bằng lai với phấn từ cây CENH3 đột biến, hoặc bằng cách lai cây trồng này làm thể cái với phấn kiểu dại (Ravi and Chan, 2010, Haploid plants produced by centromere-mediated genome elimination. Nature 464: 615-618). Các gen ứng viên khác mà có thể được cải biến để tạo ra thể kích tạo đơn bội và dòng chỉnh sửa SEDHI bao gồm KNL2 và CENPC (cả hai đều có thể hoạt động thông qua sự loại bỏ hệ gen bố hoặc mẹ qua trung gian tâm động) cũng như là MSI2 và cây hướng dương PLA2. Trong trường hợp này, hệ gen kích tạo đơn bội (là thể đực hoặc thể cái trong phép lai) còn chứa bộ máy chỉnh sửa, sao cho việc chỉnh sửa có thể đạt được trong quy trình kích tạo đơn bội, với kết quả là cây trồng con đơn bội theo dòng mẹ hoặc theo dòng cha đã được chỉnh sửa không có CENH3 đã được biến đổi hoặc sự chỉnh sửa gen chuyển. Xem, ví dụ, tài liệu công bố đơn WO 2017/004375, được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó. Locut chuyển gen biểu hiện bộ máy chỉnh sửa có thể được đưa vào cây hai lá mầm bất kỳ hoặc họ hàng kiểu dại của chúng của cây thuộc họ cải bắp, cây cà chua, cây hồ tiêu, cây rau diếp, cây cà, cây đậu tương, cây hướng dương, cây củ cải đường, cây bông, cây cỏ linh lăng, cây thuốc lá, và các cây khác. Sau đó dòng chuyển gen biểu hiện bộ máy chỉnh sửa được sử dụng làm thể cho phấn, hoặc trong trường hợp của CENH3, làm thể cho phấn hoặc thể nhận phấn, trong phép lai xa khác loài hoặc khác chi để kích tạo đơn bội và sự chỉnh sửa hệ gen đồng thời. Ví dụ, dòng CRISPR-Cas9 chuyển gen *N. africana* biểu hiện sgRNA nhắm đích gibberellin 20-oxiđaza cây thuốc lá được tạo ra thông qua sự biến nạp qua trung gian *Agrobacterium* và được sử dụng để thụ phấn hoa cây thuốc lá đã khử đực để tạo ra cây đơn bội với hệ gen đã chỉnh sửa của chúng ở locut gibberellin 20-oxiđaza. Tốt hơn là, dòng có thể biến nạp dễ dàng với số lượng phấn lớn được sử dụng làm thể cho phấn để kích tạo đơn bội và để cung cấp bộ máy chỉnh sửa tạm thời. Cây nhận đối với sự sản xuất thể đơn bội có hoa mà dễ khử đực hoặc bất thụ đực. Tốt hơn nữa là, thành phần đánh dấu màu sắc hoặc thành phần đánh dấu trực quan khác có mặt trong dòng cảm ứng hoặc được bao gồm trong locut chỉnh sửa để dễ dàng phân biệt phôi hoặc cây trồng đơn bội từ thể lưỡng bội thu được từ sự phát triển hợp tử bình thường.

Chúng tôi minh họa điều này bằng cách sử dụng dòng kích tạo đơn bội *Arabidopsis* trong kiểu sinh thái Columbia, và biến nạp nó với cấu trúc mã hóa cho sự biểu hiện của Cas9 và ARN dẫn đơn lẻ nhắm đích vào gen *GLABROUS1* (*GL1*) mà, khi được làm bất hoạt, mang lại kiểu hình thiếu cấu trúc hình lông. Chúng tôi đã lai T0 làm

thể cái bằng phấn kiểu sinh thái Landsberg Erecta (Ler), và thu được con cháu đơn bội đã chỉnh sửa *gl1*.

Nguyên liệu thể kích tạo đơn bội thu được từ phòng thí nghiệm Comai ở UC Davis. Nguyên liệu này thường được dùng làm dòng kích tạo đơn bội theo dòng cha (gây ra sự sinh sản đơn tính đực, khi lai dưới dạng thể cái với thể đực kiểu dại) nhưng cũng có thể đóng vai trò làm thể kích tạo đơn bội theo dòng mẹ (gây ra sự sinh sản đơn tính cái, khi lai dưới dạng thể đực với thể cái kiểu dại). Các dòng này đã được làm thay đổi để trở thành thể kích tạo đơn bội bằng cách thay thế gen *CENH3* tự nhiên bằng gen chuyển CENH3 *Zea Mays* như báo cáo trong (Maheshwari, *et al*, 2017, Centromere location in Arabidopsis is unaltered by extreme divergence in CENH3 protein sequence. Genome Research 27(3)).

Cụ thể là, cả hai bản sao của gen *AtCENH3* tự nhiên được làm bất hoạt và được bù bằng gen chuyển *ZmCENH3* đã được cài xen ổn định, mà không ảnh hưởng đến sự phát triển cây trồng bình thường, và không tạo ra thể đơn bội khi tự thụ phấn, nhưng tạo ra khoảng 10% thể đơn bội khi lai xa. Đây là sự thay đổi đối với ý tưởng ban đầu về *CENH3-tailswap* được mô tả chi tiết trong (Ravi and Chan, 2010, Haploid plants produced by centromere-mediated genome elimination. Nature 464: 615–618) và nhiều tài liệu công bố tiếp theo.

Sau khi chúng tôi thu được dòng CENH3* từ UC Davis, chúng tôi để chúng lớn lên, xác nhận rằng chúng có gen chuyển *ZmCENH3* và "vô hiệu" đồng hợp tử đối với gen *AtCENH3* tự nhiên. Chúng tôi thực hiện điều này bằng cách thiết kế thử nghiệm taqman qPCR đối với *ZmCENH3* (thử nghiệm #2298) và bằng cách sử dụng PCR và điện di trên gel để kiểm tra 183 cây con về tình trạng tiếp hợp giao tử của kiểu gen *AtCENH3* bằng cách chạy PCR bằng cách sử dụng các đoạn mồi xuôi và ngược Xbal (SEQ NO TKX và TKY) và hỗn hợp Reddy ở nhiệt độ ủ 60°C và cắt bằng enzym giới hạn Xbal qua đêm ở nhiệt độ 37°C. Alen kiểu dại sẽ được cắt bằng enzym này và tạo ra hai băng (189 bp, 25 bp) trong khi thể đột biến vẫn ở 215 bp. Các kiểm tra này cho thấy rằng tất cả các hạt mà UC Davis gửi đến đều là đồng hợp tử đối với alen đột biến *Atcenh3-1*, và có nhiều bản sao của gen chuyển *ZmCENH3* có mặt.

Tin chắc rằng hạt thu được này thực sự là thể kích tạo đơn bội, chúng tôi giữ 100 cây trồng và bắt đầu sự biến nạp nhúng hoa với véc-tơ hai thành phần 24075 (SEQ ID NO: 98) chứa cát xét sgRNA nhắm đích gen *Arabidopsis* (GL1) (AT3G27920) ở hai vị trí đích. Trình tự đích là 5'-GGAAAAGTTGTAGACTGAGA-3', và 5'-GCAGTGATGAACAATGACGG-3' (sợi bổ sung). Sự phá vỡ của gen GL1 tạo ra kiểu hình có thể nhìn thấy của cây không lông một phần hoặc hoàn toàn (cây không lông thiếu cấu trúc hình lông). Gen Cas9 (cCas9-05) trong véc-tơ này được điều khiển bởi vùng khởi động yếu tố kéo dài *Arabidopsis thaliana*. Véc-tơ này còn chứa hai cát xét thành phần đánh dấu chọn lọc mang lại tính kháng Kan và huỳnh quang AmCyan lần lượt được điều khiển bởi vùng khởi động CMP-02 và vùng khởi động *Glycine max* UBI-01. Véc-tơ này được chuyển vào chủng agrobacterium EHA101 và sau đó biến nạp kiểu nhúng hoa vào cây trồng *Arabidopsis* kích tạo đơn bội.

Quy trình biến nạp như sau: Vào buổi sáng chúng tôi phết 24075 EHA101 RecA Agrobacterium thu được từ đĩa vào đĩa YPSpec100Kan50. Chúng tôi nuôi cấy chúng trong tối ở 28°C trong thời gian 24 giờ. Chúng tôi điều chế môi trường thấm (4l): ½ muối XMS (8,66g), 1X vitamin B5 Gamborg (4ml), 5% (khối lượng/thể tích) sucroza (200g), 0,044 µM BAP (12,5mg---12,5mlDMSO) 40µl, sau đó là tiệt trùng bằng bộ lọc. Sau đó chúng tôi bổ sung 250µl 40 mg/ml AS (20mg/l) và 25µl SIlwet L-77(50µl/l) vào 500ml môi trường thấm. Sử dụng que cấy vòng để thu lấy Agrobacterium và cho vào ống 50ml với ~10ml o tiệt trùng bằng bộ lọc, chúng tôi tạo huyền phù Agrobacterium đến khi nó tạo ra 1L với mật độ quang học 600 bằng 0,54. Chúng tôi nhúng chồi ra hoa vào môi trường huyền phù trong thời gian 20-30 giây và sử dụng nắp để đậy khay. Chúng tôi lặp lại việc này lần thứ hai với huyền phù khác có OD600 bằng 0,552.

Khoảng 4 tuần sau khi biến nạp, xấp xỉ 100.000 hạt tự thụ phấn được thu lấy và được ủ ở nhiệt độ 4°C để xuân hóa trong hai ngày, và sau đó hạt được tiệt trùng bằng cách ngâm trong etanol 70% trong thời gian 1 phút và sau đó ngâm trong chất tẩy trắng 50% (thể tích/thể tích) với Triton X-100 0,05% (thể tích/thể tích) trong 10 phút nữa, sau đó rửa hạt trong bốn lần thay đổi nước tiệt trùng. Sau đó đặt hạt trên đĩa kanamycin (50µg/ml) để sàng lọc/chọn lọc nảy mầm trong phòng nuôi cấy mô cây trồng (23°C ban ngày, 24°C ban đêm, 16 giờ chiếu sáng). 38 thể biến nạp dương được xác định vì chúng kháng với sự chọn lọc kanamycin, và chúng được trồng thành cây con trước khi được

chuyển ra đất và lấy mẫu để kiểm tra về sự có mặt của T-ADN Cas9 (thử nghiệm #3049) cũng như là tình trạng của hai vị trí cắt ARN dẫn (thử nghiệm # 3321 và # 3322). 10 biến cố có bản sao đơn lẻ và 15 biến cố có 2 bản sao được xác định mà có cả hai alen của *GL1* bị đột biến và có kiểu hình không có lông. Các cây trồng này được ưu tiên vì chúng thể hiện bằng chứng về hoạt tính Cas9 (do *GL1* bị đột biến và kiểu hình không có lông), chúng có gen chuyển *Cas9* và chúng có gen chuyển *ZmCENH3* theo thử nghiệm qPCR. Các cây trồng này được kích ra hoa trong khoảng thời gian dài bằng cách giữ chúng trong điều kiện sinh trưởng sau đây: 16 giờ sáng, nhiệt độ 23°C ban ngày 20°C ban đêm, độ ẩm tương đối không > 60%.

Vào lúc các cây kích tạo đơn bội mà được biến nạp bằng cấu trúc Cas9 được xác định, chúng tôi gieo hạt và phát triển quần thể hạt Landsberg Erecta (Ler) thu được từ Trung tâm nguồn sinh học *Arabidopsis* (Arabidopsis Biological Resource Center) ở Ohio State University (dòng # CS20). Chúng là hạt kiểu dại và trình tự của các vị trí đích ARN dẫn GL1 ở CS20 phù hợp với trình tự của ARN dẫn trong cấu trúc của chúng tôi. Chúng tôi để cho cả hai quần thể ra hoa và tạo ra khoảng 2000 phép lai có kiểm soát, bằng cách sử dụng cây trồng Ler kiểu dại làm thể cho phấn đực, lai trên xấp xỉ 25 thể kích tạo đơn bội với cấu trúc *Cas9*, mà được sử dụng làm thể cái. Chúng tôi thực hiện 100 phép lai ở mỗi thể cái, đánh dấu hoa đã lai bằng thành phần đánh dấu màu đen và loại bỏ hoa mà chúng tôi không lai do vậy hạn chế có khả năng thu được quả tự thụ phấn. Trong hầu hết các trường hợp, chúng tôi khử đực hoa cái trước khi thụ phấn bằng cách loại bỏ bao phấn bằng kẹp, lại tránh được sự tạp nhiễm bởi hạt tự thụ phấn, nhưng trong một số trường hợp điều này không cần thiết vì bao phấn còn non hoặc đã phát triển bệnh.

Khoảng 15 ngày chúng tôi thu hoạch quả mà đã phát triển màu nâu sáng. Sau đó chúng tôi bóc quả và trồng hạt trong đất. Sau đó đặt chúng trong buồng sinh trưởng 6°C (ban ngày và chiếu sáng), độ dài ngày 8 giờ, chiếu sáng 200umal/m^2s, độ ẩm tương đối 60% trong thời gian 4 ngày. Sau đó chúng tôi chuyển chúng vào buồng sinh trưởng chiếu sáng 16 giờ, nhiệt độ ban ngày 23°C, ban đêm 20°C, độ ẩm không > 60% trong thời gian 7-10 ngày. Chúng tôi quan sát thấy tần suất cao của hạt bị thui ở hầu như tất cả các quả, trung bình khoảng 40–50% tổng số hạt. Số lượng phôi thui này rất nhất quán với năng suất của nguyên liệu kích tạo đơn bội này trong các báo cáo đã công bố. Không

mong muốn bị ràng buộc bởi lý thuyết này, suy xét rằng hạt bị thui nhiều khả năng nhất là được gây ra bởi sự loại bỏ hệ gen một phần hoặc toàn bộ trong nội nhũ dẫn đến sự mất cân bằng và hỏng nội nhũ. Đây là hiện tượng tự nhiên ở dòng kích tạo đơn bội kiểu *CENH3* trong quá trình lai xa và có khả năng không liên quan đến sự có mặt của gen chuyển Cas9. Các phôi thui này không nảy mầm. Do tỉ lệ đều đặn và đáng tin cậy của sự thui phôi ở mọi quả được lai xa, chúng tôi dừng lại ở việc sử dụng sự không có mặt của kiểu hình đó để sàng lọc ra các quả mà tình cờ tự thụ phấn. Bằng cách đó chúng tôi cho nảy mầm các quả mà đã được lai xa.

Tổng cộng chúng tôi thu lấy xấp xỉ 2000 con cháu đã nảy mầm, phần lớn chúng được lai xa. Chúng tôi xác định thể đơn bội đã chỉnh sửa thông qua dạng kết hợp của thử nghiệm thành phần đánh dấu qPCR và/hoặc sàng lọc tính trạng kiểu hình. Các thành phần đánh dấu mà chúng tôi sử dụng để phát hiện thể đơn bội đã chỉnh sửa như sau.

Thứ nhất, chúng tôi tìm kiếm điểm "0" đối với thử nghiệm *ZmCENH3*. Điều này chỉ ra rằng cây trồng là đơn bội vì hệ gen dòng mẹ đã bị mất, và do vậy gen chuyển *ZmCENH3*, mà có mặt trong nhiều bản sao của cây kích tạo đơn bội mẹ, cũng bị mất. Ngược lại, thể lưỡng bội là thể lai giữa các hệ gen theo dòng mẹ và theo dòng cha, và sẽ có điểm Taqman bằng "1" hoặc "2" hoặc cao hơn đối với thử nghiệm này, tùy thuộc vào số lượng bản sao của cây mẹ. Điều mấu chốt là tất cả các thể lưỡng bội thể hiện bằng chứng về gen chuyển này, nhưng thể đơn bội theo dòng cha, chỉ có hệ gen *Ler*, sẽ không có và do đó sẽ bằng "0."

Thứ hai, chúng tôi tìm kiếm điểm "0" đối với thử nghiệm Cas9, mà chỉ ra rằng nó không chuyển gen. Điều này cũng có thể thấy được bằng mắt bằng cách sử dụng ánh sáng huỳnh quang và tìm kiếm CFP thành phần đánh dấu huỳnh quang.

Thứ ba, chúng tôi tìm kiếm điểm "0" đối với một trong số các thử nghiệm vị trí đích *GL1* , mà chỉ ra rằng cây trồng đã được chỉnh sửa. Cây trồng lưỡng bội có thể thể hiện "0," "1" hoặc "2" đối với các thử nghiệm, nhưng thể đơn bội cho thấy "2" hoặc "0." Cái đầu tiên trong số hai ARN dẫn *GL1* dường như có hiệu quả chỉnh sửa cao hơn nhiều so với cái thứ hai, vì thử nghiệm 3321 có ưu thế cao của "0" và "1" ở thể kích tạo đơn bội T0, nhưng 3322 có hầu hết là "2".

Bằng cách sử dụng các thử nghiệm này, chúng tôi có thể xác định thể đơn bội chưa chỉnh sửa (mà có điểm "0" đối với *ZmCENH3* và *Cas9*, nhưng có điểm "2" đối với cả hai vị trí đích *GL1*) và cũng như là thể đơn bội đã chỉnh sửa (mà có điểm "0" đối với các thử nghiệm *ZmCENH3, Cas9 và GL1* (3321)). Chúng tôi cũng có thể xác định thể lai lưỡng bội mà có Cas9 (và thường được chỉnh sửa ở các vị trí GL1) và thể lai lưỡng bội mà không có Cas9 (và thường có một bản sao của GL1 đã chỉnh sửa (từ mẹ) nhưng không có bản sao kia, và do đó có điểm bằng "1" đối với thử nghiệm GL1. Chúng tôi cũng có thể xác định một vài thể đơn bội đã chỉnh sửa giả định vì chúng có điểm bằng "0" đối với thử nghiệm vị trí đích (3321), *ZmCENH3* (2298) và Cas9 (3049). Xem bảng 9 dưới đây về ví dụ của dữ liệu Taqman con cháu từ bố mẹ USR01424136 chứa ba thể đơn bội đã chỉnh sửa giả định (cây trồng 254 trong giếng F2, cây trồng 260 trong giếng D3, và cây trồng 261 ở cây trồng E3).

Bảng 9. Phân tích con cháu từ bố mẹ USR01424136.

PLATE 1045		Vị trí cắt AtGL1-1		Vị trí cắt AtGL1-2		ZmCENH3		Cas9			
Bố mẹ HI là Cas9 bản sao đơn lẻ		3321		3322		2298		3049			
Giếng	ID Cây trồng	# Bản sao thô	Mức # bản sao	# Bản Sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	Đơn bội giả định	Đã chỉnh sửa giả định
E2	usr01424136 x ler-253	0,06	0	0,87	1	4,30	>2	2,93	>2		x
F2	usr01424136 x ler-254	0,00	0	0,32	0 hoặc 1	0,00	0	0,00	0	x	x
G2	usr01424136 x ler-255	1,32	1 hoặc 2	2,06	2	3,16	>2	0,00	0		
H2	usr01424136 x ler-256	0,02	0	0,99	1	2,51	>2	2,99	>2		x
A3	usr01424136 x ler-257	0,04	0	0,87	1	2,40	2	2,84	>2		x
B3	usr01424136 x ler-258	0,03	0	1,64	2	2,99	>2	3,17	>2		x
C3	usr01424136 x ler-259	0,03	0	1,21	1	5,28	>2	5,28	>2		x
D3	USR01424136 X Ler-260	0,06	0	2,01	2	0,00	0	0,00	0	x	x

E3	USR01424136 X Ler-261	0,00	0	2,01	2	0,01	0	0,00	0	x	x
F3	USR01424136 X Ler-262	2,04	2	2,10	2	0,00	0	0,00	0	x	
G3	USR01424136 X Ler-263	1,36	1 hoặc 2	1,25	1	0,00	0	0,00	0	x	
H3	USR01424136 X Ler-264	1,75	2	1,71	2	0,00	0	0,00	0	x	
A4	USR01424136 X Ler-265	0,00	0	1,67	2	3,06	>2	3,16	>2		x
B4	USR01424136 X Ler-266	1,66	2	2,32	2	0,00	0	0,00	0	x	
C4	USR01424136 X Ler-267	2,09	2	1,94	2	3,99	>2	0,00	0		
D4	USR01424136 X Ler-268	1,47	1 hoặc 2	2,08	2	6,34	>2	1,51	1 hoặc 2		
E4	USR01424136 X Ler-269	1,95	2	1,76	2	3,19	>2	0,00	0		
F4	USR01424136 X Ler-270	1,92	2	2,17	2	4,28	>2	0,02	0		
G4	USR01424136 X Ler-271	2,02	2	1,85	2	4,31	>2	0,00	0		
H4	USR01424136 X Ler-272	0,00	0	1,71	2	1,65	2	1,12	1		x

Đơn giản là làm nảy mầm hạt và lấy mẫu để phân tích qPCR Taqman, chúng tôi có thể xác định 8 thể đơn bội đã chỉnh sửa giả định. Thể đơn bội đã chỉnh sửa cũng được xác định bằng cách sàng lọc bằng mắt tính trạng kiểu hình, và sau đó xác nhận sau bằng thử nghiệm Taqman. Chúng tôi sàng lọc về thể đơn bội đã chỉnh sửa bằng cách tìm kiếm cây trồng thiếu cấu trúc hình lông, hoặc không có lông, mà chỉ ra rằng chúng không có alen kiểu dại bất kỳ đối với gen *GL1*, và bằng cách tìm kiếm sự thiếu sự biểu hiện protein huỳnh quang cyan ("CFP") ở phôi hoặc rễ cây con. Điều này chỉ ra rằng chúng thiếu T-ADN Cas9. Chúng tôi quan sát một vài cây trồng này, và đưa chúng đi thử nghiệm Taqman. Đối với ba cây trồng mà chúng tôi đã xác định tính trạng kiểu hình, chúng tôi có thể xác nhận rằng chúng thực sự là thể đơn bội đã chỉnh sửa bằng thử nghiệm Taqman. Chúng tôi biết thực tế là có khả năng một số trong số cây không lông đó mà thiếu CFP là dương tính giả, vì CFP bị câm hoặc vì sự tự thụ phấn của cây mẹ đã được chỉnh sửa hoàn toàn khác và sự sản xuất của con cháu đơn độc vô hiệu, đã được chỉnh sửa hoàn

toàn (và do đó không có lông). Thử nghiệm Taqman có thể phát hiện và sàng lọc ra các thể dương tính giả này, vì chúng được thử nghiệm trực tiếp về sự có mặt của không chỉ gen chuyển Cas9, mà còn của alen *ZmCENH3*, mà chắc chắn có mặt trong hạt tạp nhiễm tự thụ phấn bất kỳ. Chúng tôi phát hiện thấy một vài ví dụ về hạt tự thụ phấn mà tất cả đều đến từ một cây mẹ. Sự thụ phấn ghi lại đối với cây mẹ đó chỉ ra rằng có lượng phấn dồi dào cao mà có thể dẫn đến một số sự tự thụ phấn. Chúng tôi loại trừ các con cháu này khỏi phân tích tổng thể.

Tất cả các thể đơn bội đã chỉnh sửa giả định xác định bằng thử nghiệm Taqman được xác định trình tự. Chúng tôi sử dụng PCR để khuếch đại alen đã được chỉnh sửa, và sau đó tách dòng phụ và xác định trình tự ít nhất là 8 khuẩn lạc đối với mỗi alen đã được chỉnh sửa giả định. Xem bảng 10 về sự thay đổi trình tự chúng tôi đã phát hiện thấy trong thể đơn bội đã chỉnh sửa ở vị trí đích ARN dẫn thứ nhất (thử nghiệm #3321), cũng như là dữ liệu Taqman từ bố mẹ T0. Tổng cộng, chúng tôi phát hiện thấy 19 thể đơn bội đã chỉnh sửa giả định, và chúng tôi xác nhận rằng 3321 vị trí đích có đột biến ở 11 trong số 12 thể đơn bội đã chỉnh sửa mà chúng tôi đã cố gắng xác định trình tự. Liệu 7 thể đơn bội kia cũng có các đột biến hay không sẽ được xác nhận khi xác định trình tự. Xem sự sắp xếp thẳng hàng trình tự đối với các chỉnh sửa này trên hình 24.

Bảng 10. Dữ liệu Taqman và trình tự từ 19 thể đơn bội đã chỉnh sửa.

			Vị trí cắt AtGL1-1		Vị trí cắt AtGL1-2		ZmCENH3		Cas9			
			3321		3322		2298		3049			
Đĩa	Giếng	ID Cây trồng	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	# Bản sao thô	Mức # bản sao	Đột biến vị trí đích	Xác nhận PA
1033	A3	USR0142413 5 X Ler-425	0,00	0	1,67	2	0,04	0	0,00	0	kiểu dại	Không thực hiện
1033	C3	USR0142413 5 X Ler-427	0,21	0	2,43	>2	0,01	0	0,00	0	cài xen A	Có

1033	E4	USR0142413 5 X Ler-437	0,08	0	2,04	2	0,00	0	0,00	0	cài xen T	Có
1042	E5	USR0142413 6 X Ler-25	0,16	0	2,95	>2	0,00	0	0,00	0	cài xen A	Không thực hiện
1042	G10	USR0142413 6 X Ler-67	0,00	0	2,19	2	0,00	0	0,00	0	làm mất AG	Không thực hiện
1042	G12	USR0142413 6 X Ler-83	0,00	0	1,86	2	0,00	0	0,00	0	làm mất G	Không thực hiện
1043	B11	USR0142413 6 X Ler-154	0,16	0	1,59	1 hoặc 2	0,01	0	0,00	0	Không thực hiện	Không thực hiện
1045	F2	USR0142413 6 X Ler-254	0,00	0	0,32	0 hoặc 1	0,00	0	0,00	0	làm mất 8nt*	Không thực hiện
1045	D3	USR0142413 6 X Ler-260	0,06	0	2,01	2	0,00	0	0,00	0	cài xen T	Không thực hiện
1045	E3	USR0142413 6 X Ler-261	0,00	0	2,01	2	0,01	0	0,00	0	làm mất TG	Không thực hiện
1046	D11	USR0143160 9 X Ler-111	0,09	0	1,59	1 hoặc 2	0,02	0	0,01	0	cài xen A	Không thực hiện
1046	G12	USR0143160 9 X Ler-122	0,02	0	1,62	2	0,00	0	0,00	0	cài xen T	Không thực hiện
1046	H12	USR0143160 9 X Ler-123	0,00	0	0,00	0	0,00	0	0,00	0	làm mất CTG	Có

0583	D12	USR0143160 3 X Ler-80	0,00	0	1,50	1 hoặc 2	0,00	0	0,00	0	Không thực hiện	Không thực hiện
0584	A9	USR0143160 3 X Ler-137	0,00	0	1,87	2	0,00	0	0,00	0	Không thực hiện	Không thực hiện
0584	C11	USR0143160 3 X Ler-155	0,05	0	2,06	2	0,00	0	0,17	0	Không thực hiện	Không thực hiện
0584	G11	USR0143160 3 X Ler-159	0,09	0	2,15	2	0,00	0	0,00	0	Không thực hiện	Không thực hiện
0584	C12	USR0143160 3 X Ler-163	0,00	0	1,35	1 hoặc 2	0,00	0	0,11	0	Không thực hiện	Không thực hiện
0584	F12	USR0143160 3 X Ler-166	0,00	0	1,65	2	0,00	0	0,00	0	Không thực hiện	Không thực hiện
0585	H7	USR0143160 3 X Ler-212	0,06	0	2,05	2	0,00	0	0,01	0	Không thực hiện	Không thực hiện
Cây T0 mẹ		USR01424135	0,03	0	1,42	1 hoặc 2	4,46	>2	2,98	>2	ΔG, thể khảm +T	Lưỡng bội
		USR01424136	0,03	0	1,13	1	3,59	>2	2,76	>2	Không thực hiện	Lưỡng bội
		USR01431603	0,14	0	1,25	1	2,48	>2	3,42	>2	Không thực hiện	Lưỡng bội

USR0143160 9	0,18	0	1,1	1	4,7 5	>2	5,5 7	>2	Không thực hiện	Lưỡng bội

*làm mất 16 nt cài xen CTAAACAT

Chúng tôi chạy thêm mẫu lá từ ba cây đơn bội đã chỉnh sửa thông qua phân tích mức bội thể, cùng với ba đối chứng lưỡng bội (lấy mẫu mô từ cây mẹ theo dòng mẹ), mà cho thấy rằng chúng là thể đơn bội thực (các hình 18–23). Điều này đáp ứng để xác nhận lại tình trạng của chúng dưới dạng thể đơn bội đã chỉnh sửa.

Trong ba dòng bố mẹ mà chúng tôi tin chắc rằng không có sự tạp nhiễm tự thụ phấn, chúng tôi không thực hiện việc sàng lọc trước tính trạng kiểu hình bất kỳ, thay vì đó lấy mẫu tất cả con cháu đã nảy mầm để phân tích Taqman (Bảng 11). Ba cây mẹ đối với các con cháu này là USR01431603, USR01431609, và USR01431604. Chúng tôi phát hiện thấy tỉ lệ kích tạo đơn bội là khoảng 9,7% được tính bằng cách chia số lượng con cháu mà thiếu các gen chuyển *ZmCENH3* và *Cas9* (59) cho tổng số con cháu được lấy mẫu (605). Trong số 59 thể đơn bội này chúng tôi phát hiện thấy rằng 10 trong số chúng đã được chỉnh sửa. Điều đó có nghĩa là trung bình có 16,9% thể đơn bội đã được chỉnh sửa bởi Cas9 theo dòng mẹ, trước khi loại bỏ hệ gen dòng mẹ. Không muốn bị ràng buộc bởi con số cuối cùng này, điều này có nghĩa là, bằng cách sử dụng hệ thống này, khi tỉ lệ phần trăm của tổng số con cháu, 9,7% * 16,9% = 1,64% của tất cả con cháu đã nảy mầm là thể đơn bội đã được chỉnh sửa.

Bảng 11. Dữ liệu tỉ lệ kích tạo đơn bội và tỉ lệ chỉnh sửa từ ba tập hợp con cháu, mỗi tập hợp có nguồn gốc từ cây mẹ kích tạo SEDHI khác nhau được lai với phấn Landsberg erecta.

STT	Cây trồng bố mẹ Cas9-05	Cây trồng bố mẹ cNpt2-10	Tổng số mẫu	Số lượng thể đơn bội	Tỉ lệ đơn bội	Thể đơn bội đã chỉnh sửa	Tỉ lệ thể đơn bội đã chỉnh sửa

USR01431603 X Landsberg erecta	>2	>2	230	36	15,65	7	19,44
USR01431609X Landsberg erecta	>2	>2	123	14	11,38	3	21,43
USR01431604 X Landsberg erecta	2	1	252	9	3,57	0	0,00

Tỉ lệ của sự chỉnh sửa đơn bội kiểu *CENH3** hoặc sự chỉnh sửa đơn bội theo dòng cha khác (bằng cách sử dụng dòng kích tạo đơn bội theo dòng mẹ) có thể được tăng lên thông qua việc sử dụng vùng khởi động mà gây ra sự biểu hiện của Cas9 và/hoặc ARN dẫn để tạo ra hàm lượng cao hơn trong tế bào giao tử cái trước khi thụ tinh và /hoặc trong tế bào hợp tử trong hoặc sau khi thụ tinh. Ví dụ về vùng khởi động này là vùng khởi động đối với EA1 (EGG APPARATUS1) (GRMZM2G456746), mặc dù có nhiều ví dụ khác. Người ta cũng có thể biểu hiện Cas9 trong trường hợp của chất tăng cường đặc hiệu bộ máy giao tử cái (EASE), mà là trình tự 77-bp mà kích thích sự biểu hiện của gen tiếp giáp trong tế bào giao tử cái hoặc hợp tử rất sớm (xem, ví dụ, Yang, et al. *An Egg Apparatus-Specific Enhancer of Arabidopsis, Identified by Enhancer Detection*, PLANT PHYSIOLOGY Nov 2005, 139 (3) 1421-1432; DOI: https://doi.org/10.1104/pp.105.068262).

VIII. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa bằng cách cải biến trực tiếp bazơ đích ở trình tự ADN hệ gen.

Sự gây đột biến hướng đích của trình tự ADN cũng có thể đạt được thông qua sự chuyển đổi trực tiếp của một bazơ ADN thành bazơ khác mà không cần phá vỡ sợi kép (DSB). Ví dụ, xytiđin đeaminaza APOBEC1, ađenin đeaminaza, và các thành phần tăng cường khác như Uraxin ADN glycosylaza (UDG) có thể được dung hợp với Cas9 (A840H) nickaza hoặc Cas9 chết bị bất hoạt bằng nucleaza (dCa9) để định hướng sự chỉnh sửa của trình tự ADN mà không đưa vào sự phá vỡ ADN sợi kép (Komor et al. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature doi:10.1038/nature17946; Gaudelli et al. 2017. Programmable

base editing of A:T to G:C in genomic DNA without DNA cleavage. Nature doi:10.1038/nature24644; Komor et al. 2017. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Science Advances,*

Vol. 3, no. 8, eaao4774, DOI: 10.1126/sciadv.aao4774). Loại bộ máy chỉnh sửa bazơ này cũng có thể được chuyển thông qua dòng kích tạo đơn bội để gây ra sự chỉnh sửa bazơ trong trình tự đích một cách trực tiếp ở các giống khác. Ví dụ, trình tự ARN dẫn, xZmVLHP-03 (5'-AGGCGTCGAGCAGCGAGGTG-3', SEQ ID NO: 28) được thiết kế để nhắm đích hệ thống chỉnh sửa bazơ xytiđin đeaminaza để chuyển đổi trình tự hệ gen exon 2 gen ZmVLHP 5'-AGGCGTCGAGCAGCGAGGTG-3' (SEQ ID NO: 28) thành 5'-AGGCGTTGAGCAGCGAGGTG-3' (SEQ ID NO: 29), do đó làm thay đổi bộ ba mã hóa arginin CGA thành bộ ba kết thúc (TGA) trong trình tự mã hóa và gây ra sự kết thúc sớm của trình tự protein và làm bất hoạt gen chức năng. Đột biến C thành T được gạch chân. Tương tự, hệ thống chỉnh sửa bazơ nCas9- hoặc dCas9- ađenin đeaminaza khảm có thể được dùng để gây đột biến vùng mã hóa, ghép nối chỗ tiếp giáp hoặc trình tự vùng khởi động của ZmVLHP hoặc gen khác để tạo ra các biến thể mà có hoạt tính gen đã được làm thay đổi. Cả xytiđin và ađenin đeaminaza đều đặc biệt hữu dụng để làm thay đổi vị trí ghép nối phiên mã vì chỗ tiếp giáp ghép nối chính tắc có trình tự 5'-…..AG/GT …3' (hoặc 5'- ……AC/CT…… 3' trong sợi đối diện).

IX. Sự kích tạo đơn bội đồng thời và sự chỉnh sửa bằng cách thay thế alen bằng khuôn ADN

Không chỉ có thể sử dụng hệ thống kích tạo đơn bội in vivo để đưa protein, ARN hoặc ADN vào để phân cắt hoặc biến đổi trình tự đích, nó cũng có thể được dùng để chuyển khuôn ADN để sửa chữa phụ thuộc độ tương đồng để thay thế trình tự chính xác trong vùng đích ở dạng ADN chuyển gen. ADN khuôn có thể được cài xen vào hệ gen dòng kích tạo mang bộ máy chỉnh sửa hệ gen chẳng hạn như hệ thống CRISPR-Cas9, trong cùng locut chuyển gen hoặc locut khác. Khi cả Cas9-sgRNA và ADN khuôn cùng có mặt trong phôi đơn bội đã được tạo ra, sự phân cắt của trình tự đích sẽ dẫn đến sự sửa chữa của sự phá vỡ nhiễm sắc thể với trình tự ADN chuyển gen tương đồng làm khuôn. Ví dụ, để tạo ra đột biến E149L ở gen ZmPYL-D (GRMZM2G048733_P02)

(xem tài liệu WO16033230, được kết hợp trong bản mô tả này để tham khảo), mảnh ADN chứa trình tự thể cho (5'-CCTTGGTGTTGCCGTCGGGGACGTCGACGACGAATGACAGGATGACGAGC GTCCCTGGCCGGCCGTCGATGACCT-3', SEQ ID NO: 30) được sử dụng làm thể cho sửa chữa. Cần lưu ý rằng trình tự tương đồng khác có thể được bổ sung để củng cố biên trình tự thể cho sửa chữa lõi này. Một hoặc nhiều bản sao của trình tự thể cho sửa chữa này được cài xen vào véc-tơ biểu hiện Cas9-sgRNA 23136 (SEQ ID NO: 31) mà biểu hiện ARN dẫn 5'-GTCGGGGACGTCGACGACGA-3' (SEQ ID NO: 32) để tạo thành véc-tơ cải biến alen pBSC23136-AMD. Cần lưu ý rằng vị trí PAM tiềm năng đã được loại bỏ khỏi các trình tự ADN thể cho sao cho trình tự thể cho đã tích hợp sẽ không bị phân cắt bởi phức hợp Cas9-sgRNA được biểu hiện từ pBSC23136-AMD. pBSC23136-AMD được biến nạp vào dòng kích tạo đơn bội NP2222-HI để tạo ra dòng chỉnh sửa chuyển gen. Dòng kích tạo đơn thể bội chỉnh sửa chuyển gen đơn bội tự thụ phấn để tạo ra dòng con cháu đồng hợp tử ở locut chỉnh sửa. Các dòng đồng hợp tử này được sử dụng để thụ phấn dòng lai cùng dòng ngô cao cấp đích để gây ra sự tạo thành thể đơn bội và cũng đưa vào alen được cải biến bằng Cas9-sgRNA đã được biểu hiện bằng cách sử dụng ADN thể cho có mặt một cách tạm thời trước khi nhiễm sắc thể thể cho phấn bị loại bỏ.

X. Kích tạo thể đơn bội và chỉnh sửa gen đồng thời ở cây lúa

Thu được dòng lúa HI. Ví dụ, ortholog *MATL* lúa, Os03g27610 (SEQ ID NO: 33, bị gây đột biến để tạo ra dòng HI lúa mới. Dòng này được biến nạp với véc-tơ có chứa hệ thống gây đột biến định hướng vị trí để chỉnh sửa hệ gen cây lúa, ví dụ hệ thống CRISPR/Cas9.

Dòng HI lúa được lai với dòng lúa khác, tốt hơn là dòng cao cấp, để tạo ra ít nhất là một phôi đơn bội con cháu. Trong quá trình lai để tạo ra ít nhất là một phôi đơn bội con cháu, cây lúa bố mẹ HI cũng làm cho gây ra bộ máy chỉnh sửa hệ gen, ví dụ, Cas9 cộng với ARN dẫn, được chuyển cho phôi. Tại thời điểm đó, bộ máy chỉnh sửa hoạt động để chỉnh sửa hệ gen của phôi đơn bội, và do đó thu được cây trồng con đơn bội, đã chỉnh sửa.

XI. Thử nghiệm và điều kiện Taqman.

Một vài thử nghiệm được đề cập bằng số hoặc bằng tên đích. Nêu dưới đây là bảng của các thử nghiệm được đề cập ở trên và trình tự của các đoạn mồi và các đoạn dò liên quan. Điều kiện để PCR là tiêu chuẩn đối với tất cả các thử nghiệm và như sau: Biến tính ở nhiệt độ 98°C trong thời gian 2 phút; sau đó là 35 chu kỳ của (i) biến tính ở nhiệt độ 98°C trong thời gian 30 giây, (ii) ủ ở nhiệt độ 60°C trong thời gian 30 giây, (iii) kéo dài ở nhiệt độ 72°C trong thời gian 1 phút; sau đó kéo dài lần cuối cùng ở nhiệt độ 72°C trong thời gian 10 phút với việc giữ ở nhiệt độ 4°C đến khi sẵn sàng. Thử nghiệm được thực hiện ở các điều kiện này trừ khi có chỉ dẫn khác dưới đây.

Bảng 11. Các đoạn mồi và các đoạn dò thử nghiệm.

Đích	Cas9-ở ngô		
Thử nghiệm số	2540	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE09340	TTGTGCTGCTCCACGAACA	39
Đoạn mồi ngược	FE09341	GCCAGCCACTACGAGAAGCT	40
Đoạn dò	FE09342	CTGCTTCTGCTCGTTGTCCTCCGG	41
Đích	matl		
Thử nghiệm số	2827	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10299	GCGGATGCTGGCACAGC	42
Đoạn mồi ngược	FE10300	GGCATTGCTTCCTTCTCCG	43
Đoạn dò	FE10301	CAGGGAGCGAGGTAC	44
Đích	PMI		
Thử nghiệm số	1750	Trình tự	SEQ ID NO

Đoạn mồi xuôi	FE07390	CTGGTGGCCAACGTGAAGTT	45
Đoạn mồi ngược	FE07391	GCTTCACGGGCTGGGTC	46
Đoạn dò	FE07392	AGGCCAAGCCCGCCAACCAG	47

Đích	MATL-WT		
Thử nghiệm số	2826	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10297	GCGGATGCTGGCACAGA	48
Đoạn mồi ngược	FE10298	GCATTGCTTCCTTCGCCA	49
Đoạn dò	FE10299	CAGGGAGGTACGAACC	50

Đích	TAV_4A		
Thử nghiệm số	3252	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11306	GCGGCGAAGAAGCGAA	51
Đoạn mồi ngược	FE11307	GCGGCGTCTCCAGCTTC	52
Đoạn dò	FE11308	CCAGGAACTGCG	53

Đích	TAV_4B		
Thử nghiệm số	3253	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11309	AAGAAACGCCGGCTGAGT	54
Đoạn mồi ngược	FE11310	ACCTTGCGGGGCGTT	55
Đoạn dò	FE11308	CCAGGAACTGCG	56

Đích	TAV_4D		
Thử nghiệm số	3254	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11309	AAGAAACGCCGGCTGAGT	57
Đoạn mồi ngược	FE11311	CCTTGCGCGGCGTC	58
Đoạn dò	FE11308	CCAGGAACTGCG	59
Đích	GW2-01		
Thử nghiệm số	3065	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10799	TGATCCTCGAGGCCAAGCT	60
Đoạn mồi ngược	FE10800	AGGTCGAGGTCCCCTCCA	61
Đoạn dò	FE10801	CCTGCTACCCGGGC	62
Đích	GW2-02		
Thử nghiệm số	3095	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10991	CGCGCCCTGCTACCC	63
Đoạn mồi ngược	FE10992	GCGCGTGCTTACCAGGA	64
Đoạn dò	FE10993	TCGAGGAGTGCCC	65
Đích	TaVHLP2-2A		
Thử nghiệm số	3332	Trình tự	SEQ ID NO:

Đoạn mồi xuôi	FE11312	CACCGATGAGCAGGCG	66
Đoạn mồi ngược	FE11313	AGATACACCTTCCGGCCG	67
Đoạn dò	FE11314	TTCCTCCCGGAAGC	68

Đích	TaVHLP2-2D		
Thử nghiệm số	3333	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11312	CACCGATGAGCAGGCG	69
Đoạn mồi ngược	FE11313	AGATACACCTTCCGGCCAGT	70
Đoạn dò	FE11314	CTCCTCCCGGAAGC	71

Đích			
Thử nghiệm số	3049	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10730	CAAGTTTCTGGACAAGGAGATTCTC	72
Đoạn mồi ngược	FE10731	AAGAATTCCCTTCTTAATAGCTGGAGA	73
Đoạn dò	FE10732	CACGAGCACATTGCTAACCTTGCTGG	74

Đích	TaVHLP2-2B		
Thử nghiệm số	3255	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11315	TCACCGATGAGCAGGCA	75
Đoạn mồi ngược	FE11316	ATACACCTTCCGGCCAGC	76

Đoạn dò	FE11317	TTCCTCCCGGAAGC	77

Đích

Thử nghiệm số	3321	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11540	GATAGGGCTAAAGAGATGTGGGAA	78
Đoạn mồi ngược	FE11541	CTTTGTTCACATTAGGGCTCAAATAA	79
Đoạn dò	FE11542	TAGACTGAGATGGATG	80

Đích

Thử nghiệm số	3322	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE11543	AAAACCACCGGAGAAGACGA	81
Đoạn mồi ngược	FE11544	AGGTGTGGCGGCAGTGA	82
Đoạn dò	FE11545	CACCGTCATTGTTC	83

Đích	Cas9-ở *Arabidopsis*		
Thử nghiệm số	3049	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE10730	CAAGTTTCTGGACAAGGAGATTCTC	84
Đoạn mồi ngược	FE10731	AAGAATTCCCTTCTTAATAGCTGGAGA	85
Đoạn dò	FE10732	CACGAGCACATTGCTAACCTTGCTGG	86

Đích ZmCENH3

Thử nghiệm số	2298	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE08737	GCGACGCCGGAAAGG	87
Đoạn mồi ngược	FE08738	TGGCGTGGTTTCGTCTTCTTA	88
Đoạn dò	FE08739	AAGAGCGGCGTCTGGAGGTGACTCA	89

Đích	Vị trí đích GL1 3321 (PCR)	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	3321F	AACCGCATCGTCAGAAAAAC	90
Đoạn mồi ngược	3321R	TCAACTTAACCGGCCAAATC	91
Nhiệt độ ủ	60 °C		

Đích	Vị trí đích VLHP2-2A (PCR)	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE4117	CATCCCTTCTCTTCCCTCCTG	92
Đoạn mồi ngược	FE4118	GCCAGTGTGAGTGTGTATGAGCA	93
Nhiệt độ ủ	61 °C		

Đích	Vị trí đích VLHP2-2B (PCR)	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE4120	CATCGTTTTCTCCCCTCCTCA	94

Đoạn mồi ngược	FE4121	ACTGATATGCACGGCGCCA	95
Nhiệt độ ủ	62 °C		

Đích	Vị trí đích VLHP2-2D (PCR)	Trình tự	SEQ ID NO:
Đoạn mồi xuôi	FE4121	TGCAGTAGCTTCATTTTCACCG	96
Đoạn mồi ngược	FE4122	AGGAATTGATATGTACGCCCGT	97
Nhiệt độ ủ	61 °C		

Danh mục trình tự

Đơn này có kèm theo danh mục trình tự có tên là 81189_ST25.txt, được tạo ra ngày 4 tháng 12 năm 2017, có kích thước xấp xỉ 334 kilobyte. Danh mục trình tự này được kết hợp ở đây để tham khảo đến toàn bộ nội dung của nó. Danh mục trình tự này được nộp qua EFS-Web, và tuân theo 37 C.F.R. § 1.824(a)(2)-(6) và (b).

YÊU CẦU BẢO HỘ

1. Phương pháp chỉnh sửa ADN hệ gen ở cây hai lá mầm, bao gồm các bước:

a) thu lấy cây hai lá mầm thứ nhất bao gồm đột biến trong gen histon 3 của tâm động (CENH3) và tùy ý gen chuyển hoán đổi đuôi CENH3 ở ngô, trong đó cây hai lá mầm thứ nhất biểu hiện enzym cải biến ADN và tùy ý ít nhất một axit nucleic hướng dẫn;

b) thu lấy cây hai lá mầm thứ hai, trong đó cây hai lá mầm thứ hai bao gồm ADN hệ gen cây hai lá mầm cần được chỉnh sửa;

c) thụ phấn cây hai lá mầm thứ nhất bằng phấn hoa từ cây hai lá mầm thứ hai; và

d) chọn lọc ít nhất một con cháu đơn bội được tạo ra bởi sự thụ phấn của bước (c) trong đó con cháu đơn bội bao gồm hệ gen của cây hai lá mầm thứ hai chứ không phải cây hai lá mầm thứ nhất, và hệ gen của con cháu đơn bội đã được cải biến bằng enzym cải biến ADN và tùy ý ít nhất một axit nucleic hướng dẫn được chuyển bởi cây hai lá mầm thứ nhất;

trong đó đột biến trong gen CENH3 được chọn từ nhóm bao gồm đột biến mất chức năng, đột biến mất chức năng một phần, đột biến dịch khung được phục hồi, và đột biến mất đoạn trong khung đọc.

2. Phương pháp theo điểm 1 còn bao gồm bước:

e) cho phép ít nhất một con cháu đơn bội được chọn trong bước (d) trải qua quá trình nhân đôi nhiễm sắc thể.

3. Phương pháp theo điểm 2, trong đó con cháu đơn bội được chỉnh sửa trải qua quá trình nhân đôi nhiễm sắc thể tự phát, do đó tạo ra con cháu đơn bội kép được chỉnh sửa.

4. Phương pháp theo điểm 2, trong đó con cháu đơn bội được chỉnh sửa được xử lý bằng tác nhân gây nhân đôi nhiễm sắc thể, do đó tạo ra con cháu đơn bội kép được chỉnh sửa.

5. Phương pháp theo điểm 4, trong đó tác nhân gây nhân đôi nhiễm sắc thể là colchixin, pronamit, dithipyr, trifluralin, hoặc tác nhân chống vi ống đã biết khác.

6. Phương pháp theo điểm 1, trong đó enzym cải biến ADN là nucleaza định hướng vị trí được chọn từ nhóm bao gồm meganucleaza (MN), nucleaza ngón tay kẽm (ZFN), nucleaza tác động giống chất kích hoạt phiên mã (TALEN), Cas9 nucleaza, Cpf1 nucleaza (Cas12a), dCas9-FokI, dCpf1-FokI, Cas9-xytidin deaminaza khảm, Cas9-adenin deaminaza khảm, FEN1-FokI khảm, và Mega-TAL, nickaza Cas9 (nCas9), dCas9 không phải-FokI nucleaza khảm, dCpf1 không phải-FokI nucleaza, Cpf1-xytidin deaminaza khảm, và Cpf1-adenin deaminaza.

7. Phương pháp theo điểm 1, trong đó ít nhất một axit nucleic hướng dẫn là ARN hướng dẫn.

8. Phương pháp theo điểm 7, trong đó ARN hướng dẫn là trình tự gồm 18-21 nucleotit và ít nhất 90% tương đồng với trình tự đích.

9. Phương pháp theo điểm 8, trong đó trình tự đích là SEQ ID NO: 103, và trong đó ADN hệ gen của cây hai lá mầm ADN hệ gen của *Arabidopsis*.

10. Phương pháp theo điểm 1, trong đó cây hai lá mầm thứ nhất biểu hiện gen đánh dấu.

11. Phương pháp theo điểm 10, trong đó gen đánh dấu được chọn từ nhóm bao gồm GUS, PMI, PAT, GFP, RFP, CFP, B1, C1, R-nj, sắc tố anthoxyanin.

12. Phương pháp theo điểm 1, trong đó cây hai lá mầm thứ nhất là cây hai lá mầm có thể biến nạp được chọn và/ hoặc có nguồn gốc từ nhóm bao gồm *Arabidopsis, Brassica,* đậu tương, cà chua, hướng dương, và bông.

13. Phương pháp theo điểm 1, trong đó đột biến mất chức năng một phần là đột biến dịch khung gần đầu 3' của gen CENH3.

14. Phương pháp theo điểm 1, trong đó đột biến mất chức năng là đột biến bất hoạt.

15. Phương pháp theo điểm 1, trong đó ít nhất một con cháu đơn bội bao gồm SEQ ID NO: 104, 105, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, hoặc 117, và hơn nữa trong đó ít nhất một con cháu đơn bội là *Arabidopsis*.

DANH MỤC TRÌNH TỰ

<110> Syngenta Participations AG
Que, Qiudeng
Kelliher, Timothy

<120> PHƯƠNG PHÁP CHỈNH SỬA ADN HỆ GEN Ở CÂY HAI LÁ MẦM

<130> 81189-WO-REG-ORG-P-1

<150> US 62/429260
<151> 2016-12-02

<160> 98

<170> PatentIn phiên bản 3.5

<210> 1
<211> 15722
<212> ADN
<213> Trình tự Nhân tạo

<220>
<223> vecto 23396

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05

<220>
<221> vùng khởi động
<222> (304)..(2100)

<223> prSoUbi4-04

<220>
<221> gen
<222> (2117)..(6286)
<223> cCas9-01

<220>
<221> đột biến
<222> (5606)..(5608)
<223> đột biến L thành V

<220>
<221> đột biến
<222> (5651)..(5653)
<223> đột biến I thành V

<220>
<221> trình tự kết thúc
<222> (6292)..(6544)
<223> tNOS-05-01

<220>
<221> vùng khởi động
<222> (6551)..(6925)
<223> prOsU3-01

<220>
<221> đặc điểm_hỗn tạp
<222> (6927)..(6946)
<223> xZmVLHP

<220>
<221> đặc điểm_hỗn tạp

<222> (6927)..(7032)
<223> rsgRNAZmVLHP-01

<220>
<221> vùng khởi động
<222> (7042)..(9033)
<223> prUbi1-04

<220>
<221> gen
<222> (9050)..(10228)
<223> cPMI-09

<220>
<221> trình tự kết thúc
<222> (10251)..(10503)
<223> tNOS-05-01

<220>
<221> đặc điểm_hỗn tạp
<222> (10547)..(10676)
<223> bNLB-03

<220>
<221> gen
<222> (10956)..(11744)
<223> cSpec-03

<220>
<221> vùng khởi động
<222> (11839)..(11969)
<223> prVirG-01

<220>

<221> gen

<222> (12706)..(13779)

<223> cRepA-01

<400> 1

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560

acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620

gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680

tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740

ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800

ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860

gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920

ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980

ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040

tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100

gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160

cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220

gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280

cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340

gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400

cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460

gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520

ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580

cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640

cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700

gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760

gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820

cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880

cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940

ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000

cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060

cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120

gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180

caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240

catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300

ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360

cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420

caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480

ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540

gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600

gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660

cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720

gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780

caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840
cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960
cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020
gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080
gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200
gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260
ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380
gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440
gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500
cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560
cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620
ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680
ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740
ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800
gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860
cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920
ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980
gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040
gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100
cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160
gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220
gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280
cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340
caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400
gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460
cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520
cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580
gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640
gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700
cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760
cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820

cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880
gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940
gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000
caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060
gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120
gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180
gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240
gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300
aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360
atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420
tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480
aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540
gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600
taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660
acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720
acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcagcag gaggcgtcga gcagcggttt tagagctaga 6960
aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020
gctttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080
ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140
gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttagt 7440
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500
ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa 7560
taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680
cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740
agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860

acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920
gacacccccт ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980
accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040
ccccccccccc cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100
agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160
cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220
ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280
tttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340
gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400
gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460
ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520
atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580
atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640
attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700
gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760
gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820
tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880
ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940
gatttttta gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000
tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060
gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120
catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180
cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240
cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300
gttcaaggtg ctgtgcgccg cccagcccct gagcatccag gtgcacccca acaagcacaa 9360
cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420
caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480
catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540
ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600
cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660
cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720
ccccgaggac agcggcctgt tcagcccccт gctgctgaac gtggtgaagc tgaaccccgg 9780
cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840
ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900

ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960
gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020
cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080
ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140
cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200
ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500
atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560
aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620
gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680
aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740
gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800
gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860
caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920
gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980
tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040
tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100
atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160
acctttttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220
ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280
ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340
ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400
cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460
ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520
cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580
ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640
aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700
ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760
cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820
tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880
tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940

tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000
cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060
tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120
caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180
cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240
tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300
gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360
gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420
ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480
cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540
tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600
aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740
cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860
attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980

tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100
tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160
ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220
cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280
ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340
cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400
gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460
aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640
gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760
agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240
gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc 15300
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720
ta 15722

<210> 2
<211> 20
<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> trình tự gRNA để chỉnh sửa VLHP1

<400> 2

gcaggaggcg tcgagcagcg 20

<210> 3

<211> 15722

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> vectơ 23399

<220>

<221> đặc điểm_hỗn tạp

<222> (4)..(259)

<223> bNRB-05

<220>

<221> vùng khởi động

<222> (304)..(2100)

<223> prSoUbi4-04

<220>

<221> gen

<222> (2117)..(6286)

<223> cCas9-01

<220>

<221> đột biến

<222> (5606)..(5608)

<223> đột biến L thành V

<220>

<221> đột biến

<222> (5651)..(5653)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc
<222> (6292)..(6544)
<223> tNOS-05-01
<220>
<221> vùng khởi động
<222> (6551)..(6925)
<223> prOsU3-01
<220>
<221> đặc điểm_hỗn tạp
<222> (6927)..(6946)
<223> xZmGW2
<220>
<221> đặc điểm_hỗn tạp
<222> (6927)..(7031)
<223> rsgRNAZmGW2-02
<220>
<221> vùng khởi động
<222> (7042)..(9033)
<223> prUbi1-04
<220>
<221> gen
<222> (9050)..(10228)
<223> cPMI-09
<220>
<221> trình tự kết thúc
<222> (10251)..(10503)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp
<222> (10547)..(10676)
<223> bNLB-03
<220>
<221> gen
<222> (10956)..(11744)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (11839)..(11969)

<223> prVirG-01

<220>

<221> gen

<222> (12706)..(13779)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (13822)..(14226)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14904)..(15710)

<223> oCOLE-06

<400> 3

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tccccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560

acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620

gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680

tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740

ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800

ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860

gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920

ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980

ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040

tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100

gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160

cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220

gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280

cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340

gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400

cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460

gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520

ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580

cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640

cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700

gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760

gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820

cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880

cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940

ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000

cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060
cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120
gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240
catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300
ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360
cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420
caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480
ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540
gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600
gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660
cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720
gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780
caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840
cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960
cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020
gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080
gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200
gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260
ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380
gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440
gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500
cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560
cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620
ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680
ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740
ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800
gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860
cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920
ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980
gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040

gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100
cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160
gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220
gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280
cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340
caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400
gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460
cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520
cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580
gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640
gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700
cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760
cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820
cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880
gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940
gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000
caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060
gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120
gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180
gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240
gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300
aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360
atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420
tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480
aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540
gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600
taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660
acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720
acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcaaagc tcgcgccctg ctacccgttt tagagctaga 6960
aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020
gctttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080

ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140
gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttttagt 7440
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500
ttttatttaa taatttagat ataaaatага ataaaataaa gtgactaaaa attaaacaaa 7560
taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680
cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740
agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860
acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920
gacacccccct ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980
accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040
ccccccccccc cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100
agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160
cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220
ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280
ttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340
gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400
gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460
ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520
atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580
atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640
attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700
gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760
gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820
tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880
ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940
gatttttta gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000
tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060
gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120

catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180
cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240
cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300
gttcaaggtg ctgtgcgccg cccagccccct gagcatccag gtgcacccca acaagcacaa 9360
cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420
caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480
catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540
ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600
cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660
cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720
ccccgaggac agcggcctgt tcagcccccct gctgctgaac gtggtgaagc tgaaccccgg 9780
cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840
ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900
ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960
gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020
cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080
ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140
cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200
ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500
atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560
aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620
gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680
aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740
gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800
gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860
caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920
gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980
tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040
tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100
atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160

acctttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220
ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280
ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340
ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400
cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460
ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520
cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580
ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640
aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700
ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760
cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820
tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880
tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940
tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000
cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060
tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120
caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180
cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240
tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300
gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360
gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420
ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480
cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540
tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600
aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200

gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740
cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860
attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980
tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100
tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160
ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220
cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280
ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340
cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400
gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460
aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640
gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760
agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240

gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc 15300
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720
ta 15722

<210> 4
<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> để chỉnh sửa GW2-2
<400> 4
aagctcgcgc cctgctaccc 20

<210> 5
<211> 19617
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 22808

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (304)..(2100)
<223> prSoUbi4-04

<220>
<221> gen
<222> (2119)..(5193)
<223> cTNPLAIIAFw-01
<220>
<221> trình tự kết thúc
<222> (5200)..(5452)
<223> tNOS-05-01
<220>
<221> vùng khởi động
<222> (5486)..(7478)
<223> prUbi1-10
<220>
<221> gen
<222> (7492)..(10566)
<223> cTNPLAIIARv-01
<220>
<221> trình tự kết thúc
<222> (10573)..(10825)
<223> tNOS-05-01
<220>
<221> vùng khởi động
<222> (10844)..(12835)
<223> prUbi1-04
<220>
<221> gen
<222> (12852)..(14030)
<223> cPMI-09
<220>
<221> trình tự kết thúc
<222> (14053)..(14305)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp

<222> (14349)..(14478)

<223> bNLB-03

<220>

<221> gen

<222> (14758)..(15546)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (15641)..(15771)

<223> prVirG-01

<220>

<221> gen

<222> (15846)..(16571)

<223> cVirG-01

<220>

<221> gen

<222> (16601)..(17674)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (17717)..(18121)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (18799)..(19605)

<223> oCOLE-06

<400> 5

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560

acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620

gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680

tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740

ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800

ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860

gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920

ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980

ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040

tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100

gatcgcggcc gcgccaccat gggaaaacct attcctaatc ctctgctggg cctggattct 2160

accggaggca tggcccctaa gaaaaagcgg aaggtggacg gcggagtgga cctgagaaca 2220

ctgggatatt ctcagcagca gcaggagaag atcaagccca aggtgagatc tacagtggcc 2280

cagcaccacg aagccctggt gggacacgga tttacacacg cccacattgt ggccctgtct 2340

cagcaccctg ccgccctggg aacagtggcc gtgaaatatc aggatatgat tgccgccctg 2400

cctgaggcca cacacgaagc cattgtggga gtgggaaaac agtggtctgg agccagagcc 2460

ctggaagccc tgctgacagt ggccggagaa ctgagaggac ctcctctgca gctggataca 2520

ggacagctgc tgaagattgc caaaaggggc ggagtgaccg cggtggaagc cgtgcacgcc 2580

tggagaaatg ccctgacagg agcccctctg aacctgaccc ccgaacaggt ggtggccatt 2640

gccagccacg acggcggcaa gcaggccctg gaaaccgtgc agagactgct gcccgtgctg 2700

tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctca cgacggagga 2760

aaacaggctc tggaaacagt gcagcggctg ctgcctgtgc tgtgtcaggc tcacggcttg 2820

actccagaac aggtggtggc tattgcttcc aatattgggg ggaaacaggc cctggaaact 2880

gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acaggtggtg 2940

gccattgcca gcaacaacgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc 3000

gtgctgtgcc aggcccatgg cctgacacct gaacaggtgg tggctatcgc ctctaacaac 3060

ggaggaaaac aggctctgga aacagtgcag cggctgctgc ctgtgctgtg tcaggctcac 3120

ggcttgactc cagaacaggt ggtggctatt gcttccaaca acgggggggaa acaggccctg 3180

gaaactgtgc agcgcctgct gccagtgctg tgccaggctc acgggctgac ccccgaacag 3240

gtggtggcca ttgccagcaa cggcggcggc aagcaggccc tggaaaccgt gcagagactg 3300

ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtggtggc tatcgcctct 3360

cacgacggag gaaaacaggc tctggaaaca gtgcagcggc tgctgcctgt gctgtgtcag 3420

gctcacggct tgactccaga acaggtggtg gctattgctt ccaatattgg ggggaaacag 3480

gccctggaaa ctgtgcagcg cctgctgcca gtgctgtgcc aggctcacgg cctcactccc 3540

gaacaggtgg tggccattgc cagcaacatc ggcggcaagc aggccctgga aaccgtgcag 3600

agactgctgc ccgtgctgtg ccaggcccat ggcctgacac ctgaacaggt ggtggctatc 3660

gcctctcacg acggaggaaa acaggctctg gaaacagtgc agcggctgct gcctgtgctg 3720

tgtcaggctc acggcttgac tccagaacag gtggtggcta ttgcttccaa caacgggggg 3780

aaacaggccc tggaaactgt gcagcgcctg ctgccagtgc tgtgccaggc tcacggactg 3840

acccccgaac aggtggtggc cattgccagc aacggcggcg gcaagcaggc cctggaaacc 3900

gtgcagagac tgctgcccgt gctgtgccag gcccatggcc tgacacctga acaggtggtg 3960

gctatcgcct ctaacaacgg aggaaaacaa gcactcgaga cagtgcagcg gctgctgcct 4020

gtgctgtgtc aggctcacgg cttgactcca gaacaggtgg tggctattgc ttccaacaac 4080

ggggggaaac aggccctgga aactgtgcag cgcctgctgc cagtgctgtg ccaggctcac 4140

gggctgaccc ccgaacaggt ggtggccatt gccagcaaca tcggcggcaa gcaggccctg 4200

gaaaccgtgc agagactgct gcccgtgctg tgccaggccc atggcctgac acctgaacag 4260

gtggtggcta tcgcctctaa caacggagga aaacaggctc tggaaacagt gcagcggctg 4320

ctgcctgtgc tgtgtcaggc tcacggcttg actccacagc aggtcgtggc aattgctagc 4380

aatatcggcg gacggcccgc cctggagagc attgtggccc agctgtctag acctgatcct 4440

gccctggccg ccctgacaaa tgatcacctg gtggccctgg cctgtctggg aggcagacct 4500

gccctggatg ccgtgaaaaa aggactgcct cacgcccctg ccctgattaa aagaacaaat 4560
agaagaatcc ccgagcggac ctctcacaga gtggccggat cccagctggt gaaatctgag 4620
ctggaggaga agaagtctga gctgagacac aagctgaagt acgtgcctca cgagtacatc 4680
gagctgatcg agatcgccag aaatagcacc caggatagaa tcctggagat gaaggtgatg 4740
gagttcttca tgaaagtgta cggctacaga ggaaagcatc tgggaggaag cagaaaacct 4800
gacggagcca tttatacagt gggcagccct atcgattatg gcgtgatcgt ggatacaaag 4860
gcctacagcg gaggctacaa tctgcctatt ggacaggccg atgagatgca gagatacgtg 4920
gaggagaacc aaaccaggaa caagcatatc aaccctaacg agtggtggaa ggtgtaccct 4980
tctagcgtga ccgagttcaa gttcctgttt gtgagcggcc acttcaaggg caattataag 5040
gcccagctga ccaggctgaa ccacatcaca aattgtaatg gcgccgtgct gtctgtggag 5100
gaactgctga ttggaggaga gatgattaag gccggaacac tgacactgga ggaggtgaga 5160
agaaagttca acaacggcga gatcaacttc tgaaagcttg atcgttcaaa catttggcaa 5220
taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat ataatttctg 5280
ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt tatgagatgg 5340
gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa caaaatatag 5400
cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga tcttcgaacc 5460
ctagtcgaag acaaccggtg catgcctgca gtgcagcgtg acccggtcgt gcccctctct 5520
agagataatg agcattgcat gtctaagtta taaaaaatta ccacatattt tttttgtcac 5580
acttgtttga agtgcagttt atctatcttt atacatatat ttaaacttta ctctacgaat 5640
aatataatct atagtactac aataatatca gtgttttaga gaatcatata aatgaacagt 5700
tagacatggt ctaaaggaca attgagtatt ttgacaacag gactctacag ttttatcttt 5760
ttagtgtgca tgtgttctcc tttttttttg caaatagctt cacctatata atacttcatc 5820
cattttatta gtacatccat ttagggttta gggttaatgg tttttataga ctaatttttt 5880
tagtacatct attttattct attttagcct ctaaattaag aaaactaaaa ctctatttta 5940
gtttttttat ttaataattt agatataaaa tagaataaaa taaagtgact aaaaattaaa 6000
caaataccct ttaagaaatt aaaaaaacta aggaaacatt tttcttgttt cgagtagata 6060
atgccagcct gttaaacgcc gtcgacgagt ctaacggaca ccaaccagcg aaccagcagc 6120
gtcgcgtcgg gccaagcgaa gcagacggca cggcatctct gtcgctgcct ctggacccct 6180
ctcgagagtt ccgctccacc gttggacttg ctccgctgtc ggcatccaga aattgcgtgg 6240
cggagcggca gacgtgagcc ggcacggcag gcggcctcct cctcctctca cggcaccggc 6300
agctacgggg gattcctttc ccaccgctcc ttcgctttcc cttcctcgcc cgccgtaata 6360
aatagacacc ccctccacac cctctttccc caacctcgtg ttgttcggag cgcacacaca 6420
cacaaccaga tctcccccaa atccacccgt cggcacctcc gcttcaaggt acgccgctcg 6480
tcctcccccc ccccccctct ctaccttctc tagatcggcg ttccggtcca tggttagggc 6540

ccggtagttc tacttctgtt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct 6600
gctagcgttc gtacacggat gcgacctgta cgtcagacac gttctgattg ctaacttgcc 6660
agtgtttctc tttggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt 6720
catgattttt tttgtttcgt tgcatagggt ttggtttgcc cttttccttt atttcaatat 6780
atgccgtgca cttgtttgtc gggtcatctt ttcatgcttt tttttgtctt ggttgtgatg 6840
atgtggtctg gttgggcggt cgttctagat cggagtagaa ttctgtttca aactacctgg 6900
tggatttatt aattttggat ctgtatgtgt gtgccataca tattcatagt tacgaattga 6960
agatgatgga tggaaatatc gatctaggat aggtatacat gttgatgcgg gttttactga 7020
tgcatataca gagatgcttt ttgttcgctt ggttgtgatg atgtggtgtg gttgggcggt 7080
cgttcattcg ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa 7140
ttttggaact gtatgtgtgt gtcatacatc ttcatagtta cgagtttaag atggatggaa 7200
atatcgatct aggataggta tacatgttga tgtgggtttt actgatgcat atacatgatg 7260
gcatatgcag catctattca tatgctctaa ccttgagtac ctatctatta taataaacaa 7320
gtatgtttta taattatttt gatcttgata tacttggatg atggcatatg cagcagctat 7380
atgtggattt ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt 7440
gtcgatgctc accctgttgt ttggtgttac ttctgcagcg gccgcgccac catgggaaaa 7500
cctattccta atcctctgct gggcctggat tctaccggag gcatggcccc taagaaaaag 7560
cggaaggtgg acggcggagt ggacctgaga acactgggat attctcagca gcagcaggag 7620
aagatcaagc ccaaggtgag atctacagtg gcccagcacc acgaagccct ggtgggacac 7680
ggatttacac acgcccacat tgtggccctg tctcagcacc ctgccgccct gggaacagtg 7740
gccgtgaaat atcaggatat gattgccgcc ctgcctgagg ccacacacga agccattgtg 7800
ggagtgggaa aacagtggtc tggagccaga gccctggaag ccctgctgac agtggccgga 7860
gaactgagag gacctcctct gcagctggat acaggacagc tgctgaagat tgccaaaagg 7920
ggcggagtga ccgcggtgga agccgtgcac gcctggagaa atgccctgac aggagcccct 7980
ctgaacctga cccccgaaca ggtggtggcc attgccagca acaacggcgg caagcaggcc 8040
ctggaaaccg tgcagagact gctgcccgtg ctgtgccagg cccatggcct gacacctgaa 8100
caggtggtgg ctatcgcctc tcacgacgga ggaaaacagg ctctggaaac agtgcagcgg 8160
ctgctgcctg tgctgtgtca ggctcacggc ttgactccag aacaggtggt ggctattgct 8220
tccaacggcg gggggaaaca ggccctggaa actgtgcagc gcctgctgcc agtgctgtgc 8280
caggctcacg gactgacccc cgaacaggtg gtggccattg ccagcaacgg cggcggcaag 8340
caggccctgg aaaccgtgca gagactgctg cccgtgctgt gccaggccca tggcctgaca 8400
cctgaacagg tggtggctat cgcctctcac gacggaggaa aacaggctct ggaaacagtg 8460
cagcggctgc tgcctgtgct gtgtcaggct cacggcttga ctccagaaca ggtggtggct 8520
attgcttccc acgacggggg gaaacaggcc ctggaaactg tgcagcgcct gctgccagtg 8580

ctgtgccagg ctcacgggct gacccccgaa caggtggtgg ccattgccag caacggcggc 8640
ggcaagcagg ccctggaaac cgtgcagaga ctgctgcccg tgctgtgcca ggcccatggc 8700
ctgacacctg aacaggtggt ggctatcgcc tctaacggcg gaggaaaaca ggctctggaa 8760
acagtgcagc ggctgctgcc tgtgctgtgt caggctcacg gcttgactcc agaacaggtg 8820
gtggctattg cttcccacga cgggggggaaa caggccctgg aaactgtgca gcgcctgctg 8880
ccagtgctgt gccaggctca cggcctcact cccgaacagg tggtggccat tgccagcaac 8940
aacggcggca agcaggccct ggaaaccgtg cagagactgc tgcccgtgct gtgccaggcc 9000
catggcctga cacctgaaca ggtggtggct atcgcctctc acgacggagg aaaacaggct 9060
ctggaaacag tgcagcggct gctgcctgtg ctgtgtcagg ctcacggctt gactccagaa 9120
caggtggtgg ctattgcttc ccacgacggg gggaaacagg ccctggaaac tgtgcagcgc 9180
ctgctgccag tgctgtgcca ggctcacgga ctgacccccg aacaggtggt ggccattgcc 9240
agcaacatcg gcggcaagca ggccctggaa accgtgcaga gactgctgcc cgtgctgtgc 9300
caggcccatg gcctgacacc tgaacaggtg gtggctatcg cctctaacaa cggaggaaaa 9360
caagcactcg agacagtgca gcggctgctg cctgtgctgt gtcaggctca cggcttgact 9420
ccagaacagg tggtggctat tgcttccaac ggcgggggga aacaggccct ggaaactgtg 9480
cagcgcctgc tgccagtgct gtgccaggct cacgggctga cccccgaaca ggtggtggcc 9540
attgccagcc acgacggcgg caagcaggcc ctggaaaccg tgcagagact gctgcccgtg 9600
ctgtgccagg cccatggcct gacacctgaa caggtggtgg ctatcgcctc taatatcgga 9660
ggaaaacagg ctctggaaac agtgcagcgg ctgctgcctg tgctgtgtca ggctcacggc 9720
ttgactccac agcaggtcgt ggcaattgct agccacgacg gcggacggcc cgccctggag 9780
agcattgtgg cccagctgtc tagacctgat cctgccctgg ccgccctgac aaatgatcac 9840
ctggtggccc tggcctgtct gggaggcaga cctgccctgg atgccgtgaa aaaaggactg 9900
cctcacgccc ctgccctgat taaaagaaca aatagaagaa tccccgagcg gacctctcac 9960
agagtggccg gatcccagct ggtgaaatct gagctggagg agaagaagtc tgagctgaga 10020
cacaagctga agtacgtgcc tcacgagtac atcgagctga tcgagatcgc cagaaatagc 10080
acccaggata gaatcctgga gatgaaggtg atggagttct tcatgaaagt gtacggctac 10140
agaggaaagc atctgggagg aagcagaaaa cctgacggag ccatttatac agtgggcagc 10200
cctatcgatt atggcgtgat cgtggataca aaggcctaca gcggaggcta caatctgcct 10260
attggacagg ccgatgagat gcagagatac gtggaggaga accaaaccag gaacaagcat 10320
atcaacccta acgagtggtg gaaggtgtac ccttctagcg tgaccgagtt caagttcctg 10380
tttgtgagcg gccacttcaa gggcaattat aaggcccagc tgaccaggct gaaccacatc 10440
acaaattgta atggcgccgt gctgtctgtg gaggaactgc tgattggagg agagatgatt 10500
aaggccggaa cactgacact ggaggaggtg agaagaaagt tcaacaacgg cgagatcaac 10560
ttctgaaagc ttgatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt 10620

gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt 10680
aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta 10740
tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc 10800
gcggtgtcat ctatgttact agatcttcga agacggaccg cgcctgcagt gcagcgtgac 10860
ccggtcgtgc ccctctctag agataatgag cattgcatgt ctaagttata aaaaattacc 10920
acatattttt tttgtcacac ttgtttgaag tgcagtttat ctatctttat acatatattt 10980
aaactttact ctacgaataa tataatctat agtactacaa taatatcagt gttttagaga 11040
atcatataaa tgaacagtta gacatggtct aaaggacaat tgagtatttt gacaacagga 11100
ctctacagtt ttatcttttt agtgtgcatg tgttctcctt ttttttgca aatagcttca 11160
cctatataat acttcatcca ttttattagt acatccattt agggtttagg gttaatggtt 11220
tttatagact aatttttta gtacatctat tttattctat tttagcctct aaattaagaa 11280
aactaaaact ctattttagt ttttttattt aataatttag atataaaata gaataaaata 11340
aagtgactaa aaattaaaca aatacccttt aagaaattaa aaaaactaag gaaacatttt 11400
tcttgtttcg agtagataat gccagcctgt taaacgccgt cgacgagtct aacggacacc 11460
aaccagcgaa ccagcagcgt cgcgtcgggc caagcgaagc agacggcacg gcatctctgt 11520
cgctgcctct ggacccctct cgagagttcc gctccaccgt tggacttgct ccgctgtcgg 11580
catccagaaa ttgcgtggcg gagcggcaga cgtgagccgg cacggcaggc ggcctcctcc 11640
tcctctcacg gcaccggcag ctacggggga ttcctttccc accgctcctt cgctttccct 11700
tcctcgcccg ccgtaataaa tagacacccc ctccacaccc tctttcccca acctcgtgtt 11760
gttcggagcg cacacacaca caaccagatc tcccccaaat ccacccgtcg gcacctccgc 11820
ttcaaggtac gccgctcgtc ctcccccccc cccctctcta ccttctctag atcggcgttc 11880
cggtccatgg ttagggcccg gtagttctac ttctgttcat gtttgtgtta gatccgtgtt 11940
tgtgttagat ccgtgctgct agcgttcgta cacggatgcg acctgtacgt cagacacgtt 12000
ctgattgcta acttgccagt gtttctcttt ggggaatcct gggatggctc tagccgttcc 12060
gcagacggga tcgatttcat gatttttttt gtttcgttgc atagggtttg gtttgccctt 12120
ttcctttatt tcaatatatg ccgtgcactt gtttgtcggg tcatcttttc atgctttttt 12180
ttgtcttggt tgtgatgatg tggtctggtt gggcggtcgt tctagatcgg agtagaattc 12240
tgtttcaaac tacctggtgg atttattaat tttggatctg tatgtgtgtg ccatacatat 12300
tcatagttac gaattgaaga tgatggatgg aaatatcgat ctaggatagg tatacatgtt 12360
gatgcgggtt ttactgatgc atatacagag atgctttttg ttcgcttggt tgtgatgatg 12420
tggtgtggtt gggcggtcgt tcattcgttc tagatcggag tagaatactg tttcaaacta 12480
cctggtgtat ttattaattt tggaactgta tgtgtgtgtc atacatcttc atagttacga 12540
gtttaagatg gatggaaata tcgatctagg ataggtatac atgttgatgt gggttttact 12600
gatgcatata catgatggca tatgcagcat ctattcatat gctctaacct tgagtaccta 12660

tctattataa taaacaagta tgttttataa ttattttgat cttgatatac ttggatgatg 12720
gcatatgcag cagctatatg tggatttttt tagccctgcc ttcatacgct atttatttgc 12780
ttggtactgt ttcttttgtc gatgctcacc ctgttgtttg gtgttacttc tgcagggatc 12840
cggcagcagc catgcagaag ctgatcaaca gcgtgcagaa ctacgcctgg ggcagcaaga 12900
ccgccctgac cgagctgtac ggcatggaga accccagcag ccagcccatg gccgagctgt 12960
ggatgggcgc ccaccccaag agcagcagcc gcgtgcagaa cgccgccggc gacatcgtga 13020
gcctgcgcga cgtgatcgag agcgacaaga gcaccctgct gggcgaggcc gtggccaagc 13080
gcttcggcga gctgcccttc ctgttcaagg tgctgtgcgc cgcccagccc ctgagcatcc 13140
aggtgcaccc caacaagcac aacagcgaga tcggcttcgc caaggagaac gccgccggca 13200
tccccatgga cgccgccgag cgcaactaca aggaccccaa ccacaagccc gagctggtgt 13260
tcgccctgac ccccttcctg gccatgaacg ccttccgcga gttcagcgag atcgtgagcc 13320
tgctgcagcc cgtggccggc gcccaccccg ccatcgccca cttcctgcag cagcccgacg 13380
ccgagcgcct gagcgagctg ttcgccagcc tgctgaacat gcagggcgag gagaagagcc 13440
gcgccctggc catcctgaag agcgccctgg acagccagca gggcgagccc tggcagacca 13500
tccgcctgat cagcgagttc taccccgagg acagcggcct gttcagcccc ctgctgctga 13560
acgtggtgaa gctgaacccc ggcgaggcca tgttcctgtt cgccgagacc ccccacgcct 13620
acctgcaggg cgtggccctg gaggtgatgg ccaacagcga caacgtgctg cgcgccggcc 13680
tgaccccaa gtacatcgac atccccgagc tggtggccaa cgtgaagttc gaggccaagc 13740
ccgccaacca gctgctgacc cagcccgtga agcagggcgc cgagctggac ttccccatcc 13800
ccgtggacga cttcgccttc agcctgcacg acctgagcga caaggagacc accatcagcc 13860
agcagagcgc cgccatcctg ttctgcgtgg agggcgacgc caccctgtgg aagggcagcc 13920
agcagctgca gctgaagccc ggcgagagcg ccttcatcgc cgccaacgag agccccgtga 13980
ccgtgaaggg ccacggccgc ctggcccgcg tgtacaacaa gctgtgatag gagctcgatc 14040
cgtcgacctg cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt 14100
gccggtcttg cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt 14160
aacatgtaat gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta 14220
tacatttaat acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc 14280
gcggtgtcat ctatgttact agatcggcgc gccgcaattg aagtttgggc ggccagcatg 14340
gccgtatccg caatgtgtta ttaagttgtc taagcgtcaa tttgtttaca ccacaatata 14400
tcctgccacc agccagccaa cagctccccg accggcagct cggcacaaaa tcaccactcg 14460
atacaggcag cccatcagaa ttaattctca tgtttgacag cttatcatcg actgcacggt 14520
gcaccaatgc ttctggcgtc aggcagccat cggaagctgt ggtatggctg tgcaggtcgt 14580
aaatcactgc ataattcgtg tcgctcaagg cgcactcccg ttctggataa tgtttttgc 14640
gccgacatca taacggttct ggcaaatatt ctgaaatgag ctgttgacaa ttaatcatcc 14700

ggctcgtata atgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagaccatg 14760
agggaagcgt tgatcgccga agtatcgact caactatcag aggtagttgg cgtcatcgag 14820
cgccatctcg aaccgacgtt gctggccgta catttgtacg gctccgcagt ggatggcggc 14880
ctgaagccac acagtgatat tgatttgctg gttacggtga ccgtaaggct tgatgaaaca 14940
acgcggcgag ctttgatcaa cgacctttg gaaacttcgg cttcccctgg agagagcgag 15000
attctccgcg ctgtagaagt caccattgtt gtgcacgacg acatcattcc gtggcgttat 15060
ccagctaagc gcgaactgca atttggagaa tggcagcgca atgacattct tgcaggtatc 15120
ttcgagccag ccacgatcga cattgatctg gctatcttgc tgacaaaagc aagagaacat 15180
agcgttgcct tggtaggtcc agcggcggag gaactctttg atccggttcc tgaacaggat 15240
ctatttgagg cgctaaatga aaccttaacg ctatggaact cgccgcccga ctgggctggc 15300
gatgagcgaa atgtagtgct tacgttgtcc cgcatttggt acagcgcagt aaccggcaaa 15360
atcgcgccga aggatgtcgc tgccgactgg gcaatggagc gcctgccggc ccagtatcag 15420
cccgtcatac ttgaagctag gcaggcttat cttggacaag aagatcgctt ggcctcgcgc 15480
gcagatcagt tggaagaatt tgttcactac gtgaaaggcg agatcaccaa agtagtcggc 15540
aaataaagct ctagtggatc tccgtacccg gggatctggc tcgcggcgga cgcacgacgc 15600
cggggcgaga ccataggcga tctcctaaat caatagtagc tgtaacctcg aagcgtttca 15660
cttgtaacaa cgattgagaa tttttgtcat aaaattgaaa tacttggttc gcatttttgt 15720
catccgcggt cagccgcaat tctgacgaac tgcccattta gctggagatg attgtacatc 15780
cttcacgtga aaatttctca agcgctgtga acaagggttc agattttaga ttgaaaggtg 15840
agccgttgaa acacgttctt cttgtcgatg acgacgtcgc tatgcggcat cttattattg 15900
aataccttac gatccacgcc ttcaaagtga ccgcggtagc cgacagcacc cagttcacaa 15960
gagtactctc ttccgcgacg gtcgatgtcg tggttgttga tctagattta ggtcgtgaag 16020
atgggctcga gatcgttcgt aatctggcgg caaagtctga tattccaatc ataattatca 16080
gtggcgaccg ccttgaggag acggataaag ttgttgcact cgagctagga gcaagtgatt 16140
ttatcgctaa gccgttcagt atcagagagt ttctagcacg cattcgggtt gccttgcgcg 16200
tgcgccccaa cgttgtccgc tccaaagacc gacggtcttt ttgttttact gactggacac 16260
ttaatctcag gcaacgtcgc ttgatgtccg aagctggcgg tgaggtgaaa cttacggcag 16320
gtgagttcaa tcttctcctc gcgtttttag agaaaccccg cgacgttcta tcgcgcgagc 16380
aacttctcat tgccagtcga gtacgcgacg aggaggttta tgacaggagt atagatgttc 16440
tcattttgag gctgcgccgc aaacttgagg cagatccgtc aagccctcaa ctgataaaaa 16500
cagcaagagg tgccggttat ttctttgacg cggacgtgca ggtttcgcac ggggggacga 16560
tggcagcctg agccaattcc cagatccccg aggaatcggc gtgagcggtc gcaaaccatc 16620
cggcccggta caaatcggcg cggcgctggg tgatgacctg gtggagaagt tgaaggccgc 16680
gcaggccgcc cagcggcaac gcatcgaggc agaagcacgc cccggtgaat cgtggcaagc 16740

ggccgctgat cgaatccgca aagaatcccg gcaaccgccg gcagccggtg cgccgtcgat 16800
taggaagccg cccaagggcg acgagcaacc agatttttc gttccgatgc tctatgacgt 16860
gggcacccgc gatagtcgca gcatcatgga cgtggccgtt ttccgtctgt cgaagcgtga 16920
ccgacgagct ggcgaggtga tccgctacga gcttccagac gggcacgtag aggtttccgc 16980
agggccggcc ggcatggcca gtgtgtggga ttacgacctg gtactgatgg cggtttccca 17040
tctaaccgaa tccatgaacc gataccggga agggaaggga gacaagcccg gccgcgtgtt 17100
ccgtccacac gttgcggacg tactcaagtt ctgccggcga gccgatggcg gaaagcagaa 17160
agacgacctg gtagaaacct gcattcggtt aaacaccacg cacgttgcca tgcagcgtac 17220
gaagaaggcc aagaacggcc gcctggtgac ggtatccgag ggtgaagcct tgattagccg 17280
ctacaagatc gtaaagagcg aaaccgggcg gccggagtac atcgagatcg agctagctga 17340
ttggatgtac cgcgagatca cagaaggcaa gaacccggac gtgctgacgg ttcaccccga 17400
ttacttttg atcgatcccg gcatcggccg ttttctctac cgcctggcac gccgcgccgc 17460
aggcaaggca gaagccagat ggttgttcaa gacgatctac gaacgcagtg gcagcgccgg 17520
agagttcaag aagttctgtt tcaccgtgcg caagctgatc gggtcaaatg acctgccgga 17580
gtacgatttg aaggaggagg cggggcaggc tggcccgatc ctagtcatgc gctaccgcaa 17640
cctgatcgag ggcgaagcat ccgccggttc ctaatgtacg gagcagatgc tagggcaaat 17700
tgccctagca ggggaaaaag gtcgaaaagg tctctttcct gtggatagca cgtacattgg 17760
gaacccaaag ccgtacattg ggaaccggaa cccgtacatt gggaacccaa agccgtacat 17820
tgggaaccgg tcacacatgt aagtgactga tataaaagag aaaaaaggcg atttttccgc 17880
ctaaaactct ttaaaactta ttaaaactct taaaacccgc ctggcctgtg cataactgtc 17940
tggccagcgc acagccgaag agctgcaaaa agcgcctacc cttcggtcgc tgcgctccct 18000
acgccccgcc gcttcgcgtc ggcctatcgc ggccgctggc cgctcaaaaa tggctggcct 18060
acggccaggc aatctaccag ggcgcggaca agccgcgccg tcgccactcg accgccggcg 18120
ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa tcgccccatc 18180
atccagccag aaagtgaggg agccacggtt gatgagagct ttgttgtagg tggaccagtt 18240
ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 18300
ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgccgtc ccgtcaagtc 18360
agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg 18420
agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa 18480
agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 18540
tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttcccctcg 18600
tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 18660
ggcaaaagct ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg 18720
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc 18780

ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg 18840
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct 18900
ggcgtttttc cataggctcc gccccctga cgagcatcac aaaaatcgac gctcaagtca 18960
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct 19020
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc 19080
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt 19140
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc 19200
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc 19260
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg 19320
gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc tgctgaagcc 19380
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag 19440
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga 19500
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat 19560
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttgatcc ggaatta 19617

<210> 6
<211> 55
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> trình tự đích đối với TALEN của 22808
<400> 6
tccagggtca acgtggagac agggaggtac gaaccggtga ctggcgaagg aagca 55

<210> 7
<211> 15722
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 23123

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)

<223> bNRB-05

<220>

<221> vùng khởi động

<222> (304)..(2100)

<223> prSoUbi4-04

<220>

<221> gen

<222> (2117)..(6286)

<223> cCas9-01

<220>

<221> đột biến

<222> (5606)..(5608)

<223> đột biến L thành V

<220>

<221> đột biến

<222> (5651)..(5653)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc

<222> (6292)..(6544)

<223> tNOS-05-01

<220>

<221> vùng khởi động

<222> (6551)..(6925)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(6946)

<223> xZmPLAIIA

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(7031)

<223> rsgRNAZmPLAIIA02

<220>

<221> vùng khởi động

<222> (7042)..(9033)

<223> prUbi1-04

<220>

<221> gen

<222> (9050)..(10228)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (10251)..(10503)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp

<222> (10547)..(10676)

<223> bNLB-03

<220>

<221> gen

<222> (10956)..(11744)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (11839)..(11969)

<223> prVirG-01

<220>

<221> gen

<222> (12706)..(13779)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (13822)..(14226)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14904)..(15710)

<223> oCOLE-06

<400> 7

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560

acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620

gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680

tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740

ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800

ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860

gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920

ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980
ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040
tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100
gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160
cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220
gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280
cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340
gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400
cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460
gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520
ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580
cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640
cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700
gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760
gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820
cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880
cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940
ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000
cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060
cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120
gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240
catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300
ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360
cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420
caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480
ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540
gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600
gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660
cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720
gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780
caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840
cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960

cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020
gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080
gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200
gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260
ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380
gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440
gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500
cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560
cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620
ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680
ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740
ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800
gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860
cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920
ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980
gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040
gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100
cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160
gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220
gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280
cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340
caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400
gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460
cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520
cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580
gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640
gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700
cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760
cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820
cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880
gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940
gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000

caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060
gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120
gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180
gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240
gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300
aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360
atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420
tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480
aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540
gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600
taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660
acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720
acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcagggt caacgtggag acaggggttt tagagctaga 6960
aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020
gctttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080
ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140
gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttttagt 7440
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500
ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa 7560
taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680
cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740
agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860
acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920
gacacccccc ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980
accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040

ccccccccccc cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100
agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160
cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220
ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280
tttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340
gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400
gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460
ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520
atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580
atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640
attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700
gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760
gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820
tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880
ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940
gatttttttа gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000
tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060
gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120
catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180
cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240
cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300
gttcaaggtg ctgtgcgccg cccagcccct gagcatccag gtgcacccca acaagcacaa 9360
cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420
caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480
catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540
ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600
cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660
cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720
ccccgaggac agcggcctgt tcagcccccт gctgctgaac gtggtgaagc tgaaccccgg 9780
cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840
ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900
ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960
gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020
cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080

ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140
cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200
ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500
atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560
aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620
gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680
aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740
gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800
gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860
caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920
gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980
tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040
tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100
atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160
acctttttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220
ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280
ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340
ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400
cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460
ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520
cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580
ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640
aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700
ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760
cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820
tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880
tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940
tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000
cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060
tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120

caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180
cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240
tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300
gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360
gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420
ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480
cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540
tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600
aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740
cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860
attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980
tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100
tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160

ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220
cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280
ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340
cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400
gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460
aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640
gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760
agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240
gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc 15300
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720
ta 15722

<210> 8
<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> để chỉnh sửa MTL

<400> 8

gggtcaacgt ggagacaggg 20

<210> 9

<211> 41

<212> ADN

<213> Zea mays

<400> 9

agggtcaacg tggagacagg gaggtacgaa ccggtgactg g 41

<210> 10

<211> 45

<212> ADN

<213> Zea mays

<400> 10

agggtcaacg tggagacagg cgaggaggta cgaaccggtg actgg 45

<210> 11

<211> 42

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> MTL đột biến

<400> 11

agggtcaacg tggagacaag ggaggtacga accggtgact gg 42

<210> 12

<211> 28

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> phần MTL đột biến

<400> 12

agggtcaacg tggagaaccg gtgactgg 28

<210> 13
<211> 40
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> phần MTL đột biến
<400> 13
agggtcaacg tggagacggg aggtacgaac cggtgactgg 40

<210> 14
<211> 28
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> phần MTL đột biến
<400> 14
agggtcaacg tggagaaccg gtgactgg 28

<210> 15
<211> 42
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> phần MTL đột biến
<400> 15
agggtcaacg tggagacaag ggaggtacga accggtgact gg 42

<210> 16
<211> 40
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> phần MTL đột biến

<400> 16

agggtcaacg tggagacggg aggtacgaac cggtgactgg 40

<210> 17

<211> 41

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> phần MTL không đột biến

<400> 17

agggtcaacg tggagacagg gaggtacgaa ccggtgactg g 41

<210> 18

<211> 28

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> phần MTL đột biến

<400> 18

agggtcaacg tggagaaccg gtgactgg 28

<210> 19

<211> 1371

<212> ADN

<213> Zea mays

<400> 19

agttcatcac taatcacact tattgttccc tcgacgagta tctagctagc tcattaatcg 60

atcaatcggg gtgtgcggtc gaaggcggca atggcgagct actcgtcgcg gcgtccatgc 120

aatacctgta gcacgaaggc gatggccggg agcgtggtcg gcgagcccgt cgtgctgggg 180

cagagggtga cggtgctgac ggtggacggc ggcggcgtcc ggggtctcat cccgggaacc 240

atcctcgcct tcctcgaggc caggctgcag gagctggacg gaccggaggc gaggctggcg 300

gactacttcg actacatcgc cggaaccagc accggcggtc tcatcaccgc catgctcacc 360

gcgcccggca aggacaagcg gcctctctac gctgccaagg acatcaacca cttttacatg 420

gagaactgcc cgcgcatctt ccctcagaag agcaggcttg cggccgccat gtccgcgctg 480

aggaagccaa agtacaacgg caagtgcatg cgcagcctga ttaggagcat cctcggcgag 540
acgagggcca agagcacgcc tctgaagaac gctctgctct cggacgtgtg cattggcacg 600
tccgccgcgc cgacctacct cccggcgcac tacttccaga ctgaagacgc caacggcaag 660
gagcgcgaat acaacctcat cgacggcggt gtggcggcca acaacccgac gatggttgcg 720
atgacgcaga tcaccaaaaa gatgcttgcc agcaaggaca aggccgagga gctgtaccca 780
gtgaagccgt cgaactgccg caggttcctg gtgctgtcca tcgggacggg gtcgacgtcc 840
gagcagggcc tctacacggc gcggcagtgc tcccggtggg gtatctgccg gtggctccgc 900
aacaacggca tggcccccat catcgacatc ttcatggcgg ccagctcgga cctggtggac 960
atccacgtcg ccgcgatgtt ccagtcgctc cacagcgacg gcgactacct gcgcatccag 1020
gacaactcgc tccgtggcgc cgcggccacc gtggacgcgg cgacgccgga gaacatgcgg 1080
acgctcgtcg ggatcgggga gcggatgctg gcacagaggg tgtccagggt caacgtggag 1140
acagggaggt acgaaccggt gactggcgaa ggaagcaatg ccgatgccct cggtgggctc 1200
gctaggcagc tctccgagga gaggagaaca aggctcgcgc gccgcgtgtc tgccatcaac 1260
ccaagaggct ctagatgtgc gtcgtacgat atctaagaca agtggcttta ctgtcagtca 1320
catgcttgta aataagtaga ctttatttta ataaaacata aaaatatata t 1371

<210> 20
<211> 15722
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 23397

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (304)..(2100)
<223> prSoUbi4-04
<220>
<221> gen
<222> (2117)..(6286)

<223> cCas9-01

<220>

<221> đột biến

<222> (5606)..(5608)

<223> đột biến L thành V

<220>

<221> đột biến

<222> (5651)..(5653)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc

<222> (6292)..(6544)

<223> tNOS-05-01

<220>

<221> vùng khởi động

<222> (6551)..(6925)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(6946)

<223> xZmVLHP

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(7031)

<223> rsgRNAZmVLHP-02

<220>

<221> vùng khởi động

<222> (7042)..(9033)

<223> prUbi1-04

<220>

<221> gen

<222> (9050)..(10228)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (10251)..(10503)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp

<222> (10547)..(10676)

<223> bNLB-03

<220>

<221> gen

<222> (10956)..(11744)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (11839)..(11969)

<223> prVirG-01

<220>

<221> gen

<222> (12706)..(13779)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (13822)..(14226)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14904)..(15710)

<223> oCOLE-06

<400> 20

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420
gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480
atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540
gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600
acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660
ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720
gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780
tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840
gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900
ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960
gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020
gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080
gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140
ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200
tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260
ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320
ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380
cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440
tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500
tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560
acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620
gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680
tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740
ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800
ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860
gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920
ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980
ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040
tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100
gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160
cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220
gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280
cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340
gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400

cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460
gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520
ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580
cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640
cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700
gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760
gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820
cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880
cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940
ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000
cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060
cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120
gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240
catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300
ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360
cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420
caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480
ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540
gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600
gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660
cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720
gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780
caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840
cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960
cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020
gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080
gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200
gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260
ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380
gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440

gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500

cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560

cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620

ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680

ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740

ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800

gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860

cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920

ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980

gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040

gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100

cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160

gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220

gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280

cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340

caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400

gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460

cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520

cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580

gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640

gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700

cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760

cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820

cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880

gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940

gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000

caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060

gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120

gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180

gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240

gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300

aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360

atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420

tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480

aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540
gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600
taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660
acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720
acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcagctg gagctgagct tccggggttt tagagctaga 6960
aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020
gctttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080
ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140
gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttttagt 7440
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500
ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa 7560
taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680
cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740
agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860
acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920
gacacccccct ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980
accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040
ccccccccccc cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100
agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160
cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220
ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280
tttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340
gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400
gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460
ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520

atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580
atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640
attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700
gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760
gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820
tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880
ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940
gatttttta gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000
tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060
gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120
catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180
cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240
cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300
gttcaaggtg ctgtgcgccg cccagccccc gagcatccag gtgcacccca acaagcacaa 9360
cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420
caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480
catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540
ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600
cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660
cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720
ccccgaggac agcggcctgt tcagcccccc gctgctgaac gtggtgaagc tgaaccccgg 9780
cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840
ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900
ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960
gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020
cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080
ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140
cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200
ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260
acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320
tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380
ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440
acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500
atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560

aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620

gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680

aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740

gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800

gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860

caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920

gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980

tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040

tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100

atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160

acctttttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220

ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280

ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340

ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400

cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460

ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520

cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580

ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640

aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700

ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760

cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820

tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880

tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940

tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000

cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060

tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120

caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180

cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240

tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300

gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360

gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420

ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480

cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540

tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600

aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740
cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860
attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980
tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100
tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160
ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220
cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280
ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340
cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400
gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460
aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640

gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760
agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240
gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc 15300
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720
ta 15722

<210> 21
<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> để chỉnh sửa VLHP2
<400> 21
gctggagctg agcttccggg 20

<210> 22
<211> 15722
<212> ADN
<213> Trình tự Nhân tạo
<220>

<223> vecto 23398

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (304)..(2100)
<223> prSoUbi4-04
<220>
<221> gen
<222> (2117)..(6286)
<223> cCas9-01
<220>
<221> đặc điểm_hỗn tạp
<222> (5606)..(5608)
<223> đột biến L thành V
<220>
<221> đặc điểm_hỗn tạp
<222> (5651)..(5653)
<223> đột biến I thành V
<220>
<221> trình tự kết thúc
<222> (6292)..(6544)
<223> tNOS-05-01
<220>
<221> vùng khởi động
<222> (6551)..(6925)
<223> prOsU3-01
<220>
<221> đặc điểm_hỗn tạp
<222> (6927)..(6946)
<223> xZmGW2

<220>
<221> đặc điểm_hỗn tạp
<222> (6927)..(7031)
<223> rsgRNAZmGW2-01
<220>
<221> vùng khởi động
<222> (7042)..(9033)
<223> prUbi1-04
<220>
<221> gen
<222> (9050)..(10228)
<223> cPMI-09
<220>
<221> trình tự kết thúc
<222> (10251)..(10503)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp
<222> (10547)..(10676)
<223> bNLB-03
<220>
<221> gen
<222> (10956)..(11744)
<223> cSpec-03
<220>
<221> vùng khởi động
<222> (11839)..(11969)
<223> prVirG-01
<220>
<221> gen
<222> (12706)..(13779)
<223> cRepA-01
<220>
<221> đặc điểm_hỗn tạp

<222> (13822)..(14226)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14904)..(15710)

<223> oCOLE-06

<400> 22

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560

acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620

gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680
tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740
ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800
ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860
gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920
ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980
ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040
tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100
gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160
cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220
gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280
cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340
gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400
cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460
gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520
ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580
cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640
cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700
gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760
gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820
cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880
cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940
ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000
cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060
cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120
gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240
catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300
ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360
cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420
caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480
ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540
gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600
gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660

cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720

gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780

caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840

cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900

cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960

cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020

gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080

gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140

cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200

gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260

ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320

gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380

gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440

gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500

cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560

cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620

ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680

ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740

ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800

gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860

cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920

ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980

gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040

gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100

cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160

gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220

gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280

cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340

caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400

gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460

cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520

cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580

gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640

gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700

cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760
cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820
cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880
gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940
gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000
caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060
gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120
gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180
gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240
gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300
aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360
atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420
tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480
aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540
gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600
taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660
acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720
acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcagagc ggttcacgcg gccgcagttt tagagctaga 6960
aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020
gctttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080
ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140
gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200
taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260
catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320
tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380
ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttagt 7440
acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500
ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa 7560
taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620
cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680
cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740

agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800
gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860
acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920
gacacccccc ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980
accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040
ccccccccc cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100
agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160
cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220
ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280
tttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340
gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400
gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460
ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520
atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580
atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640
attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700
gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760
gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820
tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880
ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940
gatttttta gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000
tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060
gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120
catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180
cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240
cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300
gttcaaggtg ctgtgcgccg cccagccccc gagcatccag gtgcacccca acaagcacaa 9360
cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420
caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480
catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540
ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600
cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660
cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720
ccccgaggac agcggcctgt tcagcccccc gctgctgaac gtggtgaagc tgaaccccgg 9780

cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840

ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900

ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960

gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020

cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080

ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140

cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200

ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260

acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320

tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380

ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440

acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500

atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560

aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620

gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680

aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740

gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800

gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860

caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920

gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980

tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040

tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100

atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160

acctttttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220

ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280

ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340

ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400

cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460

ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520

cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580

ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640

aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700

ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760

cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820

tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880
tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940
tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000
cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060
tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120
caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180
cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240
tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300
gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360
gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420
ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480
cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540
tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600
aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740
cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800
caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860

attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920
tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980
tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040
ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100
tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160
ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220
cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280
ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340
cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400
gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460
aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520
catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580
gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640
gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700
cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760
agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240
gtcgttcgct ccaagctggg ctgtgtgcac gaacccccgc ttcagcccga ccgctgcgcc 15300
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720
ta 15722

<210> 23

<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> để chỉnh sửa GW2-1
<400> 23
gagcggttca cgcggccgca 20

<210> 24
<211> 15721
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 23763

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (304)..(2100)
<223> prSoUbi4-04
<220>
<221> gen
<222> (2117)..(6286)
<223> cCas9-01
<220>
<221> đột biến
<222> (5606)..(5608)
<223> đột biến L thành V
<220>
<221> đột biến
<222> (5651)..(5653)

<223> đột biến I thành V

<220>

<221> đặc điểm_hỗn tạp

<222> (6221)..(6283)

<223> xSV40NLS-03

<220>

<221> trình tự kết thúc

<222> (6292)..(6544)

<223> tNOS-05-01

<220>

<221> vùng khởi động

<222> (6551)..(6925)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(6945)

<223> xTaVLHP1

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(7030)

<223> rsgRNA TaVLHP1-01

<220>

<221> vùng khởi động

<222> (7041)..(9032)

<223> prUbi1-04

<220>

<221> gen

<222> (9049)..(10227)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (10250)..(10502)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp

<222> (10546)..(10675)

<223> bNLB-03

<220>

<221> gen

<222> (10955)..(11743)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (11838)..(11968)

<223> prVirG-01

<220>

<221> gen

<222> (12705)..(13778)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (13821)..(14225)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14903)..(15709)

<223> oCOLE-06

<400> 24

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660
ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720
gaatcctcgc gatcctctca aggtactgcg agttttcgat ccccctctcg acccctcgta 780
tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840
gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900
ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960
gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020
gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080
gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140
ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200
tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260
ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320
ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380
cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440
tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500
tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560
acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620
gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680
tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740
ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800
ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860
gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920
ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980
ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040
tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100
gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160
cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220
gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280
cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340
gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400
cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460
gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520
ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580
cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640

cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700

gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760

gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820

cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880

cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940

ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000

cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060

cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120

gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180

caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240

catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300

ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatccccgc accagatcca 3360

cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420

caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480

ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540

gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600

gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660

cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720

gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780

caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840

cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900

cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960

cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020

gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080

gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140

cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200

gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260

ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320

gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380

gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440

gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500

cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560

cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620

ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680

ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740

ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800

gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860

cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920

ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980

gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040

gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100

cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160

gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220

gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280

cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340

caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400

gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460

cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520

cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580

gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640

gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700

cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760

cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820

cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880

gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940

gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000

caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060

gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120

gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180

gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240

gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300

aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360

atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420

tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480

aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540

gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600

taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660

acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720

acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780
ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840
gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900
aagagttgtg cagatgatcc gtggcagacg agcaggcgca gttccgtttt agagctagaa 6960
atagcaagtt aaaataaggc tagtccgtta tcaacttgaa aaagtggcac cgagtcggtg 7020
ctttttttt cggaccgcgc ctgcagtgca gcgtgacccg gtcgtgcccc tctctagaga 7080
taatgagcat tgcatgtcta agttataaaa aattaccaca tatttttttt gtcacacttg 7140
tttgaagtgc agtttatcta tctttataca tatatttaaa ctttactcta cgaataatat 7200
aatctatagt actacaataa tatcagtgtt ttagagaatc atataaatga acagttagac 7260
atggtctaaa ggacaattga gtattttgac aacaggactc tacagtttta tctttttagt 7320
gtgcatgtgt tctccttttt ttttgcaaat agcttcacct atataatact tcatccattt 7380
tattagtaca tccatttagg gtttagggtt aatggttttt atagactaat ttttttagta 7440
catctatttt attctatttt agcctctaaa ttaagaaaac taaaactcta ttttagtttt 7500
tttatttaat aatttagata taaaatagaa taaaataaag tgactaaaaa ttaaacaaat 7560
acccttaag aaattaaaaa aactaaggaa acatttttct tgtttcgagt agataatgcc 7620
agcctgttaa acgccgtcga cgagtctaac ggacaccaac cagcgaacca gcagcgtcgc 7680
gtcgggccaa gcgaagcaga cggcacggca tctctgtcgc tgcctctgga cccctctcga 7740
gagttccgct ccaccgttgg acttgctccg ctgtcggcat ccagaaattg cgtggcggag 7800
cggcagacgt gagccggcac ggcaggcggc ctcctcctcc tctcacggca ccggcagcta 7860
cgggggattc ctttcccacc gctccttcgc tttcccttcc tcgcccgccg taataaatag 7920
acacccccctc cacaccctct ttccccaacc tcgtgttgtt cggagcgcac acacacacaa 7980
ccagatctcc cccaaatcca cccgtcggca cctccgcttc aaggtacgcc gctcgtcctc 8040
ccccccccc ctctctacct tctctagatc ggcgttccgg tccatggtta gggcccggta 8100
gttctacttc tgttcatgtt tgtgttagat ccgtgtttgt gttagatccg tgctgctagc 8160
gttcgtacac ggatgcgacc tgtacgtcag acacgttctg attgctaact tgccagtgtt 8220
tctctttggg gaatcctggg atggctctag ccgttccgca gacgggatcg atttcatgat 8280
ttttttgtt tcgttgcata gggtttggtt tgccctttc ctttatttca atatatgccg 8340
tgcacttgtt tgtcgggtca tcttttcatg ctttttttg tcttggttgt gatgatgtgg 8400
tctggttggg cggtcgttct agatcggagt agaattctgt ttcaaactac ctggtggatt 8460
tattaatttt ggatctgtat gtgtgtgcca tacatattca tagttacgaa ttgaagatga 8520
tggatggaaa tatcgatcta ggataggtat acatgttgat gcgggtttta ctgatgcata 8580
tacagagatg ctttttgttc gcttggttgt gatgatgtgg tgtggttggg cggtcgttca 8640
ttcgttctag atcggagtag aatactgttt caaactacct ggtgtattta ttaattttgg 8700
aactgtatgt gtgtgtcata catcttcata gttacgagtt taagatggat ggaaatatcg 8760

atctaggata ggtatacatg ttgatgtggg ttttactgat gcatatacat gatggcatat 8820
gcagcatcta ttcatatgct ctaaccttga gtacctatct attataataa acaagtatgt 8880
tttataatta ttttgatctt gatatacttg gatgatggca tatgcagcag ctatatgtgg 8940
attttttag ccctgccttc atacgctatt tatttgcttg gtactgtttc ttttgtcgat 9000
gctcaccctg ttgtttggtg ttacttctgc agggatccgg cagcagccat gcagaagctg 9060
atcaacagcg tgcagaacta cgcctggggc agcaagaccg ccctgaccga gctgtacggc 9120
atggagaacc ccagcagcca gcccatggcc gagctgtgga tgggcgccca ccccaagagc 9180
agcagccgcg tgcagaacgc cgccggcgac atcgtgagcc tgcgcgacgt gatcgagagc 9240
gacaagagca ccctgctggg cgaggccgtg gccaagcgct tcggcgagct gcccttcctg 9300
ttcaaggtgc tgtgcgccgc ccagcccctg agcatccagg tgcaccccaa caagcacaac 9360
agcgagatcg gcttcgccaa ggagaacgcc gccggcatcc ccatggacgc cgccgagcgc 9420
aactacaagg accccaacca caagcccgag ctggtgttcg ccctgacccc cttcctggcc 9480
atgaacgcct tccgcgagtt cagcgagatc gtgagcctgc tgcagcccgt ggccggcgcc 9540
caccccgcca tcgcccactt cctgcagcag cccgacgccg agcgcctgag cgagctgttc 9600
gccagcctgc tgaacatgca gggcgaggag aagagccgcg ccctggccat cctgaagagc 9660
gccctggaca gccagcaggg cgagccctgg cagaccatcc gcctgatcag cgagttctac 9720
cccgaggaca gcggcctgtt cagcccccctg ctgctgaacg tggtgaagct gaaccccggc 9780
gaggccatgt tcctgttcgc cgagaccccc cacgcctacc tgcagggcgt ggccctggag 9840
gtgatggcca acagcgacaa cgtgctgcgc gccggcctga cccccaagta catcgacatc 9900
cccgagctgg tggccaacgt gaagttcgag gccaagcccg ccaaccagct gctgacccag 9960
cccgtgaagc agggcgccga gctggacttc cccatccccg tggacgactt cgccttcagc 10020
ctgcacgacc tgagcgacaa ggagaccacc atcagccagc agagcgccgc catcctgttc 10080
tgcgtggagg gcgacgccac cctgtggaag ggcagccagc agctgcagct gaagcccggc 10140
gagagcgcct tcatcgccgc caacgagagc cccgtgaccg tgaagggcca cggccgcctg 10200
gcccgcgtgt acaacaagct gtgataggag ctcgatccgt cgacctgcag atcgttcaaa 10260
catttggcaa taaagtttct taagattgaa tcctgttgcc ggtcttgcga tgattatcat 10320
ataatttctg ttgaattacg ttaagcatgt aataattaac atgtaatgca tgacgttatt 10380
tatgagatgg gtttttatga ttagagtccc gcaattatac atttaatacg cgatagaaaa 10440
caaaatatag cgcgcaaact aggataaatt atcgcgcgcg gtgtcatcta tgttactaga 10500
tcggcgcgcc gcaattgaag tttgggcggc cagcatggcc gtatccgcaa tgtgttatta 10560
agttgtctaa gcgtcaattt gtttacacca caatatatcc tgccaccagc cagccaacag 10620
ctccccgacc ggcagctcgg cacaaaatca ccactcgata caggcagccc atcagaatta 10680
attctcatgt ttgacagctt atcatcgact gcacggtgca ccaatgcttc tggcgtcagg 10740
cagccatcgg aagctgtggt atggctgtgc aggtcgtaaa tcactgcata attcgtgtcg 10800

ctcaaggcgc actcccgttc tggataatgt tttttgcgcc gacatcataa cggttctggc 10860
aaatattctg aaatgagctg ttgacaatta atcatccggc tcgtataatg tgtggaattg 10920
tgagcggata acaatttcac acaggaaaca gaccatgagg gaagcgttga tcgccgaagt 10980
atcgactcaa ctatcagagg tagttggcgt catcgagcgc catctcgaac cgacgttgct 11040
ggccgtacat ttgtacggct ccgcagtgga tggcggcctg aagccacaca gtgatattga 11100
tttgctggtt acggtgaccg taaggcttga tgaaacaacg cggcgagctt tgatcaacga 11160
ccttttggaa acttcggctt cccctggaga gagcgagatt ctccgcgctg tagaagtcac 11220
cattgttgtg cacgacgaca tcattccgtg gcgttatcca gctaagcgcg aactgcaatt 11280
tggagaatgg cagcgcaatg acattcttgc aggtatcttc gagccagcca cgatcgacat 11340
tgatctggct atcttgctga caaaagcaag agaacatagc gttgccttgg taggtccagc 11400
ggcggaggaa ctctttgatc cggttcctga acaggatcta tttgaggcgc taaatgaaac 11460
cttaacgcta tggaactcgc cgcccgactg ggctggcgat gagcgaaatg tagtgcttac 11520
gttgtcccgc atttggtaca gcgcagtaac cggcaaaatc gcgccgaagg atgtcgctgc 11580
cgactgggca atggagcgcc tgccggccca gtatcagccc gtcatacttg aagctaggca 11640
ggcttatctt ggacaagaag atcgcttggc ctcgcgcgca gatcagttgg aagaatttgt 11700
tcactacgtg aaaggcgaga tcaccaaagt agtcggcaaa taaagctcta gtggatctcc 11760
gtacccgggg atctggctcg cggcggacgc acgacgccgg ggcgagacca taggcgatct 11820
cctaaatcaa tagtagctgt aacctcgaag cgtttcactt gtaacaacga ttgagaattt 11880
ttgtcataaa attgaaatac ttggttcgca tttttgtcat ccgcggtcag ccgcaattct 11940
gacgaactgc ccatttagct ggagatgatt gtacatcctt cacgtgaaaa tttctcaagc 12000
gctgtgaaca agggttcaga ttttagattg aaaggtgagc cgttgaaaca cgttcttctt 12060
gtcgatgacg acgtcgctat gcggcatctt attattgaat accttacgat ccacgccttc 12120
aaagtgaccg cggtagccga cagcacccag ttcacaagag tactctcttc cgcgacggtc 12180
gatgtcgtgg ttgttgatct agatttaggt cgtgaagatg ggctcgagct aggagcaagt 12240
gattttatcg ctaagccgtt cagtatcaga gagtttctag cacgcattcg ggttgccttg 12300
cgcgtgcgcc ccaacgttgt ccgctccaaa gaccgacggt ctttttgttt tactgactgg 12360
acacttaatc tcaggcaacg tcgcttgatg tccgaagctg gcggtgaggt gaaacttacg 12420
gcaggtgagt tcaatcttct cctcgcgttt ttagagaaac cccgcgacgt tctatcgcgc 12480
gagcaacttc tcattgccag tcgagtacgc gacgaggagg tttatgacag gagtatagat 12540
gttctcattt tgaggctgcg ccgcaaactt gaggcagatc cgtcaagccc tcaactgata 12600
aaaacagcaa gaggtgccgg ttatttcttt gacgcggacg tgcaggtttc gcacgggggg 12660
acgatggcag cctgagccaa ttcccagatc cccgaggaat cggcgtgagc ggtcgcaaac 12720
catccggccc ggtacaaatc ggcgcggcgc tgggtgatga cctggtggag aagttgaagg 12780
ccgcgcaggc cgcccagcgg caacgcatcg aggcagaagc acgccccggt gaatcgtggc 12840

aagcggccgc tgatcgaatc cgcaaagaat cccggcaacc gccggcagcc ggtgcgccgt 12900

cgattaggaa gccgcccaag ggcgacgagc aaccagattt tttcgttccg atgctctatg 12960

acgtgggcac ccgcgatagt cgcagcatca tggacgtggc cgttttccgt ctgtcgaagc 13020

gtgaccgacg agctggcgag gtgatccgct acgagcttcc agacgggcac gtagaggttt 13080

ccgcagggcc ggccggcatg gccagtgtgt gggattacga cctggtactg atggcggttt 13140

cccatctaac cgaatccatg aaccgatacc gggaagggaa gggagacaag cccggccgcg 13200

tgttccgtcc acacgttgcg gacgtactca agttctgccg gcgagccgat ggcggaaagc 13260

agaaagacga cctggtagaa acctgcattc ggttaaacac cacgcacgtt gccatgcagc 13320

gtacgaagaa ggccaagaac ggccgcctgg tgacggtatc cgagggtgaa gccttgatta 13380

gccgctacaa gatcgtaaag agcgaaaccg ggcggccgga gtacatcgag atcgagctag 13440

ctgattggat gtaccgcgag atcacagaag gcaagaaccc ggacgtgctg acggttcacc 13500

ccgattactt tttgatcgat cccggcatcg gccgttttct ctaccgcctg gcacgccgcg 13560

ccgcaggcaa ggcagaagcc agatggttgt tcaagacgat ctacgaacgc agtggcagcg 13620

ccggagagtt caagaagttc tgtttcaccg tgcgcaagct gatcgggtca aatgacctgc 13680

cggagtacga tttgaaggag gaggcggggc aggctggccc gatcctagtc atgcgctacc 13740

gcaacctgat cgagggcgaa gcatccgccg gttcctaatg tacggagcag atgctagggc 13800

aaattgccct agcaggggaa aaaggtcgaa aaggtctctt tcctgtggat agcacgtaca 13860

ttgggaaccc aaagccgtac attgggaacc ggaacccgta cattgggaac ccaaagccgt 13920

acattgggaa ccggtcacac atgtaagtga ctgatataaa agagaaaaaa ggcgattttt 13980

ccgcctaaaa ctctttaaaa cttattaaaa ctcttaaaac ccgcctggcc tgtgcataac 14040

tgtctggcca gcgcacagcc gaagagctgc aaaaagcgcc taccgttcgg tcgctgcgct 14100

ccctacgccc cgccgcttcg cgtcggccta tcgcggccgc tggccgctca aaaatggctg 14160

gcctacggcc aggcaatcta ccagggcgcg gacaagccgc gccgtcgcca ctcgaccgcc 14220

ggcgctgagg tctgcctcgt gaagaaggtg ttgctgactc ataccaggcc tgaatcgccc 14280

catcatccag ccagaaagtg agggagccac ggttgatgag agctttgttg taggtggacc 14340

agttggtgat tttgaacttt tgctttgcca cggaacggtc tgcgttgtcg ggaagatgcg 14400

tgatctgatc cttcaactca gcaaaagttc gatttattca acaaagccgc cgtcccgtca 14460

agtcagcgta atgctctgcc agtgttacaa ccaattaacc aattctgatt agaaaaactc 14520

atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac catatttttg 14580

aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatggcaag 14640

atcctggtat cggtctgcga ttccgactcg tccaacatca atacaaccta ttaatttccc 14700

ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg aatccggtga 14760

gaatggcaaa agctctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta 14820

ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc 14880

gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg 14940
caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt 15000
tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa 15060
gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct 15120
ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc 15180
cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg 15240
tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct 15300
tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag 15360
cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga 15420
agtggtggcc taactacggc tacactagaa gaacagtatt tggtatctgc gctctgctga 15480
agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg 15540
gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag 15600
aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag 15660
ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttttg atccggaatt 15720
a 15721

<210> 25
<211> 19
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> để chỉnh sửa VLHP1 ở lúa mỳ
<400> 25
gacgagcagg cgcagttcc 19

<210> 26
<211> 20
<212> ADN
<213> Triticum aestivum
<400> 26
gctggagctg agcttccggg 20

<210> 27
<211> 20

<212> ADN

<213> Triticum aestivum

<400> 27

tctggagctg agcttccggg 20

<210> 28

<211> 20

<212> ADN

<213> Zea mays

<400> 28

aggcgtcgag cagcgaggtg 20

<210> 29

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> phần ZmVLHP-03 đã chỉnh sửa

<400> 29

aggcgttgag cagcgaggtg 20

<210> 30

<211> 75

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> khuôn thể cho sửa chữa để tạo ra đột biến E149L ở ZmPYL-D

<400> 30

ccttggtgtt gccgtcgggg acgtcgacga cgaatgacag gatgacgagc gtccctggcc 60

ggccgtcgat gacct 75

<210> 31

<211> 15722

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> vecto 23136

<220>

<221> đặc điểm_hỗn tạp

<222> (4)..(259)

<223> bNRB-05

<220>

<221> vùng khởi động

<222> (304)..(2100)

<223> prSoUbi4-04

<220>

<221> gen

<222> (2117)..(6286)

<223> cCas9-01

<220>

<221> đột biến

<222> (5606)..(5608)

<223> đột biến L thành V

<220>

<221> đột biến

<222> (5651)..(5653)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc

<222> (6292)..(6544)

<223> tNOS-05-01

<220>

<221> vùng khởi động

<222> (6551)..(6925)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(6946)

<223> xZmPYL-D

<220>

<221> đặc điểm_hỗn tạp

<222> (6927)..(7031)

<223> rsgRBAZmPYLd-02

<220>

<221> vùng khởi động

<222> (7042)..(9033)

<223> prUbi1-04

<220>

<221> gen

<222> (9050)..(10228)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (10251)..(10503)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp

<222> (10547)..(10676)

<223> bNLB-03

<220>

<221> gen

<222> (10956)..(11744)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (11839)..(11969)

<223> prVirG-01

<220>

<221> gen

<222> (12706)..(13779)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (13822)..(14226)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (14904)..(15710)

<223> oCOLE-06

<400> 31

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccattatg tggtctaggt aggttctata tataagaaaa cttgaaatgt tctaaaaaaa 360

aattcaagcc catgcatgat tgaagcaaac ggtatagcaa cggtgttaac ctgatctagt 420

gatctcttgc aatccttaac ggccacctac cgcaggtagc aaacggcgtc cccctcctcg 480

atatctccgc ggcgacctct ggctttttcc gcggaattgc gcggtgggga cggattccac 540

gagaccgcga cgcaaccgcc tctcgccgct gggccccaca ccgctcggtg ccgtagcctc 600

acgggactct ttctccctcc tcccccgtta taaattggct tcatcccctc cttgcctcat 660

ccatccaaat cccagtcccc aatcccatcc cttcgtagga gaaattcatc gaagctaagc 720

gaatcctcgc gatcctctca aggtactgcg agttttcgat cccctctctcg acccctcgta 780

tgtttgtgtt tgtcgtagcg tttgattagg tatgctttcc ctgtttgtgt tcgtcgtagc 840

gtttgattag gtatgctttc cctgttcgtg ttcatcgtag tgtttgatta ggtcgtgtga 900

ggcgatggcc tgctcgcgtc cttcgatctg tagtcgattt gcgggtcgtg gtgtagatct 960

gcgggctgtg atgaagttat ttggtgtgat ctgctcgcct gattctgcgg gttggctcga 1020

gtagatatga tggttggacc ggttggttcg tttaccgcgc tagggttggg ctgggatgat 1080

gttgcatgcg ccgttgcgcg tgatcccgca gcaggacttg cgtttgattg ccagatctcg 1140

ttacgattat gtgatttggt ttggactttt tagatctgta gcttctgctt atgtgccaga 1200

tgcgcctact gctcatatgc ctgatgataa tcataaatgg ctgtggaact aactagttga 1260

ttgcggagtc atgtatcagc tacaggtgta gggactagct acaggtgtag ggacttgcgt 1320

ctaattgttt ggtcctttac tcatgttgca attatgcaat ttagtttaga ttgtttgttc 1380

cactcatcta ggctgtaaaa gggacactgc ttagattgct gtttaatctt tttagtagat 1440

tatattatat tggtaactta ttacccctat tacatgccat acgtgacttc tgctcatgcc 1500

tgatgataat catagatcac tgtggaatta attagttgat tgttgaatca tgtttcatgt 1560
acataccacg gcacaattgc ttagttcctt aacaaatgca aattttactg atccatgtat 1620
gatttgcgtg gttctctaat gtgaaatact atagctactt gttagtaaga atcaggttcg 1680
tatgcttaat gctgtatgtg ccttctgctc atgcctgatg ataatcatat atcactggaa 1740
ttaattagtt gatcgtttaa tcatatatca agtacatacc atgccacaat ttttagtcac 1800
ttaacccatg cagattgaac tggtccctgc atgttttgct aaattgttct attctgatta 1860
gaccatatat catgtatttt tttttggtaa tggttctctt attttaaatg ctatatagtt 1920
ctggtacttg ttagaaagat ctgcttcata gtttagttgc ctatccctcg aattaggatg 1980
ctgagcagct gatcctatag ctttgtttca tgtatcaatt cttttgtgtt caacagtcag 2040
tttttgttag attcattgta acttatggtc gcttactctt ctggtcctca atgcttgcag 2100
gatcgcggcc gctcatatgg acaagaagta cagcatcggc ctggacatcg gcaccaacag 2160
cgtgggctgg gccgtgatca ccgacgagta caaggtgccg agcaagaagt tcaaggtgct 2220
gggcaacacc gacaggcaca gcatcaagaa gaacctgatc ggcgccctgc tgttcgacag 2280
cggcgagacc gccgaggcca ccaggctgaa gaggaccgcc aggaggaggt acaccaggag 2340
gaagaacagg atctgctacc tgcaggagat cttcagcaac gagatggcca aggtggacga 2400
cagcttcttc cacaggctgg aggagagctt cctggtggag gaggacaaga agcacgagag 2460
gcacccgatc ttcggcaaca tcgtggacga ggtggcctac cacgagaagt acccgaccat 2520
ctaccacctg aggaagaagc tggtggacag caccgacaag gccgacctga ggctgatcta 2580
cctggccctg gcccacatga tcaagttcag gggccacttc ctgatcgagg gcgacctgaa 2640
cccggacaac agcgacgtgg acaagctgtt catccagctg gtgcagacct acaaccagct 2700
gttcgaggag aacccgatca acgccagcgg cgtggacgcc aaggccatcc tgagcgccag 2760
gctgagcaag agcaggaggc tggagaacct gatcgcccag ctgccgggcg agaagaagaa 2820
cggcctgttc ggcaacctga tcgccctgag cctgggcctg accccgaact tcaagagcaa 2880
cttcgacctg gccgaggacg ccaagctgca gctgagcaag gacacctacg acgacgacct 2940
ggacaacctg ctggcccaga tcggcgacca gtacgccgac ctgttcctgg ccgccaagaa 3000
cctgagcgac gccatcctgc tgagcgacat cctgagggtg aacaccgaga tcaccaaggc 3060
cccgctgagc gccagcatga tcaagaggta cgacgagcac caccaggacc tgaccctgct 3120
gaaggccctg gtgaggcagc agctgccgga gaagtacaag gagatcttct tcgaccagag 3180
caagaacggc tacgccggct acatcgacgg cggcgccagc caggaggagt tctacaagtt 3240
catcaagccg atcctggaga agatggacgg caccgaggag ctgctggtga agctgaacag 3300
ggaggacctg ctgaggaagc agaggacctt cgacaacggc agcatcccgc accagatcca 3360
cctgggcgag ctgcacgcca tcctgaggag gcaggaggac ttctacccgt tcctgaagga 3420
caacagggag aagatcgaga agatcctgac cttccgcatc ccgtactacg tgggcccgct 3480
ggccaggggc aacagcaggt tcgcctggat gaccaggaag agcgaggaga ccatcacccc 3540

gtggaacttc gaggaggtgg tggacaaggg cgccagcgcc cagagcttca tcgagaggat 3600
gaccaacttc gacaagaacc tgccgaacga gaaggtgctg ccgaagcaca gcctgctgta 3660
cgagtacttc accgtgtaca acgagctgac caaggtgaag tacgtgaccg agggcatgag 3720
gaagccggcc ttcctgagcg gcgagcagaa gaaggccatc gtggacctgc tgttcaagac 3780
caacaggaag gtgaccgtga agcagctgaa ggaggactac ttcaagaaga tcgagtgctt 3840
cgacagcgtg gagatcagcg gcgtggagga caggttcaac gccagcctgg gcacctacca 3900
cgacctgctg aagatcatca aggacaagga cttcctggac aacgaggaga acgaggacat 3960
cctggaggac atcgtgctga ccctgaccct gttcgaggac agggagatga tcgaggagag 4020
gctgaagacc tacgcccacc tgttcgacga caaggtgatg aagcagctga agaggaggag 4080
gtacaccggc tggggcaggc tgagcaggaa gctgatcaac ggcatcaggg acaagcagag 4140
cggcaagacc atcctggact tcctgaagag cgacggcttc gccaacagga acttcatgca 4200
gctgatccac gacgacagcc tgaccttcaa ggaggacatc cagaaggccc aggtgagcgg 4260
ccagggcgac agcctgcacg agcacatcgc caacctggcc ggcagcccgg ccatcaagaa 4320
gggcatcctg cagaccgtga aggtggtgga cgagctggtg aaggtgatgg gcaggcacaa 4380
gccggagaac atcgtgatcg agatggccag ggagaaccag accacccaga agggccagaa 4440
gaacagcagg gagaggatga agaggatcga ggagggcatc aaggagctgg gcagccagat 4500
cctgaaggag cacccggtgg agaacaccca gctgcagaac gagaagctgt acctgtacta 4560
cctgcagaac ggcagggaca tgtacgtgga ccaggagctg gacatcaaca ggctgagcga 4620
ctacgacgtg gaccacatcg tgccgcagag cttcctgaag gacgacagca tcgacaacaa 4680
ggtgctgacc aggagcgaca agaacagggg caagagcgac aacgtgccga gcgaggaggt 4740
ggtgaagaag atgaaaaact actggaggca gctgctgaac gccaagctga tcacccagag 4800
gaagttcgac aacctgacca aggccgagag gggcggcctg agcgagctgg acaaggccgg 4860
cttcattaaa aggcagctgg tggagaccag gcagatcacc aagcacgtgg cccagatcct 4920
ggacagcagg atgaacacca agtacgacga gaacgacaag ctgatcaggg aggtgaaggt 4980
gatcaccctg aagagcaagc tggtgagcga cttcaggaag gacttccagt tctacaaggt 5040
gagggagatc aataattacc accacgccca cgacgcctac ctgaacgccg tggtgggcac 5100
cgccctgatt aaaaagtacc cgaagctgga gagcgagttc gtgtacggcg actacaaggt 5160
gtacgacgtg aggaagatga tcgccaagag cgagcaggag atcggcaagg ccaccgccaa 5220
gtacttcttc tacagcaaca tcatgaactt cttcaagacc gagatcaccc tggccaacgg 5280
cgagatcagg aagaggccgc tgatcgagac caacggcgag accggcgaga tcgtgtggga 5340
caagggcagg gacttcgcca ccgtgaggaa ggtgctgtcc atgccgcagg tgaacatcgt 5400
gaagaagacc gaggtgcaga ccggcggctt cagcaaggag agcatcctgc cgaagaggaa 5460
cagcgacaag ctgatcgcca ggaagaagga ctgggacccg aagaagtacg gcggcttcga 5520
cagcccgacc gtggcctaca gcgtgctggt ggtggccaag gtggagaagg gcaagagcaa 5580

gaagctgaag agcgtgaagg agctggtggg catcaccatc atggagagga gcagcttcga 5640

gaagaaccca gtggacttcc tggaggccaa gggctacaag gaggtgaaga aggacctgat 5700

cattaaactg ccgaagtaca gcctgttcga gctggagaac ggcaggaaga ggatgctggc 5760

cagcgccggc gagctgcaga agggcaacga gctggccctg ccgagcaagt acgtgaactt 5820

cctgtacctg gccagccact acgagaagct gaagggcagc ccggaggaca acgagcagaa 5880

gcagctgttc gtggagcagc acaagcacta cctggacgag atcatcgagc agatcagcga 5940

gttcagcaag agggtgatcc tggccgacgc caacctggac aaggtgctga gcgcctacaa 6000

caagcacagg gacaagccga tcagggagca ggccgagaac atcatccacc tgttcaccct 6060

gaccaacctg ggcgccccgg ccgccttcaa gtacttcgac accaccatcg acaggaagag 6120

gtacaccagc accaaggagg tgctggacgc caccctgatc caccagagca tcaccggcct 6180

gtacgagacc aggatcgacc tgagccagct gggcggcgac agcagcccgc cgaagaagaa 6240

gaggaaggtg agctggaagg acgccagcgg ctggagcagg atgtgaagct tgatcgttca 6300

aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc gatgattatc 6360

atataatttc tgttgaatta cgttaagcat gtaataatta acatgtaatg catgacgtta 6420

tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata cgcgatagaa 6480

aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cggtgtcatc tatgttacta 6540

gatcttcgaa gggatcttta aacatacgaa cagatcactt aaagttcttc tgaagcaact 6600

taaagttatc aggcatgcat ggatcttgga ggaatcagat gtgcagtcag ggaccatagc 6660

acaggacagg cgtcttctac tggtgctacc agcaaatgct ggaagccggg aacactgggt 6720

acgttggaaa ccacgtgatg tggagtaaga taaactgtag gagaaaagca tttcgtagtg 6780

ggccatgaag cctttcagga catgtattgc agtatgggcc ggcccattac gcaattggac 6840

gacaacaaag actagtatta gtaccacctc ggctatccac atagatcaaa gctggtttaa 6900

aagagttgtg cagatgatcc gtggcagtcg gggacgtcga cgacgagttt tagagctaga 6960

aatagcaagt taaaataagg ctagtccgtt atcaacttga aaaagtggca ccgagtcggt 7020

gcttttttt tcggaccgcg cctgcagtgc agcgtgaccc ggtcgtgccc ctctctagag 7080

ataatgagca ttgcatgtct aagttataaa aaattaccac atatttttt tgtcacactt 7140

gtttgaagtg cagtttatct atctttatac atatatttaa actttactct acgaataata 7200

taatctatag tactacaata atatcagtgt tttagagaat catataaatg aacagttaga 7260

catggtctaa aggacaattg agtattttga caacaggact ctacagtttt atctttttag 7320

tgtgcatgtg ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcatccatt 7380

ttattagtac atccatttag ggtttagggt taatggtttt tatagactaa ttttttttagt 7440

acatctattt tattctattt tagcctctaa attaagaaaa ctaaaactct attttagttt 7500

ttttatttaa taatttagat ataaaataga ataaaataaa gtgactaaaa attaaacaaa 7560

taccctttaa gaaattaaaa aaactaagga aacatttttc ttgtttcgag tagataatgc 7620

cagcctgtta aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg 7680

cgtcgggcca agcgaagcag acggcacggc atctctgtcg ctgcctctgg acccctctcg 7740

agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt gcgtggcgga 7800

gcggcagacg tgagccggca cggcaggcgg cctcctcctc ctctcacggc accggcagct 7860

acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata 7920

gacacccccc ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca 7980

accagatctc ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct 8040

ccccccccca cctctctacc ttctctagat cggcgttccg gtccatggtt agggcccggt 8100

agttctactt ctgttcatgt ttgtgttaga tccgtgtttg tgttagatcc gtgctgctag 8160

cgttcgtaca cggatgcgac ctgtacgtca gacacgttct gattgctaac ttgccagtgt 8220

ttctctttgg ggaatcctgg gatggctcta gccgttccgc agacgggatc gatttcatga 8280

tttttttttgt ttcgttgcat agggtttggt ttgccctttt cctttatttc aatatatgcc 8340

gtgcacttgt ttgtcgggtc atcttttcat gctttttttt gtcttggttg tgatgatgtg 8400

gtctggttgg gcggtcgttc tagatcggag tagaattctg tttcaaacta cctggtggat 8460

ttattaattt tggatctgta tgtgtgtgcc atacatattc atagttacga attgaagatg 8520

atggatggaa atatcgatct aggataggta tacatgttga tgcgggtttt actgatgcat 8580

atacagagat gctttttgtt cgcttggttg tgatgatgtg gtgtggttgg gcggtcgttc 8640

attcgttcta gatcggagta gaatactgtt tcaaactacc tggtgtattt attaattttg 8700

gaactgtatg tgtgtgtcat acatcttcat agttacgagt ttaagatgga tggaaatatc 8760

gatctaggat aggtatacat gttgatgtgg gttttactga tgcatataca tgatggcata 8820

tgcagcatct attcatatgc tctaaccttg agtacctatc tattataata aacaagtatg 8880

ttttataatt attttgatct tgatatactt ggatgatggc atatgcagca gctatatgtg 8940

gatttttttа gccctgcctt catacgctat ttatttgctt ggtactgttt cttttgtcga 9000

tgctcaccct gttgtttggt gttacttctg cagggatccg gcagcagcca tgcagaagct 9060

gatcaacagc gtgcagaact acgcctgggg cagcaagacc gccctgaccg agctgtacgg 9120

catggagaac cccagcagcc agcccatggc cgagctgtgg atgggcgccc accccaagag 9180

cagcagccgc gtgcagaacg ccgccggcga catcgtgagc ctgcgcgacg tgatcgagag 9240

cgacaagagc accctgctgg gcgaggccgt ggccaagcgc ttcggcgagc tgcccttcct 9300

gttcaaggtg ctgtgcgccg cccagcccct gagcatccag gtgcacccca acaagcacaa 9360

cagcgagatc ggcttcgcca aggagaacgc cgccggcatc cccatggacg ccgccgagcg 9420

caactacaag gaccccaacc acaagcccga gctggtgttc gccctgaccc ccttcctggc 9480

catgaacgcc ttccgcgagt tcagcgagat cgtgagcctg ctgcagcccg tggccggcgc 9540

ccaccccgcc atcgcccact tcctgcagca gcccgacgcc gagcgcctga gcgagctgtt 9600

cgccagcctg ctgaacatgc agggcgagga gaagagccgc gccctggcca tcctgaagag 9660

cgccctggac agccagcagg gcgagccctg gcagaccatc cgcctgatca gcgagttcta 9720

ccccgaggac agcggcctgt tcagcccccc gctgctgaac gtggtgaagc tgaaccccgg 9780

cgaggccatg ttcctgttcg ccgagacccc ccacgcctac ctgcagggcg tggccctgga 9840

ggtgatggcc aacagcgaca acgtgctgcg cgccggcctg acccccaagt acatcgacat 9900

ccccgagctg gtggccaacg tgaagttcga ggccaagccc gccaaccagc tgctgaccca 9960

gcccgtgaag cagggcgccg agctggactt ccccatcccc gtggacgact tcgccttcag 10020

cctgcacgac ctgagcgaca aggagaccac catcagccag cagagcgccg ccatcctgtt 10080

ctgcgtggag ggcgacgcca ccctgtggaa gggcagccag cagctgcagc tgaagcccgg 10140

cgagagcgcc ttcatcgccg ccaacgagag ccccgtgacc gtgaagggcc acggccgcct 10200

ggcccgcgtg tacaacaagc tgtgatagga gctcgatccg tcgacctgca gatcgttcaa 10260

acatttggca ataaagtttc ttaagattga atcctgttgc cggtcttgcg atgattatca 10320

tataatttct gttgaattac gttaagcatg taataattaa catgtaatgc atgacgttat 10380

ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac gcgatagaaa 10440

acaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct atgttactag 10500

atcggcgcgc cgcaattgaa gtttgggcgg ccagcatggc cgtatccgca atgtgttatt 10560

aagttgtcta agcgtcaatt tgtttacacc acaatatatc ctgccaccag ccagccaaca 10620

gctccccgac cggcagctcg gcacaaaatc accactcgat acaggcagcc catcagaatt 10680

aattctcatg tttgacagct tatcatcgac tgcacggtgc accaatgctt ctggcgtcag 10740

gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc 10800

gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg 10860

caaatattct gaaatgagct gttgacaatt aatcatccgg ctcgtataat gtgtggaatt 10920

gtgagcggat aacaatttca cacaggaaac agaccatgag ggaagcgttg atcgccgaag 10980

tatcgactca actatcagag gtagttggcg tcatcgagcg ccatctcgaa ccgacgttgc 11040

tggccgtaca tttgtacggc tccgcagtgg atggcggcct gaagccacac agtgatattg 11100

atttgctggt tacggtgacc gtaaggcttg atgaaacaac gcggcgagct ttgatcaacg 11160

acctttttgga aacttcggct tcccctggag agagcgagat tctccgcgct gtagaagtca 11220

ccattgttgt gcacgacgac atcattccgt ggcgttatcc agctaagcgc gaactgcaat 11280

ttggagaatg gcagcgcaat gacattcttg caggtatctt cgagccagcc acgatcgaca 11340

ttgatctggc tatcttgctg acaaaagcaa gagaacatag cgttgccttg gtaggtccag 11400

cggcggagga actctttgat ccggttcctg aacaggatct atttgaggcg ctaaatgaaa 11460

ccttaacgct atggaactcg ccgcccgact gggctggcga tgagcgaaat gtagtgctta 11520

cgttgtcccg catttggtac agcgcagtaa ccggcaaaat cgcgccgaag gatgtcgctg 11580

ccgactgggc aatggagcgc ctgccggccc agtatcagcc cgtcatactt gaagctaggc 11640

aggcttatct tggacaagaa gatcgcttgg cctcgcgcgc agatcagttg gaagaatttg 11700

ttcactacgt gaaaggcgag atcaccaaag tagtcggcaa ataaagctct agtggatctc 11760
cgtacccggg gatctggctc gcggcggacg cacgacgccg gggcgagacc ataggcgatc 11820
tcctaaatca atagtagctg taacctcgaa gcgtttcact tgtaacaacg attgagaatt 11880
tttgtcataa aattgaaata cttggttcgc atttttgtca tccgcggtca gccgcaattc 11940
tgacgaactg cccatttagc tggagatgat tgtacatcct tcacgtgaaa atttctcaag 12000
cgctgtgaac aagggttcag attttagatt gaaaggtgag ccgttgaaac acgttcttct 12060
tgtcgatgac gacgtcgcta tgcggcatct tattattgaa taccttacga tccacgcctt 12120
caaagtgacc gcggtagccg acagcaccca gttcacaaga gtactctctt ccgcgacggt 12180
cgatgtcgtg gttgttgatc tagatttagg tcgtgaagat gggctcgagc taggagcaag 12240
tgattttatc gctaagccgt tcagtatcag agagtttcta gcacgcattc gggttgcctt 12300
gcgcgtgcgc cccaacgttg tccgctccaa agaccgacgg tctttttgtt ttactgactg 12360
gacacttaat ctcaggcaac gtcgcttgat gtccgaagct ggcggtgagg tgaaacttac 12420
ggcaggtgag ttcaatcttc tcctcgcgtt tttagagaaa ccccgcgacg ttctatcgcg 12480
cgagcaactt ctcattgcca gtcgagtacg cgacgaggag gtttatgaca ggagtataga 12540
tgttctcatt ttgaggctgc gccgcaaact tgaggcagat ccgtcaagcc ctcaactgat 12600
aaaaacagca agaggtgccg gttatttctt tgacgcggac gtgcaggttt cgcacggggg 12660
gacgatggca gcctgagcca attcccagat ccccgaggaa tcggcgtgag cggtcgcaaa 12720
ccatccggcc cggtacaaat cggcgcggcg ctgggtgatg acctggtgga gaagttgaag 12780
gccgcgcagg ccgcccagcg gcaacgcatc gaggcagaag cacgccccgg tgaatcgtgg 12840
caagcggccg ctgatcgaat ccgcaaagaa tcccggcaac cgccggcagc cggtgcgccg 12900
tcgattagga agccgcccaa gggcgacgag caaccagatt ttttcgttcc gatgctctat 12960
gacgtgggca cccgcgatag tcgcagcatc atggacgtgg ccgttttccg tctgtcgaag 13020
cgtgaccgac gagctggcga ggtgatccgc tacgagcttc cagacgggca cgtagaggtt 13080
tccgcagggc cggccggcat ggccagtgtg tgggattacg acctggtact gatggcggtt 13140
tcccatctaa ccgaatccat gaaccgatac cgggaaggga agggagacaa gcccggccgc 13200
gtgttccgtc cacacgttgc ggacgtactc aagttctgcc ggcgagccga tggcggaaag 13260
cagaaagacg acctggtaga aacctgcatt cggttaaaca ccacgcacgt tgccatgcag 13320
cgtacgaaga aggccaagaa cggccgcctg gtgacggtat ccgagggtga agccttgatt 13380
agccgctaca agatcgtaaa gagcgaaacc gggcggccgg agtacatcga gatcgagcta 13440
gctgattgga tgtaccgcga gatcacagaa ggcaagaacc cggacgtgct gacggttcac 13500
cccgattact ttttgatcga tcccggcatc ggccgttttc tctaccgcct ggcacgccgc 13560
gccgcaggca aggcagaagc cagatggttg ttcaagacga tctacgaacg cagtggcagc 13620
gccggagagt tcaagaagtt ctgtttcacc gtgcgcaagc tgatcgggtc aaatgacctg 13680
ccggagtacg atttgaagga ggaggcgggg caggctggcc cgatcctagt catgcgctac 13740

cgcaacctga tcgagggcga agcatccgcc ggttcctaat gtacggagca gatgctaggg 13800

caaattgccc tagcagggga aaaaggtcga aaaggtctct ttcctgtgga tagcacgtac 13860

attgggaacc caaagccgta cattgggaac cggaacccgt acattgggaa cccaaagccg 13920

tacattggga accggtcaca catgtaagtg actgatataa aagagaaaaa aggcgatttt 13980

tccgcctaaa actctttaaa acttattaaa actcttaaaa cccgcctggc ctgtgcataa 14040

ctgtctggcc agcgcacagc cgaagagctg caaaaagcgc ctacccttcg gtcgctgcgc 14100

tccctacgcc ccgccgcttc gcgtcggcct atcgcggccg ctggccgctc aaaaatggct 14160

ggcctacggc caggcaatct accagggcgc ggacaagccg cgccgtcgcc actcgaccgc 14220

cggcgctgag gtctgcctcg tgaagaaggt gttgctgact cataccaggc ctgaatcgcc 14280

ccatcatcca gccagaaagt gagggagcca cggttgatga gagctttgtt gtaggtggac 14340

cagttggtga ttttgaactt ttgctttgcc acggaacggt ctgcgttgtc gggaagatgc 14400

gtgatctgat ccttcaactc agcaaaagtt cgatttattc aacaaagccg ccgtcccgtc 14460

aagtcagcgt aatgctctgc cagtgttaca accaattaac caattctgat tagaaaaact 14520

catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt 14580

gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa 14640

gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc 14700

cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg 14760

agaatggcaa aagctctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 14820

attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940

gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120

tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180

ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag 15240

gtcgttcgct ccaagctggg ctgtgtgcac gaacccccgg ttcagcccga ccgctgcgcc 15300

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420

aagtggtggc ctaactacgg ctacactaga agaacagtat ttggtatctg cgctctgctg 15480

aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540

ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 15600

gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660

gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt gatccggaat 15720

ta 15722

<210> 32
<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> mã hóa cho gRNA đối với vecto 23136
<400> 32
gtcggggacg tcgacgacga 20

<210> 33
<211> 1823
<212> ADN
<213> Oryza sativa
<400> 33
acagtgacta gtgacaaacg atcgatcgat ccctccatcc acaaaccctc ctcgatctca 60
tcttccttcg tctcgtcaat ggcggcgagc tactcgtgcc ggcggacatg cgaggcgtgc 120
agcacgaggg cgatggccgg gtgcgtggtg ggcgagccgg cgtcggcgcc ggggcagcgg 180
gtgacgttgc tggcgatcga cggcggcggc atcaggggcc tcatcccggg caccatcctc 240
gccttcctcg aggccaggct gcaggagctg gatggccccg acgcgcgcct cgccgattac 300
ttcgactgca tcgccgggac cagcaccggc ggcctcatca ccgccatgct cgccgcgccc 360
ggcgaccacg gccgcccgct cttcgccgcc agcgacatca accgcttcta cctcgacaac 420
ggcccactca tcttcccaca aaagtaactg atcacctcga attcgatctc ctctcttcga 480
tctctgcatt atttgatttg attggggatt gtgggcggcg tggcgtggcg tccaggaggt 540
gcggcatggc ggcggccatg gcggcgctga cgaggccgag gtacaacggc aagtacctgc 600
aggggaagat caggaagatg ctgggcgaga cgagggtgcg cgacacgctg acgaacgtcg 660
tcatccccac gttcgacgtc aggctgctcc agccaaccat cttctccaca tacgacgtgc 720
gtgcgttgat tccatccgca ttggcgttgg aatcagctga ttgtttgatt gatcgaacaa 780
ttgatcggtt aaaattttgc aggcgaagag catgccgctc aagaacgcgc tcctctccga 840
catctgcatc agcacatccg cggcgccgac ctacctcccc gcgcactgct tccagaccac 900
cgacgacgcc accggcaagg tccgcgagtt cgacctcatc gacggcggcg tcgccgccaa 960
caacccggta actaatcaat caagcaatcc atcaaacgaa gatccacatg tgcattcctg 1020
tggtacaaat gctgatcgat cgatggatgg atcgattttc gcgagaacgt acagacgatg 1080
gtggccatga cgcagatcac caagaagata atggtgaagg acaaggagga gctgtacccg 1140

gtaaagccgt cggactgcgg taagttcctg gtgctgtccg tgggcaccgg gtcgacgtcg 1200
gaccagggga tgtacacggc gaggcagtgc tcgcggtggg ggatcgtccg gtggctgcgc 1260
aacaagggga tggcgcccat catcgacatc ttcatggcgg ccagctccga cctcgtcgac 1320
atccacgccg ccgtcatgtt ccagtcgctg cacagcgacg gcgactacct ccgcatccag 1380
gacaacacgc tccacggcga cgccgccacg gtggacgccg ccaccaggga caacatgcgg 1440
gcgctcgtcg ggatcggcga gcggatgctg gcgcagcggg tgtcgagggt caacgtcgag 1500
accggcaggt acgtcgaggt gcccggcgcc ggcagcaacg ccgacgcgct gaggggcttc 1560
gccaggcagc tctccgagga gaggagggcg aggctaggtc ggcgaaacgc ctgcggcggc 1620
ggcggcgaag gagagcccag cggcgtggcg tgcaagcgtt agtaactgta cacgcatcat 1680
gctgacgcga tctttttttat ttttcttttt ttttttttac ctttctagcg gacatgggga 1740
ataacaagac gtgacagtag tgcaatcggt ttgtaacgtg cgtataccaa cattgatcca 1800
tttcttcatc acagtttcag ttc 1823

<210> 34
<211> 15921
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> Vecto 24038

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (313)..(1149)
<223> prZmGRMZM5G876285-01
<220>
<221> gen
<222> (1152)..(5412)
<223> cCas9-12
<220>
<221> trình tự kết thúc

<222> (5419)..(6736)

<223> tZmGRMZM5G876285-01

<220>

<221> vùng khởi động

<222> (6750)..(7124)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7126)..(7145)

<223> xZmVLHP2

<220>

<221> đặc điểm_hỗn tạp

<222> (7126)..(7230)

<223> rsgRNAZmVLHP-02

<220>

<221> đặc điểm_hỗn tạp

<222> (7146)..(7157)

<223> rCrRNA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7146)..(7230)

<223> rsgRNAbase-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7162)..(7230)

<223> rTracrRNA-01

<220>

<221> vùng khởi động

<222> (7241)..(9232)

<223> prUbi-04

<220>

<221> gen

<222> (9249)..(10427)

<223> cPMI-09

<220>
<221> trình tự kết thúc
<222> (10450)..(10702)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp
<222> (10746)..(10875)
<223> bNLB-03
<220>
<221> gen
<222> (11155)..(11943)
<223> cSpec-03
<220>
<221> vùng khởi động
<222> (12038)..(12168)
<223> prVirG-01
<220>
<221> gen
<222> (12243)..(12875)
<223> cVirG-09
<220>
<221> gen
<222> (12905)..(13978)
<223> cRepA-01
<220>
<221> đặc điểm_hỗn tạp
<222> (14021)..(14425)
<223> oVA1-02
<220>
<221> đặc điểm_hỗn tạp
<222> (15103)..(15909)
<223> oCOLE-06
<400> 34
attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120
tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180
attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240
aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300
taccgggacc ctaagtaatc ttgtgctaca aatttatttt tcagacagaa aaatctattt 360
agctaactaa ttaatacaaa ttaataccaa gcaacgatag atgaacatct agttgtctaa 420
ttagctaact aattaataca aattaagtag aatccttacc gtggggagat ggggcgcgac 480
gaagtgctcg agcttggggc gcggcgaccg gcgacgtgaa gcttgggggc gcgggggccg 540
gacggcgctg cgggcggcat ggcgggcggc tgcgggcggc ggcgcgggcg caggaaacaa 600
acgacgggag tgggaggaag gagaaagcgg cgcgccggtt tagtcctagc tcggcgccaa 660
gatctgtggc gccgagctag gtgccacgat ggccgccgcg tcagcaaagc tcggcgccaa 720
ggcatgttgc gccgagccgt gttagctcgg cgtcatagct catggtgccg agttttgggt 780
ctaaaattgc gtttaagtat tctagggatc taaacgcaaa tatttttcga aaatagggcc 840
gaaaaacaaa aaaaaatcgg tcgtttcgtc gagcacatcg tccagcctat cttgcatgtc 900
catcctctct atggttcgcg agccgcgcgc atggcgctcc aaaggagggg cgaggttgaa 960
tatagacaga tggaatgggt ggttctctat ttatagcgca tgcagtcgtc ccctggcaca 1020
cctatttata tgtgagcgtt cctggcacta gagagatcga tcgatcgagc ttaattgcgc 1080
cactgctcgt tatcctcctc ttgcattgca ttgcaggtcg tagttgagca gcagcaacca 1140
ctgcacaggc catggacaag aagtacagca tcggcctgga catcggcacc aacagcgtgg 1200
gctgggccgt gatcaccgac gagtacaagg tgataccaat ttgcatgatc cttgttcgtt 1260
ctagctcttg catgccgatc agttgaatca cgcggtttcc ttctgcgcat ttgcatccag 1320
gtgccgagca agaagttcaa ggtgctgggc aacaccgaca ggcacagcat caagaagaac 1380
ctgatcggcg ccctgctgtt cgacagcggc gagaccgccg aggccaccag gctgaagagg 1440
accgccagga ggaggtacac caggaggaag aacaggatct gctacctgca ggagatcttc 1500
agcaacgaga tggccaaggt ggacgacagc ttcttccaca ggctggagga gagcttcctg 1560
gtggaggagg acaagaagca cgagaggcac ccgatcttcg gcaacatcgt ggacgaggtg 1620
gcctaccacg agaagtaccc gaccatctac cacctgagga agaagctggt ggacagcacc 1680
gacaaggccg acctgaggct gatctacctg gccctggccc acatgatcaa gttcaggggc 1740
cacttcctga tcgagggcga cctgaacccg gacaacagcg acgtggacaa gctgttcatc 1800
cagctggtgc agacctacaa ccagctgttc gaggagaacc cgatcaacgc cagcggcgtg 1860
gacgccaagg ccatcctgag cgccaggctg agcaagagca ggaggctgga gaacctgatc 1920
gcccagctgc cgggcgagaa gaagaacggc ctgttcggca acctgatcgc cctgagcctg 1980
ggcctgaccc cgaacttcaa gagcaacttc gacctggccg aggacgccaa gctgcagctg 2040
agcaaggaca cctacgacga cgacctggac aacctgctgg cccagatcgg cgaccagtac 2100

gccgacctgt tcctggccgc caagaacctg agcgacgcca tcctgctgag cgacatcctg 2160

agggtgaaca ccgagatcac caaggccccg ctgagcgcca gcatgatcaa gaggtacgac 2220

gagcaccacc aggacctgac cctgctgaag gccctggtga ggcagcagct gccggagaag 2280

tacaaggaga tcttcttcga ccagagcaag aacggctacg ccggctacat cgacggcggc 2340

gccagccagg aggagttcta caagttcatc aagccgatcc tggagaagat ggacggcacc 2400

gaggagctgc tggtgaagct gaacagggag gacctgctga ggaagcagag gaccttcgac 2460

aacggcagca tcccgcacca gatccacctg ggcgagctgc acgccatcct gaggaggcag 2520

gaggacttct acccgttcct gaaggacaac agggagaaga tcgagaagat cctgaccttc 2580

cgcatcccgt actacgtggg cccgctggcc aggggcaaca gcaggttcgc ctggatgacc 2640

aggaagagcg aggagaccat caccccgtgg aacttcgagg aggtggtgga caagggcgcc 2700

agcgcccaga gcttcatcga gaggatgacc aacttcgaca agaacctgcc gaacgagaag 2760

gtgctgccga agcacagcct gctgtacgag tacttcaccg tgtacaacga gctgaccaag 2820

gtgaagtacg tgaccgaggg catgaggaag ccggccttcc tgagcggcga gcagaagaag 2880

gccatcgtgg acctgctgtt caagaccaac aggaaggtga ccgtgaagca gctgaaggag 2940

gactacttca agaagatcga gtgcttcgac agcgtggaga tcagcggcgt ggaggacagg 3000

ttcaacgcca gcctgggcac ctaccacgac ctgctgaaga tcatcaagga caaggacttc 3060

ctggacaacg aggagaacga ggacatcctg gaggacatcg tgctgaccct gaccctgttc 3120

gaggacaggg agatgatcga ggagaggctg aagacctacg cccacctgtt cgacgacaag 3180

gtgatgaagc agctgaagag gaggaggtac accggctggg gcaggctgag caggaagctg 3240

atcaacggca tcagggacaa gcagagcggc aagaccatcc tggacttcct gaagagcgac 3300

ggcttcgcca acaggaactt catgcagctg atccacgacg acagcctgac cttcaaggag 3360

gacatccaga aggcccaggt gagcggccag ggcgacagcc tgcacgagca catcgccaac 3420

ctggccggca gcccggccat caagaagggc atcctgcaga ccgtgaaggt ggtggacgag 3480

ctggtgaagg tgatgggcag gcacaagccg gagaacatcg tgatcgagat ggccagggag 3540

aaccagacca cccagaaggg ccagaagaac agcagggaga ggatgaagag gatcgaggag 3600

ggcatcaagg agctgggcag ccagatcctg aaggagcacc cggtggagaa cacccagctg 3660

cagaacgaga agctgtacct gtactacctg cagaacggca gggacatgta cgtggaccag 3720

gagctggaca tcaacaggct gagcgactac gacgtggacc acatcgtgcc gcagagcttc 3780

ctgaaggacg acagcatcga caacaaggtg ctgaccagga gcgacaagaa caggggcaag 3840

agcgacaacg tgccgagcga ggaggtggtg aagaagatga aaaactactg gaggcagctg 3900

ctgaacgcca agctgatcac ccagaggaag ttcgacaacc tgaccaaggc cgagaggggc 3960

ggcctgagcg agctggacaa ggccggcttc attaaaaggc agctggtgga gaccaggcag 4020

atcaccaagc acgtggccca gatcctggac agcaggatga acaccaagta cgacgagaac 4080

gacaagctga tcagggaggt gaaggtgatc accctgaaga gcaagctggt gagcgacttc 4140

aggaaggact tccagttcta caaggtgagg gagatcaata attaccacca cgcccacgac 4200
gcctacctga acgccgtggt gggcaccgcc ctgattaaaa agtacccgaa gctggagagc 4260
gagttcgtgt acggcgacta caaggtgtac gacgtgagga agatgatcgc caagagcgag 4320
caggagatcg gcaaggccac cgccaagtac ttcttctaca gcaacatcat gaacttcttc 4380
aagaccgaga tcaccctggc caacggcgag atcaggaaga ggccgctgat cgagaccaac 4440
ggcgagaccg gcgagatcgt gtgggacaag ggcagggact tcgccaccgt gaggaaggtg 4500
ctgtccatgc cgcaggtgaa catcgtgaag aagaccgagg tgcagaccgg cggcttcagc 4560
aaggagagca tcctgccgaa gaggaacagc gacaagctga tcgccaggaa gaaggactgg 4620
gacccgaaga agtacggcgg cttcgacagc ccgaccgtgg cctacagcgt gctggtggtg 4680
gccaaggtgg agaagggcaa gagcaagaag ctgaagagcg tgaaggagct ggtgggcatc 4740
accatcatgg agaggagcag cttcgagaag aacccagtgg acttcctgga ggccaagggc 4800
tacaaggagg tgaagaagga cctgatcatt aaactgccga agtacagcct gttcgagctg 4860
gagaacggca ggaagaggat gctggccagc gccggcgagc tgcagaaggg caacgagctg 4920
gccctgccga gcaagtacgt gaacttcctg tacctggcca gccactacga gaagctgaag 4980
ggcagcccgg aggacaacga gcagaagcag ctgttcgtgg agcagcacaa gcactacctg 5040
gacgagatca tcgagcagat cagcgagttc agcaagaggg tgatcctggc cgacgccaac 5100
ctggacaagg tgctgagcgc ctacaacaag cacagggaca agccgatcag ggagcaggcc 5160
gagaacatca tccacctgtt caccctgacc aacctgggcg ccccggccgc cttcaagtac 5220
ttcgacacca ccatcgacag gaagaggtac accagcacca aggaggtgct ggacgccacc 5280
ctgatccacc agagcatcac cggcctgtac gagaccagga tcgacctgag ccagctgggc 5340
ggcgacagca gcccgccgaa gaagaagagg aaggtgagct ggaaggacgc cagcggctgg 5400
agcaggatgt gagctctaat gcatccaaac aacgacacca acgccaacat taattaatta 5460
gtagtctcca tgccctggga ttgtgcgtgg ccgctccgtt gaacaccacc catccttcgt 5520
tcggcatttt ttccccccctt gtttatataa ttttattgta tcgttttggc aaataatttt 5580
gtgattcgac cccaaagcaa gtttggttgt cttacgattt gtaaacctgg aacaatatat 5640
aatgtgattg aactgctttg tctattcttt ttgtagtacg ataatatgta tatgtattcc 5700
atgcgatctc ttctagggcg acgactaatg tgcaagtgtg tgtttgcatg cgctgagcac 5760
ggagtttgta ttcaggggtc aatatctttc gattccttta tctaaaaagg tgttgcatat 5820
atctaaaaaa aagaaaaaaa aggcttacaa ctgttgaaaa aataagcatt tttagtttta 5880
atttaattca gaaaatcata gtgatatatg tgacgatatg catgtgcata tgtatcacta 5940
ctcacataaa cagtaaacaa cagtaaaata tgtataaata caaaaataac aaagtgtacc 6000
ctgcggaggg accgatgttc aaggcatctg tggctccatt cacacgagac atctcgtgtg 6060
tatgttcgat gtagtcatac gcagtcgagg cagtcagatg tacgcagtgc agtccctcga 6120
tcggcgccgg cgacgaggaa cttgatcagt gctggtcgag cggacgaagc gagcagtcgc 6180

gagtacgctc ccgaaaaaca tgatcgctcg cacacccatg caagtgtcgc tctgcggacg 6240
acgatttcgg aagcctacgc gtatgagaat gtttgtatgt gtgttctctc gtaaccagaa 6300
gcctcatctc ctccgtatat atacacgcgc agagggaggc caacagatag taacggtgga 6360
aggaatactc ggaccaaggt ccgatctacc atggccacgg cccggcctgg ccagcggcgc 6420
gtgcgtgtgg cagtccttca tccttttatc agcttatcaa tagatgcacc aaagatccac 6480
ctatttaagt tgattgaatt gtctcttgta cttccggtat gttactaaag taataataca 6540
ccgtagcatt aaattgggcc tttagcattg gctattattg aatattaatt tgagccagac 6600
ccaccaccag atgctaagtc acaccaaaat gctctcatca tctcaaacat ttcatatact 6660
ggtgtttcga tggagactat taagttgaac atccacctag aatctagatt acacttgacc 6720
acaactacat aatggacgga ccgttcgaag ggatctttaa acatacgaac agatcactta 6780
aagttcttct gaagcaactt aaagttatca ggcatgcatg gatcttggag gaatcagatg 6840
tgcagtcagg gaccatagca caggacaggc gtcttctact ggtgctacca gcaaatgctg 6900
gaagccggga acactgggta cgttggaaac cacgtgatgt ggagtaagat aaactgtagg 6960
agaaaagcat ttcgtagtgg gccatgaagc ctttcaggac atgtattgca gtatgggccg 7020
gcccattacg caattggacg acaacaaaga ctagtattag taccacctcg gctatccaca 7080
tagatcaaag ctggtttaaa agagttgtgc agatgatccg tggcagctgg agctgagctt 7140
ccggggtttt agagctagaa atagcaagtt aaaataaggc tagtccgtta tcaacttgaa 7200
aaagtggcac cgagtcggtg ctttttttttt cggaccgcgc ctgcagtgca gcgtgacccg 7260
gtcgtgcccc tctctagaga taatgagcat tgcatgtcta agttataaaa aattaccaca 7320
tatttttttt gtcacacttg tttgaagtgc agtttatcta tctttataca tatatttaaa 7380
ctttactcta cgaataatat aatctatagt actacaataa tatcagtgtt ttagagaatc 7440
atataaatga acagttagac atggtctaaa ggacaattga gtattttgac aacaggactc 7500
tacagtttta tctttttagt gtgcatgtgt tctccttttt ttttgcaaat agcttcacct 7560
atataatact tcatccattt tattagtaca tccatttagg gtttagggtt aatggttttt 7620
atagactaat ttttttagta catctatttt attctatttt agcctctaaa ttaagaaaac 7680
taaaactcta ttttagtttt tttatttaat aatttagata taaaatagaa taaaataaag 7740
tgactaaaaa ttaaacaaat acccttaag aaattaaaaa aactaaggaa acatttttct 7800
tgtttcgagt agataatgcc agcctgttaa acgccgtcga cgagtctaac ggacaccaac 7860
cagcgaacca gcagcgtcgc gtcgggccaa gcgaagcaga cggcacggca tctctgtcgc 7920
tgcctctgga cccctctcga gagttccgct ccaccgttgg acttgctccg ctgtcggcat 7980
ccagaaattg cgtggcggag cggcagacgt gagccggcac ggcaggcggc ctcctcctcc 8040
tctcacggca ccggcagcta cgggggattc ctttcccacc gctccttcgc tttcccttcc 8100
tcgcccgccg taataaatag acacccccctc cacaccctct ttccccaacc tcgtgttgtt 8160
cggagcgcac acacacacaa ccagatctcc cccaaatcca cccgtcggca cctccgcttc 8220

aaggtacgcc gctcgtcctc ccccccccc ctctctacct tctctagatc ggcgttccgg 8280

tccatggtta gggcccggta gttctacttc tgttcatgtt tgtgttagat ccgtgtttgt 8340

gttagatccg tgctgctagc gttcgtacac ggatgcgacc tgtacgtcag acacgttctg 8400

attgctaact tgccagtgtt tctctttggg gaatcctggg atggctctag ccgttccgca 8460

gacgggatcg atttcatgat ttttttgtt tcgttgcata gggtttggtt tgccctttc 8520

ctttatttca atatatgccg tgcacttgtt tgtcgggtca tcttttcatg ctttttttg 8580

tcttggttgt gatgatgtgg tctggttggg cggtcgttct agatcggagt agaattctgt 8640

ttcaaactac ctggtggatt tattaatttt ggatctgtat gtgtgtgcca tacatattca 8700

tagttacgaa ttgaagatga tggatggaaa tatcgatcta ggataggtat acatgttgat 8760

gcgggtttta ctgatgcata tacagagatg ctttttgttc gcttggttgt gatgatgtgg 8820

tgtggttggg cggtcgttca ttcgttctag atcggagtag aatactgttt caaactacct 8880

ggtgtattta ttaattttgg aactgtatgt gtgtgtcata catcttcata gttacgagtt 8940

taagatggat ggaaatatcg atctaggata ggtatacatg ttgatgtggg ttttactgat 9000

gcatatacat gatggcatat gcagcatcta ttcatatgct ctaaccttga gtacctatct 9060

attataataa acaagtatgt tttataatta ttttgatctt gatatacttg gatgatggca 9120

tatgcagcag ctatatgtgg attttttttag ccctgccttc atacgctatt tatttgcttg 9180

gtactgtttc ttttgtcgat gctcaccctg ttgtttggtg ttacttctgc agggatccgg 9240

cagcagccat gcagaagctg atcaacagcg tgcagaacta cgcctggggc agcaagaccg 9300

ccctgaccga gctgtacggc atggagaacc ccagcagcca gcccatggcc gagctgtgga 9360

tgggcgccca ccccaagagc agcagccgcg tgcagaacgc cgccggcgac atcgtgagcc 9420

tgcgcgacgt gatcgagagc gacaagagca ccctgctggg cgaggccgtg gccaagcgct 9480

tcggcgagct gcccttcctg ttcaaggtgc tgtgcgccgc ccagcccctg agcatccagg 9540

tgcaccccaa caagcacaac agcgagatcg gcttcgccaa ggagaacgcc gccggcatcc 9600

ccatggacgc cgccgagcgc aactacaagg accccaacca caagcccgag ctggtgttcg 9660

ccctgacccc cttcctggcc atgaacgcct tccgcgagtt cagcgagatc gtgagcctgc 9720

tgcagcccgt ggccggcgcc caccccgcca tcgcccactt cctgcagcag cccgacgccg 9780

agcgcctgag cgagctgttc gccagcctgc tgaacatgca gggcgaggag aagagccgcg 9840

ccctggccat cctgaagagc gccctggaca gccagcaggg cgagccctgg cagaccatcc 9900

gcctgatcag cgagttctac cccgaggaca gcggcctgtt cagcccctg ctgctgaacg 9960

tggtgaagct gaaccccggc gaggccatgt tcctgttcgc cgagaccccc cacgcctacc 10020

tgcagggcgt ggccctggag gtgatggcca acagcgacaa cgtgctgcgc gccggcctga 10080

cccccaagta catcgacatc cccgagctgg tggccaacgt gaagttcgag gccaagcccg 10140

ccaaccagct gctgacccag cccgtgaagc agggcgccga gctggacttc cccatccccg 10200

tggacgactt cgccttcagc ctgcacgacc tgagcgacaa ggagaccacc atcagccagc 10260

agagcgccgc catcctgttc tgcgtggagg gcgacgccac cctgtggaag ggcagccagc 10320
agctgcagct gaagcccggc gagagcgcct tcatcgccgc caacgagagc cccgtgaccg 10380
tgaagggcca cggccgcctg gcccgcgtgt acaacaagct gtgataggag ctcgatccgt 10440
cgacctgcag atcgttcaaa catttggcaa taaagtttct taagattgaa tcctgttgcc 10500
ggtcttgcga tgattatcat ataatttctg ttgaattacg ttaagcatgt aataattaac 10560
atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc gcaattatac 10620
atttaatacg cgatagaaaa caaaatatag cgcgcaaact aggataaatt atcgcgcgcg 10680
gtgtcatcta tgttactaga tcggcgcgcc gcaattgaag tttgggcggc cagcatggcc 10740
gtatccgcaa tgtgttatta agttgtctaa gcgtcaattt gtttacacca caatatatcc 10800
tgccaccagc cagccaacag ctccccgacc ggcagctcgg cacaaaatca ccactcgata 10860
caggcagccc atcagaatta attctcatgt ttgacagctt atcatcgact gcacggtgca 10920
ccaatgcttc tggcgtcagg cagccatcgg aagctgtggt atggctgtgc aggtcgtaaa 10980
tcactgcata attcgtgtcg ctcaaggcgc actcccgttc tggataatgt tttttgcgcc 11040
gacatcataa cggttctggc aaatattctg aaatgagctg ttgacaatta atcatccggc 11100
tcgtataatg tgtggaattg tgagcggata acaatttcac acaggaaaca gaccatgagg 11160
gaagcgttga tcgccgaagt atcgactcaa ctatcagagg tagttggcgt catcgagcgc 11220
catctcgaac cgacgttgct ggccgtacat ttgtacggct ccgcagtgga tggcggcctg 11280
aagccacaca gtgatattga tttgctggtt acggtgaccg taaggcttga tgaaacaacg 11340
cggcgagctt tgatcaacga cctttttggaa acttcggctt cccctggaga gagcgagatt 11400
ctccgcgctg tagaagtcac cattgttgtg cacgacgaca tcattccgtg gcgttatcca 11460
gctaagcgcg aactgcaatt tggagaatgg cagcgcaatg acattcttgc aggtatcttc 11520
gagccagcca cgatcgacat tgatctggct atcttgctga caaaagcaag agaacatagc 11580
gttgccttgg taggtccagc ggcggaggaa ctctttgatc cggttcctga acaggatcta 11640
tttgaggcgc taaatgaaac cttaacgcta tggaactcgc cgcccgactg ggctggcgat 11700
gagcgaaatg tagtgcttac gttgtcccgc atttggtaca gcgcagtaac cggcaaaatc 11760
gcgccgaagg atgtcgctgc cgactgggca atggagcgcc tgccggccca gtatcagccc 11820
gtcatacttg aagctaggca ggcttatctt ggacaagaag atcgcttggc ctcgcgcgca 11880
gatcagttgg aagaatttgt tcactacgtg aaaggcgaga tcaccaaagt agtcggcaaa 11940
taaagctcta gtggatctcc gtacccgggg atctggctcg cggcggacgc acgacgccgg 12000
ggcgagacca taggcgatct cctaaatcaa tagtagctgt aacctcgaag cgtttcactt 12060
gtaacaacga ttgagaattt ttgtcataaa attgaaatac ttggttcgca tttttgtcat 12120
ccgcggtcag ccgcaattct gacgaactgc ccatttagct ggagatgatt gtacatcctt 12180
cacgtgaaaa tttctcaagc gctgtgaaca agggttcaga ttttagattg aaaggtgagc 12240
cgttgaaaca cgttcttctt gtcgatgacg acgtcgctat gcggcatctt attattgaat 12300

accttacgat ccacgccttc aaagtgaccg cggtagccga cagcacccag ttcacaagag 12360
tactctcttc cgcgacggtc gatgtcgtgg ttgttgatct agatttaggt cgtgaagatg 12420
ggctcgagct aggagcaagt gattttatcg ctaagccgtt cagtatcaga gagtttctag 12480
cacgcattcg ggttgccttg cgcgtgcgcc ccaacgttgt ccgctccaaa gaccgacggt 12540
ctttttgttt tactgactgg acacttaatc tcaggcaacg tcgcttgatg tccgaagctg 12600
gcggtgaggt gaaacttacg gcaggtgagt tcaatcttct cctcgcgttt ttagagaaac 12660
cccgcgacgt tctatcgcgc gagcaacttc tcattgccag tcgagtacgc gacgaggagg 12720
tttatgacag gagtatagat gttctcattt tgaggctgcg ccgcaaactt gaggcagatc 12780
cgtcaagccc tcaactgata aaaacagcaa gaggtgccgg ttatttcttt gacgcggacg 12840
tgcaggtttc gcacgggggg acgatggcag cctgagccaa ttcccagatc cccgaggaat 12900
cggcgtgagc ggtcgcaaac catccggccc ggtacaaatc ggcgcggcgc tgggtgatga 12960
cctggtggag aagttgaagg ccgcgcaggc cgcccagcgg caacgcatcg aggcagaagc 13020
acgccccggt gaatcgtggc aagcggccgc tgatcgaatc cgcaaagaat cccggcaacc 13080
gccggcagcc ggtgcgccgt cgattaggaa gccgcccaag ggcgacgagc aaccagattt 13140
tttcgttccg atgctctatg acgtgggcac ccgcgatagt cgcagcatca tggacgtggc 13200
cgttttccgt ctgtcgaagc gtgaccgacg agctggcgag gtgatccgct acgagcttcc 13260
agacgggcac gtagaggttt ccgcagggcc ggccggcatg gccagtgtgt gggattacga 13320
cctggtactg atggcggttt cccatctaac cgaatccatg aaccgatacc gggaagggaa 13380
gggagacaag cccggccgcg tgttccgtcc acacgttgcg gacgtactca agttctgccg 13440
gcgagccgat ggcggaaagc agaaagacga cctggtagaa acctgcattc ggttaaacac 13500
cacgcacgtt gccatgcagc gtacgaagaa ggccaagaac ggccgcctgg tgacggtatc 13560
cgagggtgaa gccttgatta gccgctacaa gatcgtaaag agcgaaaccg ggcggccgga 13620
gtacatcgag atcgagctag ctgattggat gtaccgcgag atcacagaag gcaagaaccc 13680
ggacgtgctg acggttcacc ccgattactt tttgatcgat cccggcatcg gccgttttct 13740
ctaccgcctg gcacgccgcg ccgcaggcaa ggcagaagcc agatggttgt tcaagacgat 13800
ctacgaacgc agtggcagcg ccggagagtt caagaagttc tgtttcaccg tgcgcaagct 13860
gatcgggtca aatgacctgc cggagtacga tttgaaggag gaggcggggc aggctggccc 13920
gatcctagtc atgcgctacc gcaacctgat cgagggcgaa gcatccgccg gttcctaatg 13980
tacggagcag atgctagggc aaattgccct agcaggggaa aaaggtcgaa aaggtctctt 14040
tcctgtggat agcacgtaca ttgggaaccc aaagccgtac attgggaacc ggaacccgta 14100
cattgggaac ccaaagccgt acattgggaa ccggtcacac atgtaagtga ctgatataaa 14160
agagaaaaaa ggcgattttt ccgcctaaaa ctctttaaaa cttattaaaa ctcttaaaac 14220
ccgcctggcc tgtgcataac tgtctggcca gcgcacagcc gaagagctgc aaaaagcgcc 14280
taccttcgg tcgctgcgct ccctacgccc cgccgcttcg cgtcggccta tcgcggccgc 14340

tggccgctca aaaatggctg gcctacggcc aggcaatcta ccagggcgcg gacaagccgc 14400
gccgtcgcca ctcgaccgcc ggcgctgagg tctgcctcgt gaagaaggtg ttgctgactc 14460
ataccaggcc tgaatcgccc catcatccag ccagaaagtg agggagccac ggttgatgag 14520
agctttgttg taggtggacc agttggtgat tttgaacttt tgctttgcca cggaacggtc 14580
tgcgttgtcg ggaagatgcg tgatctgatc cttcaactca gcaaaagttc gatttattca 14640
acaaagccgc cgtcccgtca agtcagcgta atgctctgcc agtgttacaa ccaattaacc 14700
aattctgatt agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga 14760
ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg 14820
cagttccata ggatggcaag atcctggtat cggtctgcga ttccgactcg tccaacatca 14880
atacaaccta ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga 14940
gtgacgactg aatccggtga gaatggcaaa agctctgcat taatgaatcg gccaacgcgc 15000
ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg 15060
ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc 15120
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 15180
gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca 15240
tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca 15300
ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 15360
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 15420
gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccccgt 15480
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 15540
cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 15600
cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa gaacagtatt 15660
tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 15720
cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg 15780
cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg 15840
gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 15900
gatccttttg atccggaatt a 15921

<210> 35
<211> 17954
<212> ADN
<213> trình tự nhân tạo
<220>
<223> vecto 24039

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (315)..(1729)
<223> prZmGRMZM2G020852-01
<220>
<221> gen
<222> (1731)..(5979)
<223> cCas9-13
<220>
<221> trình tự kết thúc
<222> (5989)..(8769)
<223> tZmGRMZM2G020852-01
<220>
<221> vùng khởi động
<222> (8783)..(9157)
<223> prOsU3-01
<220>
<221> đặc điểm_hỗn tạp
<222> (9159)..(9178)
<223> xZmVLHP2
<220>
<221> đặc điểm_hỗn tạp
<222> (9159)..(9263)
<223> rsgRNAZmVLHP-02
<220>
<221> đặc điểm_hỗn tạp
<222> (9179)..(9190)
<223> rCrRNA-01
<220>

<221> đặc điểm_hỗn tạp
<222> (9195)..(9263)
<223> rTracrRNA-01
<220>
<221> vùng khởi động
<222> (9274)..(11265)
<223> prUbi1-04
<220>
<221> gen
<222> (11282)..(12460)
<223> cPMI-09
<220>
<221> trình tự kết thúc
<222> (12483)..(12735)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp
<222> (12779)..(12908)
<223> bNLB-03
<220>
<221> gen
<222> (13188)..(13976)
<223> cSpec-03
<220>
<221> vùng khởi động
<222> (14071)..(14201)
<223> prVirG-01
<220>
<221> gen
<222> (14276)..(14908)
<223> cVirG-09
<220>
<221> gen
<222> (14938)..(16011)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (16054)..(16458)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (17136)..(17942)

<223> oCOLE-o6

<400> 35

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

taccggaccg ttataacagt gaatacaaaa atgacattcg tgttatttag cacaagttac 360

gatctatttc aggaacatgc cggaattttc gaacaccatt ctcacaaaac atgaccttga 420

acttgcgatc cagttgtttt aaaattatat aaaacaaaaa caaagtcaga aaatcatgaa 480

acttgtcgac atgtcatgat atcatatgta gagactctaa taaaaagttg agattgtttc 540

atgaaagttg tcacacacta tgtgtagaaa cttagcccgt ctacattgaa gttctatgat 600

ttcatgtgaa ggacacctag gcatcgatgt ttatgataat atcttatgtt tgtttggaca 660

aaatattaaa aacaaataaa aggggtccct gatcactttg acgagcattg cattcagcaa 720

agggtgcctt tgttgagtgc aatggtcata gaactcggta gaaaagacat acataaacat 780

cgggaaactt gctttaccgc acgctatggc caagacactc ggcaaactag gctcctttgt 840

tgagtgccat ctcaagcact cgacattgga actacgacta ggcctcacgg aagctttctt 900

tgccgagtgc cactaagcga ggaactcgga cactcagcaa cagctctgtc atcgtcacga 960

tgtcttttct ttgtcgtgta ccagttggca ctcggttaag actttactga gtgcccgata 1020

gaaagtactc ggcaaagaga ccgttgccga cgtttggttc actgagggct ctttgctgcc 1080

ttttggactt gacaaagaag tcatctccag tactgtctcc taggacgcag gatttatgtt 1140

ttttcccgga gctcgatctg tgggacatca cagatggtcc aatctggtga tctaaaatgg 1200

acggtttgcc aagcccacag agaagtcttt aagatcttcc acgatgcacg catgctttaa 1260

ggttagatag tgtttggtcc aaaaaagcgt caacaatcag gaaattagaa ctaaaattat 1320

taaaggacag atcaaaaggc atgcatgttc ttcttctata gtgtgtgttg agcctgagtt 1380

ttgattttag gctttattag gggactcgca gtctagctaa ggagttgtat tgatgttctg 1440

acaaatatta tgttcgatcg tcacagtggt cttgtgcgga tcgattaggc ccgatcatgg 1500

tgaaataaac taaccaccgg taagcccggg cagccctaga gcatgcagcg gcctacgtga 1560

agcccgcgtg tcgcatcgtc gtccgtcaga cgctaacggc aggccgctgc atgcgttgcc 1620

ggcgaactct ctcctgagcc actcgtcatc catataagta gacatcccat cactgtcgtc 1680

tatcaacaac acacagagcg acatttcgaa taacacagtt gagcgcgacc atggacaaga 1740

agtacagcat cggcctggac atcggcacca acagcgtggg ctgggccgtg atcaccgacg 1800

agtacaaggt acgagcggga tacatgttta tactcctcct gtaggtcgct ccttcatgta 1860

atgtgttgcg attaaaacgg tgcgcaggtg ccgagcaaga agttcaaggt gctgggcaac 1920

accgacaggc acagcatcaa gaagaacctg atcggcgccc tgctgttcga cagcggcgag 1980

accgccgagg ccaccaggct gaagaggacc gccaggagga ggtacaccag gaggaagaac 2040

aggatctgct acctgcagga gatcttcagc aacgagatgg ccaaggtgga cgacagcttc 2100

ttccacaggc tggaggagag cttcctggtg gaggaggaca agaagcacga gaggcacccg 2160

atcttcggca acatcgtgga cgaggtggcc taccacgaga agtacccgac catctaccac 2220

ctgaggaaga agctggtgga cagcaccgac aaggccgacc tgaggctgat ctacctggcc 2280

ctggcccaca tgatcaagtt caggggccac ttcctgatcg agggcgacct gaacccggac 2340

aacagcgacg tggacaagct gttcatccag ctggtgcaga cctacaacca gctgttcgag 2400

gagaacccga tcaacgccag cggcgtggac gccaaggcca tcctgagcgc caggctgagc 2460

aagagcagga ggctggagaa cctgatcgcc cagctgccgg gcgagaagaa gaacggcctg 2520

ttcggcaacc tgatcgccct gagcctgggc ctgaccccga acttcaagag caacttcgac 2580

ctggccgagg acgccaagct gcagctgagc aaggacacct acgacgacga cctggacaac 2640

ctgctggccc agatcggcga ccagtacgcc gacctgttcc tggccgccaa gaacctgagc 2700

gacgccatcc tgctgagcga catcctgagg gtgaacaccg agatcaccaa ggccccgctg 2760

agcgccagca tgatcaagag gtacgacgag caccaccagg acctgaccct gctgaaggcc 2820

ctggtgaggc agcagctgcc ggagaagtac aaggagatct tcttcgacca gagcaagaac 2880

ggctacgccg gctacatcga cggcggcgcc agccaggagg agttctacaa gttcatcaag 2940

ccgatcctgg agaagatgga cggcaccgag gagctgctgg tgaagctgaa cagggaggac 3000

ctgctgagga agcagaggac cttcgacaac ggcagcatcc cgcaccagat ccacctgggc 3060

gagctgcacg ccatcctgag gaggcaggag gacttctacc cgttcctgaa ggacaacagg 3120

gagaagatcg agaagatcct gaccttccgc atcccgtact acgtgggccc gctggccagg 3180

ggcaacagca ggttcgcctg gatgaccagg aagagcgagg agaccatcac cccgtggaac 3240

ttcgaggagg tggtggacaa gggcgccagc gcccagagct tcatcgagag gatgaccaac 3300

ttcgacaaga acctgccgaa cgagaaggtg ctgccgaagc acagcctgct gtacgagtac 3360

ttcaccgtgt acaacgagct gaccaaggtg aagtacgtga ccgagggcat gaggaagccg 3420

gccttcctga gcggcgagca gaagaaggcc atcgtggacc tgctgttcaa gaccaacagg 3480

aaggtgaccg tgaagcagct gaaggaggac tacttcaaga agatcgagtg cttcgacagc 3540

gtggagatca gcggcgtgga ggacaggttc aacgccagcc tgggcaccta ccacgacctg 3600

ctgaagatca tcaaggacaa ggacttcctg gacaacgagg agaacgagga catcctggag 3660

gacatcgtgc tgaccctgac cctgttcgag gacagggaga tgatcgagga gaggctgaag 3720

acctacgccc acctgttcga cgacaaggtg atgaagcagc tgaagaggag gaggtacacc 3780

ggctggggca ggctgagcag gaagctgatc aacggcatca gggacaagca gagcggcaag 3840

accatcctgg acttcctgaa gagcgacggc ttcgccaaca ggaacttcat gcagctgatc 3900

cacgacgaca gcctgacctt caaggaggac atccagaagg cccaggtgag cggccagggc 3960

gacagcctgc acgagcacat cgccaacctg gccggcagcc cggccatcaa gaagggcatc 4020

ctgcagaccg tgaaggtggt ggacgagctg gtgaaggtga tgggcaggca caagccggag 4080

aacatcgtga tcgagatggc cagggagaac cagaccaccc agaagggcca gaagaacagc 4140

agggagagga tgaagaggat cgaggagggc atcaaggagc tgggcagcca gatcctgaag 4200

gagcacccgg tggagaacac ccagctgcag aacgagaagc tgtacctgta ctacctgcag 4260

aacggcaggg acatgtacgt ggaccaggag ctggacatca acaggctgag cgactacgac 4320

gtggaccaca tcgtgccgca gagcttcctg aaggacgaca gcatcgacaa caaggtgctg 4380

accaggagcg acaagaacag gggcaagagc gacaacgtgc cgagcgagga ggtggtgaag 4440

aagatgaaaa actactggag gcagctgctg aacgccaagc tgatcaccca gaggaagttc 4500

gacaacctga ccaaggccga gaggggcggc ctgagcgagc tggacaaggc cggcttcatt 4560

aaaaggcagc tggtggagac caggcagatc accaagcacg tggcccagat cctggacagc 4620

aggatgaaca ccaagtacga cgagaacgac aagctgatca gggaggtgaa ggtgatcacc 4680

ctgaagagca agctggtgag cgacttcagg aaggacttcc agttctacaa ggtgagggag 4740

atcaataatt accaccacgc ccacgacgcc tacctgaacg ccgtggtggg caccgccctg 4800

attaaaaagt acccgaagct ggagagcgag ttcgtgtacg gcgactacaa ggtgtacgac 4860

gtgaggaaga tgatcgccaa gagcgagcag gagatcggca aggccaccgc caagtacttc 4920

ttctacagca acatcatgaa cttcttcaag accgagatca ccctggccaa cggcgagatc 4980

aggaagaggc cgctgatcga gaccaacggc gagaccggcg agatcgtgtg ggacaagggc 5040

agggacttcg ccaccgtgag gaaggtgctg tccatgccgc aggtgaacat cgtgaagaag 5100

accgaggtgc agaccggcgg cttcagcaag gagagcatcc tgccgaagag gaacagcgac 5160

aagctgatcg ccaggaagaa ggactgggac ccgaagaagt acggcggctt cgacagcccg 5220

accgtggcct acagcgtgct ggtggtggcc aaggtggaga agggcaagag caagaagctg 5280

aagagcgtga aggagctggt gggcatcacc atcatggaga ggagcagctt cgagaagaac 5340

ccagtggact tcctggaggc caagggctac aaggaggtga agaaggacct gatcattaaa 5400

ctgccgaagt acagcctgtt cgagctggag aacggcagga agaggatgct ggccagcgcc 5460

ggcgagctgc agaagggcaa cgagctggcc ctgccgagca agtacgtgaa cttcctgtac 5520

ctggccagcc actacgagaa gctgaagggc agcccggagg acaacgagca gaagcagctg 5580
ttcgtggagc agcacaagca ctacctggac gagatcatcg agcagatcag cgagttcagc 5640
aagagggtga tcctggccga cgccaacctg gacaaggtgc tgagcgccta caacaagcac 5700
agggacaagc cgatcaggga gcaggccgag aacatcatcc acctgttcac cctgaccaac 5760
ctgggcgccc cggccgcctt caagtacttc gacaccacca tcgacaggaa gaggtacacc 5820
agcaccaagg aggtgctgga cgccaccctg atccaccaga gcatcaccgg cctgtacgag 5880
accaggatcg acctgagcca gctgggcggc gacagcagcc cgccgaagaa gaagaggaag 5940
gtgagctgga aggacgccag cggctggagc aggatgtgag ctcaattaac tttgaattcc 6000
cttcgattca tccggcgcgg tgggctatgg acctgcagca gcaagctaat taagtttata 6060
tatattgcat gagagagcat gcaccgctaa ccatatatac tactgagact tctgaattct 6120
agtatatgta atccttttgt ttgggtttag gaggcaattc taatcatgta tgccgaattc 6180
caaagagtgg aaaacaagca aaatgttaaa tatacatgcc attttcggag gcaatttttt 6240
tcatgagggc atgttgctat aattccgggg accttggact tcttggagca ccttcctgtg 6300
acttaggcat acatgattag attataatcc aattagttaa gtcatagaaa attacctcat 6360
tctcatctcc atctccattt ctctatttct tctcaatcaa ggaccaaaat agcacttttg 6420
ctaaaaaaca agttagattg caaaccaaag tgcacaatac atagtaaaag gtatatgcaa 6480
catatttgaa tactcaaacc tctcatactt acattttcca tcattttgtt ccatttagcc 6540
tgtttgagct cggggttgga ctccaaaacc tcatgtcaac ataacttgat ccttttagca 6600
aactatgagc tctaacacca tacaatggtc aacaagaact attccaaaca taggaatgac 6660
ccaaactaca agtcaaagta tacttagctc tttgggcact tacaggttct aactttgata 6720
attctgtact tcttgtgacc atgactctgc tcgagctagg atcttgagcc ttatgactta 6780
aacaattaaa ccacaaacat tacctcaatg gttgtaagcc acgtccatat atcacagact 6840
tcaatgcatt cagactattc acagcttgac caaccttgac ctcttgcaag aacctcttct 6900
tctttgtgac cttaggtact ttagtcttct tgaccttctc ccttgctctt cataccttga 6960
agtccttctt gccttcacct tagttcaatc agctatctcc aagtcatgca cattgagttc 7020
cacttagtca atgtccatcc ttcaacttga cttgtgatgt ccacaattca tagtcatctc 7080
agtctatggg tccatcatgc ttgactccat gtgatgaacc ttgtaaggtt ttcactaagt 7140
acatgctcag acctttaatt gtgttgccat ccaaaaaaac caaaacctag attggaccat 7200
tcattatatt catcaatcat tgtacttgca agagtgatca aggtcatatt atttctctca 7260
actactccat tttgttgagg ggtgtcagtt gtggagactt cttgtttgat cccaacctca 7320
tcacaatact catgaatata gttgttgtca aattcatttc cattgtcact tcttattttt 7380
cttgattttg caatcaaact cattttgtac tttcatggta aatttattca atgttgatgc 7440
aacttttgac ttttcttgaa gaaagaacac tcaattacat ctagagaaat catcaacaac 7500
gaccaaacaa tacaggtttc ccccaacact agcatattat gtaggaccaa ataaatccat 7560

gtgaagtaac tctagtggtc ttggtgttga cataaaagcg tttgtaggat gtgtattggc 7620

aacttgtttt ccagcttgac atgcactata aaagattttc ctttttcaaa cacaacatct 7680

ttcaaatctc taaccatttc tttctttgga agcttcttgt tggggaaatg atccccggac 7740

cctaggaccc accggtcaga gagcgcgagg aagagccccc ggtcgctggg acccgttggt 7800

ccgctggaaa atgtggttac gtcaaccctg aaagaacccg cccctggttg agccccgtgg 7860

caccgagcct agggtcgagc gcggtggaat ctgacaggag gggccagaca tgttggaggg 7920

gaaccactca agtggatccc gcgcctggcc ccagaatgac ccgtcattaa tacccaacca 7980

cattaaccat gcctggcacc gagccatagc acggacgtcg gtccacttcc cactcatgac 8040

ctacgaacca gttgggctgc atagcactca tgaccgatag gttgaaggct tggcttcgca 8100

gagtgaaagg cgctgcatac atgtgaaggc tcgacttctt tttcttttcc tttcttttct 8160

tttctatttt taggtttcca atttaaattc caatttttt gtggagttca tatttggatc 8220

aaatagacaa attcacctat cagtatgaat agatgcattt attttgttta tatctatttt 8280

cttcatattt atatagtatt tcccttattc tttatatcat tttcaatttg taattggtaa 8340

gtttggtctt aaattcccca tttgggcact aatatatttt tattaatatt attattatta 8400

ttattattat tatttataga tgcacaaaca cataaactcc gacatgatgc atagattatt 8460

ttagatgtca ctagttaatg gttcacttta aatatggtta ttcccatgtt ctaatgagta 8520

gagggcaaag catatattga ggtcaactct ttccttatta tttacaaatt ggggaaattc 8580

tattcataac tcttcttctc tctcccaagt agcttaatct tcaccatggt gatttcattg 8640

cactttgcac attttgatca ctttattcct tgtaacccga gtcaaagtgt caatgatctt 8700

gataggatac tccgtgcagg ttagatcacc ttgcacactg agttcttcca ttggtaactg 8760

ttcctctggc ggaccgttcg aagggatctt taaacatacg aacagatcac ttaaagttct 8820

tctgaagcaa cttaaagtta tcaggcatgc atggatcttg gaggaatcag atgtgcagtc 8880

agggaccata gcacaggaca ggcgtcttct actggtgcta ccagcaaatg ctggaagccg 8940

ggaacactgg gtacgttgga aaccacgtga tgtggagtaa gataaactgt aggagaaaag 9000

catttcgtag tgggccatga agcctttcag gacatgtatt gcagtatggg ccggcccatt 9060

acgcaattgg acgacaacaa agactagtat tagtaccacc tcggctatcc acatagatca 9120

aagctggttt aaaagagttg tgcagatgat ccgtggcagc tggagctgag cttccggggt 9180

tttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt gaaaaagtgg 9240

caccgagtcg gtgctttttt tttcggaccg cgcctgcagt gcagcgtgac ccggtcgtgc 9300

ccctctctag agataatgag cattgcatgt ctaagttata aaaaattacc acatattttt 9360

tttgtcacac ttgtttgaag tgcagtttat ctatctttat acatatattt aaactttact 9420

ctacgaataa tataatctat agtactacaa taatatcagt gttttagaga atcatataaa 9480

tgaacagtta gacatggtct aaaggacaat tgagtatttt gacaacagga ctctacagtt 9540

ttatcttttt agtgtgcatg tgttctcctt ttttttgca aatagcttca cctatataat 9600

acttcatcca ttttattagt acatccattt agggtttagg gttaatggtt tttatagact 9660
aatttttta gtacatctat tttattctat tttagcctct aaattaagaa aactaaaact 9720
ctattttagt ttttttattt aataatttag atataaaata gaataaaata aagtgactaa 9780
aaattaaaca aataccttt aagaaattaa aaaaactaag gaaacatttt tcttgtttcg 9840
agtagataat gccagcctgt taaacgccgt cgacgagtct aacggacacc aaccagcgaa 9900
ccagcagcgt cgcgtcgggc caagcgaagc agacggcacg gcatctctgt cgctgcctct 9960
ggacccctct cgagagttcc gctccaccgt tggacttgct ccgctgtcgg catccagaaa 10020
ttgcgtggcg gagcggcaga cgtgagccgg cacggcaggc ggcctcctcc tcctctcacg 10080
gcaccggcag ctacggggga ttcctttccc accgctcctt cgctttccct tcctcgcccg 10140
ccgtaataaa tagacacccc ctccacaccc tctttcccca acctcgtgtt gttcggagcg 10200
cacacacaca caaccagatc tcccccaaat ccacccgtcg gcacctccgc ttcaaggtac 10260
gccgctcgtc ctcccccccc cccctctcta ccttctctag atcggcgttc cggtccatgg 10320
ttagggcccg gtagttctac ttctgttcat gtttgtgtta gatccgtgtt tgtgttagat 10380
ccgtgctgct agcgttcgta cacggatgcg acctgtacgt cagacacgtt ctgattgcta 10440
acttgccagt gtttctcttt ggggaatcct gggatggctc tagccgttcc gcagacggga 10500
tcgatttcat gatttttttt gtttcgttgc atagggtttg gtttgccctt ttcctttatt 10560
tcaatatatg ccgtgcactt gtttgtcggg tcatcttttc atgcttttt ttgtcttggt 10620
tgtgatgatg tggtctggtt gggcggtcgt tctagatcgg agtagaattc tgtttcaaac 10680
tacctggtgg atttattaat tttggatctg tatgtgtgtg ccatacatat tcatagttac 10740
gaattgaaga tgatggatgg aaatatcgat ctaggatagg tatacatgtt gatgcgggtt 10800
ttactgatgc atatacagag atgctttttg ttcgcttggt tgtgatgatg tggtgtggtt 10860
gggcggtcgt tcattcgttc tagatcggag tagaatactg tttcaaacta cctggtgtat 10920
ttattaattt tggaactgta tgtgtgtgtc atacatcttc atagttacga gtttaagatg 10980
gatggaaata tcgatctagg ataggtatac atgttgatgt gggttttact gatgcatata 11040
catgatggca tatgcagcat ctattcatat gctctaacct tgagtaccta tctattataa 11100
taaacaagta tgttttataa ttattttgat cttgatatac ttggatgatg gcatatgcag 11160
cagctatatg tggatttttt tagccctgcc ttcatacgct atttatttgc ttggtactgt 11220
ttcttttgtc gatgctcacc ctgttgtttg gtgttacttc tgcagggatc cggcagcagc 11280
catgcagaag ctgatcaaca gcgtgcagaa ctacgcctgg ggcagcaaga ccgccctgac 11340
cgagctgtac ggcatggaga accccagcag ccagcccatg gccgagctgt ggatgggcgc 11400
ccaccccaag agcagcagcc gcgtgcagaa cgccgccggc gacatcgtga gcctgcgcga 11460
cgtgatcgag agcgacaaga gcaccctgct gggcgaggcc gtggccaagc gcttcggcga 11520
gctgcccttc ctgttcaagg tgctgtgcgc cgcccagccc ctgagcatcc aggtgcaccc 11580
caacaagcac aacagcgaga tcggcttcgc caaggagaac gccgccggca tccccatgga 11640

cgccgccgag cgcaactaca aggaccccaa ccacaagccc gagctggtgt tcgccctgac 11700
ccccttcctg gccatgaacg ccttccgcga gttcagcgag atcgtgagcc tgctgcagcc 11760
cgtggccggc gcccaccccg ccatcgccca cttcctgcag cagcccgacg ccgagcgcct 11820
gagcgagctg ttcgccagcc tgctgaacat gcagggcgag gagaagagcc gcgccctggc 11880
catcctgaag agcgccctgg acagccagca gggcgagccc tggcagacca tccgcctgat 11940
cagcgagttc taccccgagg acagcggcct gttcagcccc ctgctgctga acgtggtgaa 12000
gctgaacccc ggcgaggcca tgttcctgtt cgccgagacc ccccacgcct acctgcaggg 12060
cgtggccctg gaggtgatgg ccaacagcga caacgtgctg cgcgccggcc tgaccccaa 12120
gtacatcgac atccccgagc tggtggccaa cgtgaagttc gaggccaagc ccgccaacca 12180
gctgctgacc cagcccgtga agcagggcgc cgagctggac ttccccatcc ccgtggacga 12240
cttcgccttc agcctgcacg acctgagcga caaggagacc accatcagcc agcagagcgc 12300
cgccatcctg ttctgcgtgg agggcgacgc caccctgtgg aagggcagcc agcagctgca 12360
gctgaagccc ggcgagagcg ccttcatcgc cgccaacgag agccccgtga ccgtgaaggg 12420
ccacggccgc ctggcccgcg tgtacaacaa gctgtgatag gagctcgatc cgtcgacctg 12480
cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctgtt gccggtcttg 12540
cgatgattat catataattt ctgttgaatt acgttaagca tgtaataatt aacatgtaat 12600
gcatgacgtt atttatgaga tgggttttta tgattagagt cccgcaatta tacatttaat 12660
acgcgataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc gcggtgtcat 12720
ctatgttact agatcggcgc gccgcaattg aagtttgggc ggccagcatg gccgtatccg 12780
caatgtgtta ttaagttgtc taagcgtcaa tttgtttaca ccacaatata tcctgccacc 12840
agccagccaa cagctccccg accggcagct cggcacaaaa tcaccactcg atacaggcag 12900
cccatcagaa ttaattctca tgtttgacag cttatcatcg actgcacggt gcaccaatgc 12960
ttctggcgtc aggcagccat cggaagctgt ggtatggctg tgcaggtcgt aaatcactgc 13020
ataattcgtg tcgctcaagg cgcactcccg ttctggataa tgtttttttgc gccgacatca 13080
taacggttct ggcaaatatt ctgaaatgag ctgttgacaa ttaatcatcc ggctcgtata 13140
atgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagaccatg agggaagcgt 13200
tgatcgccga agtatcgact caactatcag aggtagttgg cgtcatcgag cgccatctcg 13260
aaccgacgtt gctggccgta catttgtacg gctccgcagt ggatggcggc ctgaagccac 13320
acagtgatat tgatttgctg gttacggtga ccgtaaggct tgatgaaaca acgcggcgag 13380
ctttgatcaa cgacctttttg gaaacttcgg cttcccctgg agagagcgag attctccgcg 13440
ctgtagaagt caccattgtt gtgcacgacg acatcattcc gtggcgttat ccagctaagc 13500
gcgaactgca atttggagaa tggcagcgca atgacattct tgcaggtatc ttcgagccag 13560
ccacgatcga cattgatctg gctatcttgc tgacaaaagc aagagaacat agcgttgcct 13620
tggtaggtcc agcggcggag gaactctttg atccggttcc tgaacaggat ctatttgagg 13680

cgctaaatga aaccttaacg ctatggaact cgccgcccga ctgggctggc gatgagcgaa 13740
atgtagtgct tacgttgtcc cgcatttggt acagcgcagt aaccggcaaa atcgcgccga 13800
aggatgtcgc tgccgactgg gcaatggagc gcctgccggc ccagtatcag cccgtcatac 13860
ttgaagctag gcaggcttat cttggacaag aagatcgctt ggcctcgcgc gcagatcagt 13920
tggaagaatt tgttcactac gtgaaaggcg agatcaccaa agtagtcggc aaataaagct 13980
ctagtggatc tccgtacccg gggatctggc tcgcggcgga cgcacgacgc cggggcgaga 14040
ccataggcga tctcctaaat caatagtagc tgtaacctcg aagcgtttca cttgtaacaa 14100
cgattgagaa tttttgtcat aaaattgaaa tacttggttc gcatttttgt catccgcggt 14160
cagccgcaat tctgacgaac tgcccattta gctggagatg attgtacatc cttcacgtga 14220
aaatttctca agcgctgtga acaagggttc agattttaga ttgaaaggtg agccgttgaa 14280
acacgttctt cttgtcgatg acgacgtcgc tatgcggcat cttattattg aataccttac 14340
gatccacgcc ttcaaagtga ccgcggtagc cgacagcacc cagttcacaa gagtactctc 14400
ttccgcgacg gtcgatgtcg tggttgttga tctagattta ggtcgtgaag atgggctcga 14460
gctaggagca agtgatttta tcgctaagcc gttcagtatc agagagtttc tagcacgcat 14520
tcgggttgcc ttgcgcgtgc gccccaacgt tgtccgctcc aaagaccgac ggtctttttg 14580
ttttactgac tggacactta atctcaggca acgtcgcttg atgtccgaag ctggcggtga 14640
ggtgaaactt acggcaggtg agttcaatct tctcctcgcg tttttagaga aaccccgcga 14700
cgttctatcg cgcgagcaac ttctcattgc cagtcgagta cgcgacgagg aggtttatga 14760
caggagtata gatgttctca ttttgaggct gcgccgcaaa cttgaggcag atccgtcaag 14820
ccctcaactg ataaaaacag caagaggtgc cggttatttc tttgacgcgg acgtgcaggt 14880
ttcgcacggg gggacgatgg cagcctgagc caattcccag atccccgagg aatcggcgtg 14940
agcggtcgca aaccatccgg cccggtacaa atcggcgcgg cgctgggtga tgacctggtg 15000
gagaagttga aggccgcgca ggccgcccag cggcaacgca tcgaggcaga agcacgcccc 15060
ggtgaatcgt ggcaagcggc cgctgatcga atccgcaaag aatcccggca accgccggca 15120
gccggtgcgc cgtcgattag gaagccgccc aagggcgacg agcaaccaga ttttttcgtt 15180
ccgatgctct atgacgtggg cacccgcgat agtcgcagca tcatggacgt ggccgttttc 15240
cgtctgtcga agcgtgaccg acgagctggc gaggtgatcc gctacgagct tccagacggg 15300
cacgtagagg tttccgcagg gccggccggc atggccagtg tgtgggatta cgacctggta 15360
ctgatggcgg tttcccatct aaccgaatcc atgaaccgat accgggaagg gaagggagac 15420
aagcccggcc gcgtgttccg tccacacgtt gcggacgtac tcaagttctg ccggcgagcc 15480
gatggcggaa agcagaaaga cgacctggta gaaacctgca ttcggttaaa caccacgcac 15540
gttgccatgc agcgtacgaa gaaggccaag aacggccgcc tggtgacggt atccgagggt 15600
gaagccttga ttagccgcta caagatcgta aagagcgaaa ccgggcggcc ggagtacatc 15660
gagatcgagc tagctgattg gatgtaccgc gagatcacag aaggcaagaa cccggacgtg 15720

ctgacggttc accccgatta ctttttgatc gatcccggca tcggccgttt tctctaccgc 15780
ctggcacgcc gcgccgcagg caaggcagaa gccagatggt tgttcaagac gatctacgaa 15840
cgcagtggca gcgccggaga gttcaagaag ttctgtttca ccgtgcgcaa gctgatcggg 15900
tcaaatgacc tgccggagta cgatttgaag gaggaggcgg ggcaggctgg cccgatccta 15960
gtcatgcgct accgcaacct gatcgagggc gaagcatccg ccggttccta atgtacggag 16020
cagatgctag ggcaaattgc cctagcaggg gaaaaaggtc gaaaaggtct ctttcctgtg 16080
gatagcacgt acattgggaa cccaaagccg tacattggga accggaaccc gtacattggg 16140
aacccaaagc cgtacattgg gaaccggtca cacatgtaag tgactgatat aaaagagaaa 16200
aaaggcgatt tttccgccta aaactcttta aaacttatta aaactcttaa aacccgcctg 16260
gcctgtgcat aactgtctgg ccagcgcaca gccgaagagc tgcaaaaagc gcctaccctt 16320
cggtcgctgc gctccctacg ccccgccgct tcgcgtcggc ctatcgcggc cgctggccgc 16380
tcaaaaatgg ctggcctacg gccaggcaat ctaccagggc gcggacaagc cgcgccgtcg 16440
ccactcgacc gccggcgctg aggtctgcct cgtgaagaag gtgttgctga ctcataccag 16500
gcctgaatcg ccccatcatc cagccagaaa gtgagggagc cacggttgat gagagctttg 16560
ttgtaggtgg accagttggt gattttgaac ttttgctttg ccacggaacg gtctgcgttg 16620
tcgggaagat gcgtgatctg atccttcaac tcagcaaaag ttcgatttat tcaacaaagc 16680
cgccgtcccg tcaagtcagc gtaatgctct gccagtgtta caaccaatta accaattctg 16740
attagaaaaa ctcatcgagc atcaaatgaa actgcaattt attcatatca ggattatcaa 16800
taccatattt ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc 16860
ataggatggc aagatcctgg tatcggtctg cgattccgac tcgtccaaca tcaatacaac 16920
ctattaattt cccctcgtca aaaataaggt tatcaagtga gaaatcacca tgagtgacga 16980
ctgaatccgg tgagaatggc aaaagctctg cattaatgaa tcggccaacg cgcggggaga 17040
ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc 17100
gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa 17160
tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 17220
aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa 17280
aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 17340
cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 17400
tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 17460
agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 17520
gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta 17580
tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct 17640
acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc 17700
tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 17760

caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa 17820
aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 17880
aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 17940
ttgatccgga atta 17954

<210> 36
<211> 17045
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 24079

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<220>
<221> vùng khởi động
<222> (312)..(2356)
<223> prGRMZM2G146551-01
<220>
<221> gen
<222> (2358)..(6527)
<223> cCas9-01
<220>
<221> đột biến
<222> (5847)..(5849)
<223> đột biến L thành V
<220>
<221> đột biến
<222> (5892)..(5894)
<223> đột biến I thành V
<220>
<221> trình tự kết thúc
<222> (6542)..(7860)

<223> tGRMZM2G146551-01

<220>

<221> vùng khởi động

<222> (7874)..(8248)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (8249)..(8354)

<223> rsgRNAZmVLHP-02

<220>

<221> đặc điểm_hỗn tạp

<222> (8250)..(8269)

<223> đích ZmVLHP2

<220>

<221> đặc điểm_hỗn tạp

<222> (8270)..(8281)

<223> rCrRNA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (8286)..(8354)

<223> rTracrRNA-01

<220>

<221> vùng khởi động

<222> (8365)..(10356)

<223> prUbi1-04

<220>

<221> gen

<222> (10373)..(11551)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (11574)..(11826)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp
<222> (11870)..(11999)
<223> bNLB-03
<220>
<221> gen
<222> (12279)..(13067)
<223> cSpec-03
<220>
<221> vùng khởi động
<222> (13162)..(13292)
<223> prVirG-01
<220>
<221> gen
<222> (13367)..(13999)
<223> cVirG-09
<220>
<221> gen
<222> (14029)..(15102)
<223> cRepA-01
<220>
<221> đặc điểm_hỗn tạp
<222> (15145)..(15549)
<223> oVS1-02
<220>
<221> đặc điểm_hỗn tạp
<222> (16227)..(17033)
<223> oCOLE-06
<400> 36

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60
taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120
tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180
attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240
aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300
taccgggacc catgtagtat cacatgagtg agtcaaggac taagtattat gcattttgtt 360

tctcactcac ggattagctc gcaatcatca tagtgaaatc tagctactgg cactatcgaa 420
atctagctct ttgccgagtg cactttatcg agcactcgac aaagcattct ttatcgagtg 480
ccagtcttgg cgaaataaga ctctcgacaa agaccttgtt taccgaggga gaaacactcg 540
gcgtaaaaag acactcggca aagaagactt tgctgagtgt caaaccctca gcgaaatgcg 600
accctcggca aaggaccgtc agcagccatc tatagttgat ggctattaac ttcgcgagtg 660
tcaggcgttg acacacgaca aaatatcttt tttgtcgagt gtcactgggc aaacacttgg 720
taaacctatg ttttgccgag tgtctttcct tgacactcga caaagtatat ttgtttttc 780
ttttttcccca aacttttttgt ggtgtgtttc tacaatatat agacctattt gttcaatttt 840
ggcacaatta taaaagtgtt tgctataact atcagattta gtttgcttaa ttggatttct 900
ttggataatt cagatttgaa ctacaagcca cttgaaaaat ggaaaacagt gaatacaaaa 960
atgacattca tgttatttag cacaagttat gatctatttc aggaacatgc gagaattttc 1020
gaacaccatt ctcacaaaac atgattgcgg acttgtgatc aagttgtttt aaaattgtat 1080
aaaacaaaaa caaagtcaga aaatcatgaa acttgttgac atgtcatgat atcatatgta 1140
gagactctaa taaaaatttg agattgtttc atgaaagttg tcacgcgcta tgtgtagaaa 1200
cctagcccgt ctacattgag gttctatgat ttcatgtgaa ggacatctag gcatcaatgt 1260
ttatgataat atcttatgtt tgtttggacg aaatattaaa aacaaataaa aaggggtcct 1320
tgatcacttt gacgagcatt gcactcagca aagggtgcct ttgctgagtg caatggtcat 1380
agaactcggt agaaaaacat acatagacat agggaaactt gctttaccgc gtgctatggc 1440
caagacactc ggcaaactag gctcctttgt cgagttccat cccaagcact cgacattgga 1500
actgcgactg ggcctcacag aagctttctt tgccgagtgc cactaagcga ggaactcgga 1560
tgctcagcaa aggctctgtc atcgtcacga tgtcttttgt ttgtcgtgta ccagttggca 1620
ctcggtaaag actttactga gtgcccgata gaaagtactc gacaaagaga ccgttgccaa 1680
cgtttggttc actgagggct ctttgctgcc ttttggactt gacaaagaag ccgtctccag 1740
tagtgtctcc tgggaggcgg gatttatgtt ttttcccgga gctctgtggg acatcatgga 1800
cggtccagtc tggtgatcta aaatagacgg tttgccaagc tcacagagaa gtctttaaga 1860
tcttccacga tgcacgcatg ctttaaggtt agttagtgtt tggtctgaaa aagcgtcaac 1920
aattaggaaa caagaactaa aattattaaa ggacagatca ggaagcatgc atgttcttct 1980
tctatagtgt gtgttgagcc tgagtttggc cttttaggct ttattagggg gctcacagtc 2040
taactaagga gttgtattga tgtgctgaca aatattatgt tcgatcgtca cagtgttctt 2100
atgcggatcg attaggcccg atcatggtga aataaactaa ccaccggtaa gcccgggcag 2160
ccctagagca tgcagcggcc tacgtgaagc ccgcacatcg catcgtcgtc cgtcaggcgc 2220
taacggccgg ccgctgcatg cgtcgccggc gaactctctg ctgagccacc cgtcctccct 2280
ataagtagct atcccagcac cgtcgtctat caaccacaca cagagcggca tttcgaataa 2340
cacaggtgag cgcgaccatg gacaagaagt acagcatcgg cctggacatc ggcaccaaca 2400

gcgtgggctg ggccgtgatc accgacgagt acaaggtgcc gagcaagaag ttcaaggtgc 2460
tgggcaacac cgacaggcac agcatcaaga agaacctgat cggcgccctg ctgttcgaca 2520
gcggcgagac cgccgaggcc accaggctga agaggaccgc caggaggagg tacaccagga 2580
ggaagaacag gatctgctac ctgcaggaga tcttcagcaa cgagatggcc aaggtggacg 2640
acagcttctt ccacaggctg gaggagagct tcctggtgga ggaggacaag aagcacgaga 2700
ggcacccgat cttcggcaac atcgtggacg aggtggccta ccacgagaag tacccgacca 2760
tctaccacct gaggaagaag ctggtggaca gcaccgacaa ggccgacctg aggctgatct 2820
acctggccct ggcccacatg atcaagttca ggggccactt cctgatcgag ggcgacctga 2880
acccggacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc tacaaccagc 2940
tgttcgagga gaacccgatc aacgccagcg gcgtggacgc caaggccatc ctgagcgcca 3000
ggctgagcaa gagcaggagg ctggagaacc tgatcgccca gctgccgggc gagaagaaga 3060
acggcctgtt cggcaacctg atcgccctga gcctgggcct gaccccgaac ttcaagagca 3120
acttcgacct ggccgaggac gccaagctgc agctgagcaa ggacacctac gacgacgacc 3180
tggacaacct gctggcccag atcggcgacc agtacgccga cctgttcctg gccgccaaga 3240
acctgagcga cgccatcctg ctgagcgaca tcctgagggt gaacaccgag atcaccaagg 3300
ccccgctgag cgccagcatg atcaagaggt acgacgagca ccaccaggac ctgaccctgc 3360
tgaaggccct ggtgaggcag cagctgccgg agaagtacaa ggagatcttc ttcgaccaga 3420
gcaagaacgg ctacgccggc tacatcgacg gcggcgccag ccaggaggag ttctacaagt 3480
tcatcaagcc gatcctggag aagatggacg gcaccgagga gctgctggtg aagctgaaca 3540
gggaggacct gctgaggaag cagaggacct tcgacaacgg cagcatcccg caccagatcc 3600
acctgggcga gctgcacgcc atcctgagga ggcaggagga cttctacccg ttcctgaagg 3660
acaacaggga gaagatcgag aagatcctga ccttccgcat cccgtactac gtgggcccgc 3720
tggccagggg caacagcagg ttcgcctgga tgaccaggaa gagcgaggag accatcaccc 3780
cgtggaactt cgaggaggtg gtggacaagg gcgccagcgc ccagagcttc atcgagagga 3840
tgaccaactt cgacaagaac ctgccgaacg agaaggtgct gccgaagcac agcctgctgt 3900
acgagtactt caccgtgtac aacgagctga ccaaggtgaa gtacgtgacc gagggcatga 3960
ggaagccggc cttcctgagc ggcgagcaga agaaggccat cgtggacctg ctgttcaaga 4020
ccaacaggaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag atcgagtgct 4080
tcgacagcgt ggagatcagc ggcgtggagg acaggttcaa cgccagcctg ggcacctacc 4140
acgacctgct gaagatcatc aaggacaagg acttcctgga caacgaggag aacgaggaca 4200
tcctggagga catcgtgctg accctgaccc tgttcgagga cagggagatg atcgaggaga 4260
ggctgaagac ctacgcccac ctgttcgacg acaaggtgat gaagcagctg aagaggagga 4320
ggtacaccgg ctggggcagg ctgagcagga agctgatcaa cggcatcagg gacaagcaga 4380
gcggcaagac catcctggac ttcctgaaga gcgacggctt cgccaacagg aacttcatgc 4440

agctgatcca cgacgacagc ctgaccttca aggaggacat ccagaaggcc caggtgagcg 4500
gccagggcga cagcctgcac gagcacatcg ccaacctggc cggcagcccg gccatcaaga 4560
agggcatcct gcagaccgtg aaggtggtgg acgagctggt gaaggtgatg ggcaggcaca 4620
agccggagaa catcgtgatc gagatggcca gggagaacca gaccacccag aagggccaga 4680
agaacagcag ggagaggatg aagaggatcg aggagggcat caaggagctg ggcagccaga 4740
tcctgaagga gcacccggtg gagaacaccc agctgcagaa cgagaagctg tacctgtact 4800
acctgcagaa cggcagggac atgtacgtgg accaggagct ggacatcaac aggctgagcg 4860
actacgacgt ggaccacatc gtgccgcaga gcttcctgaa ggacgacagc atcgacaaca 4920
aggtgctgac caggagcgac aagaacaggg gcaagagcga caacgtgccg agcgaggagg 4980
tggtgaagaa gatgaaaaac tactggaggc agctgctgaa cgccaagctg atcacccaga 5040
ggaagttcga caacctgacc aaggccgaga ggggcggcct gagcgagctg gacaaggccg 5100
gcttcattaa aaggcagctg gtggagacca ggcagatcac caagcacgtg gcccagatcc 5160
tggacagcag gatgaacacc aagtacgacg agaacgacaa gctgatcagg gaggtgaagg 5220
tgatcaccct gaagagcaag ctggtgagcg acttcaggaa ggacttccag ttctacaagg 5280
tgagggagat caataattac caccacgccc acgacgccta cctgaacgcc gtggtgggca 5340
ccgccctgat taaaaagtac ccgaagctgg agagcgagtt cgtgtacggc gactacaagg 5400
tgtacgacgt gaggaagatg atcgccaaga gcgagcagga gatcggcaag gccaccgcca 5460
agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc ctggccaacg 5520
gcgagatcag gaagaggccg ctgatcgaga ccaacggcga gaccggcgag atcgtgtggg 5580
acaagggcag ggacttcgcc accgtgagga aggtgctgtc catgccgcag gtgaacatcg 5640
tgaagaagac cgaggtgcag accggcggct tcagcaagga gagcatcctg ccgaagagga 5700
acagcgacaa gctgatcgcc aggaagaagg actgggaccc gaagaagtac ggcggcttcg 5760
acagcccgac cgtggcctac agcgtgctgg tggtggccaa ggtggagaag ggcaagagca 5820
agaagctgaa gagcgtgaag gagctggtgg gcatcaccat catggagagg agcagcttcg 5880
agaagaaccc agtggacttc ctggaggcca agggctacaa ggaggtgaag aaggacctga 5940
tcattaaact gccgaagtac agcctgttcg agctggagaa cggcaggaag aggatgctgg 6000
ccagcgccgg cgagctgcag aagggcaacg agctggccct gccgagcaag tacgtgaact 6060
tcctgtacct ggccagccac tacgagaagc tgaagggcag cccggaggac aacgagcaga 6120
agcagctgtt cgtggagcag cacaagcact acctggacga gatcatcgag cagatcagcg 6180
agttcagcaa gagggtgatc ctggccgacg ccaacctgga caaggtgctg agcgcctaca 6240
acaagcacag ggacaagccg atcagggagc aggccgagaa catcatccac ctgttcaccc 6300
tgaccaacct gggcgccccg gccgccttca agtacttcga caccaccatc gacaggaaga 6360
ggtacaccag caccaaggag gtgctggacg ccaccctgat ccaccagagc atcaccggcc 6420
tgtacgagac caggatcgac ctgagccagc tgggcggcga cagcagcccg ccgaagaaga 6480

agaggaaggt gagctggaag gacgccagcg gctggagcag gatgtgacca tggagctcta 6540
aactttgaat tcccttcgat tcatccggca cagcgggcta tggaccttca gcagcaagct 6600
aattaagttg gcagcatgca ccgctaacct tatatactac tgagacttcc aaattctagt 6660
atatgtaatc cttttgttcg ggttcatgat cgaattccaa agagtggaaa acaagcaaaa 6720
ggttaaatat acatgccatt tttggaggca ttttttttcat gagggcatgt ttcgatatat 6780
ggaccactaa atatacatat catttacttt cctacaaatt tgctacatcc ttggaaatgc 6840
atagtctgtc tccaagaaaa agatactctg attacatcac tagtacacac agcctctata 6900
gtggcggttc tagagacatt ttcactggcg cttttcagtg ccgccagtgt taggggccag 6960
tggaaatcgc catttccatt caataaccgc cagtggaaaa agcatttcca ctggcggttt 7020
tcttaagcaa ccgccagtgg aaatgtttcc cgtctttttt taaattttcg tactgaaatt 7080
tatatattta cacacacaaa catatatata tatatattga tattgataaa catgtagtat 7140
tgatactaaa agcaacatga aattaaattc tatcatacat ttatatacat caaagtcttg 7200
tttacaacca tgtatgcatc acacattata tacatcaaag ttttcactta agctctaata 7260
actatctcgg ctaagagata gtctactaat ttctgttagt attctaaact ctggcaaagc 7320
taatgttccg gaagcatcgt gatatttccc ttctgcggga atgacctctt tcaatatgaa 7380
tgtgcacagg tcctcaacta tgccatacaa tgcaccttca gtcaagttct ccgggcttcc 7440
tttttgaaat tgctgtaaag gaagtttata aacatcatct atttatactc aataataaca 7500
catttgcatc tttaatgaca taaatacata cgtgactatt actaataata ccttgccagg 7560
gttcgtgatg tatcgtccat tcattctcat aaactcgcac acgtagaacc cacataggac 7620
cgatccgggt ggttgcttgt ggcactacat aacgggagat tggttattta gttgcaacat 7680
tgtcctatgt acgtacatgt atgatatgta ttcataaatt cacatactta ctggccagtt 7740
ataatggatg tctagtggca caccttttt ggacgtgtcg tactttccac catgtagctt 7800
ataaaaccta aatgccctgt gatctcaaat agaatcacca tgttattcta caattctcat 7860
gggacccttc gaagggatct ttaaacatac gaacagatca cttaaagttc ttctgaagca 7920
acttaaagtt atcaggcatg catggatctt ggaggaatca gatgtgcagt cagggaccat 7980
agcacaggac aggcgtcttc tactggtgct accagcaaat gctggaagcc gggaacactg 8040
ggtacgttgg aaaccacgtg atgtggagta agataaactg taggagaaaa gcatttcgta 8100
gtgggccatg aagcctttca ggacatgtat tgcagtatgg gccggcccat tacgcaattg 8160
gacgacaaca aagactagta ttagtaccac ctcggctatc cacatagatc aaagctggtt 8220
taaaagagtt gtgcagatga tccgtggcag ctggagctga gcttccgggg ttttagagct 8280
agaaatagca agttaaaata aggctagtcc gttatcaact tgaaaaagtg gcaccgagtc 8340
ggtgcttttt ttttcggacc gcgcctgcag tgcagcgtga cccggtcgtg cccctctcta 8400
gagataatga gcattgcatg tctaagttat aaaaaattac cacatatttt ttttgtcaca 8460
cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac tctacgaata 8520

atataatcta tagtactaca ataatatcag tgttttagag aatcatataa atgaacagtt 8580
agacatggtc taaaggacaa ttgagtattt tgacaacagg actctacagt tttatctttt 8640
tagtgtgcat gtgttctcct ttttttttgc aaatagcttc acctatataa tacttcatcc 8700
attttattag tacatccatt tagggtttag ggttaatggt ttttatagac taatttttt 8760
agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac tctattttag 8820
tttttttatt taataattta gatataaaat agaataaaat aaagtgacta aaaattaaac 8880
aaataccctt taagaaatta aaaaaactaa ggaaacattt ttcttgtttc gagtagataa 8940
tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga accagcagcg 9000
tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg tcgctgcctc tggacccctc 9060
tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcatccagaa attgcgtggc 9120
ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctcctctcac ggcaccggca 9180
gctacggggg attcctttcc caccgctcct tcgctttccc ttcctcgccc gccgtaataa 9240
atagacaccc cctccacacc ctctttcccc aacctcgtgt tgttcggagc gcacacacac 9300
acaaccagat ctcccccaaa tccacccgtc ggcacctccg cttcaaggta cgccgctcgt 9360
cctccccccc ccccctctct accttctcta gatcggcgtt ccggtccatg gttagggccc 9420
ggtagttcta cttctgttca tgtttgtgtt agatccgtgt ttgtgttaga tccgtgctgc 9480
tagcgttcgt acacggatgc gacctgtacg tcagacacgt tctgattgct aacttgccag 9540
tgtttctctt tggggaatcc tgggatggct ctagccgttc cgcagacggg atcgatttca 9600
tgatttttt tgtttcgttg catagggttt ggtttgccct tttcctttat ttcaatatat 9660
gccgtgcact tgtttgtcgg gtcatctttt catgcttttt tttgtcttgg ttgtgatgat 9720
gtggtctggt tgggcggtcg ttctagatcg gagtagaatt ctgtttcaaa ctacctggtg 9780
gatttattaa ttttggatct gtatgtgtgt gccatacata ttcatagtta cgaattgaag 9840
atgatggatg gaaatatcga tctaggatag gtatacatgt tgatgcgggt tttactgatg 9900
catatacaga gatgcttttt gttcgcttgg ttgtgatgat gtggtgtggt tgggcggtcg 9960
ttcattcgtt ctagatcgga gtagaatact gtttcaaact acctggtgta tttattaatt 10020
ttggaactgt atgtgtgtgt catacatctt catagttacg agtttaagat ggatggaaat 10080
atcgatctag gataggtata catgttgatg tgggttttac tgatgcatat acatgatggc 10140
atatgcagca tctattcata tgctctaacc ttgagtacct atctattata ataaacaagt 10200
atgttttata attattttga tcttgatata cttggatgat ggcatatgca gcagctatat 10260
gtggattttt ttagccctgc cttcatacgc tatttatttg cttggtactg tttcttttgt 10320
cgatgctcac cctgttgttt ggtgttactt ctgcagggat ccggcagcag ccatgcagaa 10380
gctgatcaac agcgtgcaga actacgcctg gggcagcaag accgccctga ccgagctgta 10440
cggcatggag aaccccagca gccagcccat ggccgagctg tggatgggcg cccaccccaa 10500
gagcagcagc cgcgtgcaga acgccgccgg cgacatcgtg agcctgcgcg acgtgatcga 10560

gagcgacaag agcaccctgc tgggcgaggc cgtggccaag cgcttcggcg agctgccctt 10620
cctgttcaag gtgctgtgcg ccgcccagcc cctgagcatc caggtgcacc ccaacaagca 10680
caacagcgag atcggcttcg ccaaggagaa cgccgccggc atccccatgg acgccgccga 10740
gcgcaactac aaggacccca accacaagcc cgagctggtg ttcgccctga ccccccttcct 10800
ggccatgaac gccttccgcg agttcagcga gatcgtgagc ctgctgcagc ccgtggccgg 10860
cgcccacccc gccatcgccc acttcctgca gcagcccgac gccgagcgcc tgagcgagct 10920
gttcgccagc ctgctgaaca tgcagggcga ggagaagagc cgcgccctgg ccatcctgaa 10980
gagcgccctg gacagccagc agggcgagcc ctggcagacc atccgcctga tcagcgagtt 11040
ctaccccgag gacagcggcc tgttcagccc cctgctgctg aacgtggtga agctgaaccc 11100
cggcgaggcc atgttcctgt tcgccgagac cccccacgcc tacctgcagg gcgtggccct 11160
ggaggtgatg gccaacagcg acaacgtgct gcgcgccggc ctgaccccca agtacatcga 11220
catccccgag ctggtggcca acgtgaagtt cgaggccaag cccgccaacc agctgctgac 11280
ccagcccgtg aagcagggcg ccgagctgga cttccccatc cccgtggacg acttcgcctt 11340
cagcctgcac gacctgagcg acaaggagac caccatcagc cagcagagcg ccgccatcct 11400
gttctgcgtg gagggcgacg ccaccctgtg gaagggcagc cagcagctgc agctgaagcc 11460
cggcgagagc gccttcatcg ccgccaacga gagccccgtg accgtgaagg gccacggccg 11520
cctggcccgc gtgtacaaca agctgtgata ggagctcgat ccgtcgacct gcagatcgtt 11580
caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt gcgatgatta 11640
tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa tgcatgacgt 11700
tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa tacgcgatag 11760
aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca tctatgttac 11820
tagatcggcg cgccgcaatt gaagtttggg cggccagcat ggccgtatcc gcaatgtgtt 11880
attaagttgt ctaagcgtca atttgtttac accacaatat atcctgccac cagccagcca 11940
acagctcccc gaccggcagc tcggcacaaa atcaccactc gatacaggca gcccatcaga 12000
attaattctc atgtttgaca gcttatcatc gactgcacgg tgcaccaatg cttctggcgt 12060
caggcagcca tcggaagctg tggtatggct gtgcaggtcg taaatcactg cataattcgt 12120
gtcgctcaag gcgcactccc gttctggata atgtttttg cgccgacatc ataacggttc 12180
tggcaaatat tctgaaatga gctgttgaca attaatcatc cggctcgtat aatgtgtgga 12240
attgtgagcg gataacaatt tcacacagga aacagaccat gagggaagcg ttgatcgccg 12300
aagtatcgac tcaactatca gaggtagttg gcgtcatcga gcgccatctc gaaccgacgt 12360
tgctggccgt acatttgtac ggctccgcag tggatggcgg cctgaagcca cacagtgata 12420
ttgatttgct ggttacggtg accgtaaggc ttgatgaaac aacgcggcga gctttgatca 12480
acgacctttt ggaaacttcg gcttcccctg gagagagcga gattctccgc gctgtagaag 12540
tcaccattgt tgtgcacgac gacatcattc cgtggcgtta tccagctaag cgcgaactgc 12600

aatttggaga atggcagcgc aatgacattc ttgcaggtat cttcgagcca gccacgatcg 12660

acattgatct ggctatcttg ctgacaaaag caagagaaca tagcgttgcc ttggtaggtc 12720

cagcggcgga ggaactcttt gatccggttc ctgaacagga tctatttgag gcgctaaatg 12780

aaaccttaac gctatggaac tcgccgcccg actgggctgg cgatgagcga aatgtagtgc 12840

ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa aatcgcgccg aaggatgtcg 12900

ctgccgactg ggcaatggag cgcctgccgg cccagtatca gcccgtcata cttgaagcta 12960

ggcaggctta tcttggacaa gaagatcgct tggcctcgcg cgcagatcag ttggaagaat 13020

ttgttcacta cgtgaaaggc gagatcacca aagtagtcgg caaataaagc tctagtggat 13080

ctccgtaccc ggggatctgg ctcgcggcgg acgcacgacg ccggggcgag accataggcg 13140

atctcctaaa tcaatagtag ctgtaacctc gaagcgtttc acttgtaaca acgattgaga 13200

atttttgtca taaaattgaa atacttggtt cgcatttttg tcatccgcgg tcagccgcaa 13260

ttctgacgaa ctgcccattt agctggagat gattgtacat ccttcacgtg aaaatttctc 13320

aagcgctgtg aacaagggtt cagattttag attgaaaggt gagccgttga aacacgttct 13380

tcttgtcgat gacgacgtcg ctatgcggca tcttattatt gaatacctta cgatccacgc 13440

cttcaaagtg accgcggtag ccgacagcac ccagttcaca agagtactct cttccgcgac 13500

ggtcgatgtc gtggttgttg atctagattt aggtcgtgaa gatgggctcg agctaggagc 13560

aagtgatttt atcgctaagc cgttcagtat cagagagttt ctagcacgca ttcgggttgc 13620

cttgcgcgtg cgccccaacg ttgtccgctc caaagaccga cggtcttttt gttttactga 13680

ctggacactt aatctcaggc aacgtcgctt gatgtccgaa gctggcggtg aggtgaaact 13740

tacggcaggt gagttcaatc ttctcctcgc gtttttagag aaaccccgcg acgttctatc 13800

gcgcgagcaa cttctcattg ccagtcgagt acgcgacgag gaggtttatg acaggagtat 13860

agatgttctc attttgaggc tgcgccgcaa acttgaggca gatccgtcaa gccctcaact 13920

gataaaaaca gcaagaggtg ccggttattt ctttgacgcg gacgtgcagg tttcgcacgg 13980

ggggacgatg gcagcctgag ccaattccca gatccccgag gaatcggcgt gagcggtcgc 14040

aaaccatccg gcccggtaca aatcggcgcg gcgctgggtg atgacctggt ggagaagttg 14100

aaggccgcgc aggccgccca gcggcaacgc atcgaggcag aagcacgccc cggtgaatcg 14160

tggcaagcgg ccgctgatcg aatccgcaaa gaatcccggc aaccgccggc agccggtgcg 14220

ccgtcgatta ggaagccgcc caagggcgac gagcaaccag atttttttcgt tccgatgctc 14280

tatgacgtgg gcacccgcga tagtcgcagc atcatggacg tggccgtttt ccgtctgtcg 14340

aagcgtgacc gacgagctgg cgaggtgatc cgctacgagc ttccagacgg gcacgtagag 14400

gtttccgcag ggccggccgg catggccagt gtgtgggatt acgacctggt actgatggcg 14460

gtttcccatc taaccgaatc catgaaccga taccgggaag ggaagggaga caagcccggc 14520

cgcgtgttcc gtccacacgt tgcggacgta ctcaagttct gccggcgagc cgatggcgga 14580

aagcagaaag acgacctggt agaaacctgc attcggttaa acaccacgca cgttgccatg 14640

cagcgtacga agaaggccaa gaacggccgc ctggtgacgg tatccgaggg tgaagccttg 14700
attagccgct acaagatcgt aaagagcgaa accgggcggc cggagtacat cgagatcgag 14760
ctagctgatt ggatgtaccg cgagatcaca gaaggcaaga acccggacgt gctgacggtt 14820
caccccgatt acttttgat cgatcccggc atcggccgtt ttctctaccg cctggcacgc 14880
cgcgccgcag gcaaggcaga agccagatgg ttgttcaaga cgatctacga acgcagtggc 14940
agcgccggag agttcaagaa gttctgtttc accgtgcgca agctgatcgg gtcaaatgac 15000
ctgccggagt acgatttgaa ggaggaggcg gggcaggctg gcccgatcct agtcatgcgc 15060
taccgcaacc tgatcgaggg cgaagcatcc gccggttcct aatgtacgga gcagatgcta 15120
gggcaaattg ccctagcagg ggaaaaaggt cgaaaaggtc tctttcctgt ggatagcacg 15180
tacattggga acccaaagcc gtacattggg aaccggaacc cgtacattgg gaacccaaag 15240
ccgtacattg ggaaccggtc acacatgtaa gtgactgata taaaagagaa aaaaggcgat 15300
ttttccgcct aaaactcttt aaaacttatt aaaactctta aaacccgcct ggcctgtgca 15360
taactgtctg gccagcgcac agccgaagag ctgcaaaaag cgcctaccct tcggtcgctg 15420
cgctccctac gccccgccgc ttcgcgtcgg cctatcgcgg ccgctggccg ctcaaaaatg 15480
gctggcctac ggccaggcaa tctaccaggg cgcggacaag ccgcgccgtc gccactcgac 15540
cgccggcgct gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca ggcctgaatc 15600
gccccatcat ccagccagaa agtgagggag ccacggttga tgagagcttt gttgtaggtg 15660
gaccagttgg tgattttgaa cttttgcttt gccacggaac ggtctgcgtt gtcgggaaga 15720
tgcgtgatct gatccttcaa ctcagcaaaa gttcgattta ttcaacaaag ccgccgtccc 15780
gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt aaccaattct gattagaaaa 15840
actcatcgag catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt 15900
tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg 15960
caagatcctg gtatcggtct gcgattccga ctcgtccaac atcaatacaa cctattaatt 16020
tcccctcgtc aaaaataagg ttatcaagtg agaaatcacc atgagtgacg actgaatccg 16080
gtgagaatgg caaaagctct gcattaatga atcggccaac gcgcggggag aggcggtttg 16140
cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 16200
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 16260
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 16320
gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc 16380
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 16440
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt 16500
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg 16560
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 16620
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 16680

gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 16740
ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg 16800
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 16860
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 16920
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 16980
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttgatccgg 17040
aatta 17045

<210> 37
<211> 16776
<212> ADN
<213> trình tự nhân tạo
<220>
<223> vecto 24091

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (330)..(2417)
<223> prZmGRMZM2G471240-01
<220>
<221> gen
<222> (2420)..(6589)
<223> cCas9-01
<220>
<221> đột biến
<222> (5909)..(5911)
<223> đột biến L thành V
<220>
<221> đột biến
<222> (5954)..(5956)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc

<222> (6596)..(7591)

<223> tZmGRMZM2G471240-01

<220>

<221> vùng khởi động

<222> (7605)..(7979)

<223> prOsU3-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7981)..(8085)

<223> rsgRNAZmVLHP-02

<220>

<221> đặc điểm_hỗn tạp

<222> (7981)..(8000)

<223> xZmVLHP2

<220>

<221> đặc điểm_hỗn tạp

<222> (8001)..(8012)

<223> rCrRNA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (8017)..(8085)

<223> rTracrRNA-01

<220>

<221> vùng khởi động

<222> (8096)..(10087)

<223> prUbi1-04

<220>

<221> gen

<222> (10104)..(11282)

<223> cPMI-09

<220>

<221> trình tự kết thúc

<222> (11305)..(11557)

<223> tNOS-05-01

<220>

<221> đặc điểm_hỗn tạp

<222> (11601)..(11730)

<223> bNLB-03

<220>

<221> gen

<222> (12010)..(12798)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (12893)..(13023)

<223> prVirG-01

<220>

<221> gen

<222> (13098)..(13730)

<223> cVirG-09

<220>

<221> gen

<222> (13760)..(14833)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (14876)..(15280)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (15958)..(16764)

<223> oCOLE-06

<400> 37

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180
attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240
aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300
tacctcgcga atgcatctag atgggaccct atttgtactc attccatgtc tcataaactt 360
tgggcaccat ccatccaaca catccaatct aaacacacca aacgatgggg aatggaaaga 420
gcagtattcg attcaacaat ggcaaacaaa tatcactgaa ttagaccaag aataaaccta 480
attagacaac gacctcccaa ccatcattcg tcaggctgta aagaagataa agctgccttg 540
gggcatggat caagcagaac accagagatg aatccaaaca cacagaaaat cacgcgcgct 600
gtctacaatg acaacaagcc ccacatttca ttgcagtaca ctgggctaca aaggcacgta 660
caacaaagag ctagggaaac attgcggagg gcacgagaga gcagctaact tgacaatata 720
gcagactgag cttgcactgt tagcaggcga ggaagggaat catggggacg gagaatgggg 780
tccatgcccg cgaaggagaa ggcggacgcc gccacggtgg caccggcgca cgcgcacaca 840
gggaacccgc acaggcagcc aaggatgctg cctcgccatt gcgccggtcg tctctgccac 900
gctcctctct ctctcccgct gcatcgccgt ggatggggca agcagagagc agggactgcg 960
acgatctggg cggaggactc gccttggaga gcgcggacgc agacgggatt ctagggagag 1020
agcgaagacg gggcgcgcgc ggcgctcgcg cggcgtggtg gcggcgagat tagcgggggt 1080
gggggggaggg cggagccgtg gtgagggtgt ggacgccctc cttaccctct taagtagtag 1140
tagagatata atccgttcca aaatatccat ccgttcaatt tatatttcgt ttgatctttt 1200
taccctaaat ttgattgact catcttatta aaaaagttca taactattat taatctttat 1260
tgagatatca tttagcatat aatatacttt aagtgtggtt ttagattttt tttaaaaaaa 1320
aaaattcgca aaaattaaat gaaacgaccc aatcaaactt gaaaagtaaa actaattata 1380
aatttgaacg gaaggagtaa gaggatgttt gaatgtacta gagctaatag ttggttgctt 1440
taaaatttgc tagtagaatt agctagctaa taaatatcta gataactatt agctaatttg 1500
ctaaaacagc taatagttga actattagct agattgtttg gatgtattcg gctaatttta 1560
atggctaact attagctata gtacaatatt caaacacctc ctaattaaaa tggacaaata 1620
tctcttcttt tggtcccttg cgttagattt ttcatatctc cttatttagt ataaaagaat 1680
catcaaaaag tggacaaccc ctagtggaac accattttag tagtggttgc atgaaacctt 1740
tcgcgcacca gtttctatgt gtcactctaa aaatgggaca gcatgtacgt agtgcctata 1800
tatatacaag tcatctatcg ttgcctcctc agttcatcac taatcacact tattgtgccc 1860
tcgacgagta tctatagcta gctcattaat cgattcgggg gtgtgttgtc gaaggcggca 1920
ttggcgagct actcgtcgcg gcgtccaagc aatacctgta gcacgaaggc gatcgccggg 1980
agcgtggtcg gcgagcccgt cgtgctgggg cagagggtga cggtgctgac ggtggacggc 2040
ggcggcgtcc ggggtctcat cccgggaacc atcctcgcct tcctggaggc caggctgcag 2100
gagctggacg caccggaggc gaggctggcg gactacttcg actacatcgc cggaaccagc 2160

accggcggtc tcatcaccgc cttgctgacc gcgcccggca aggacaagcg gcctctctag 2220
gctgccaagg acatcaacca cttttacatc cataactgcc cgcgcatctt tcctcagaag 2280
tgagtccgat gctgccgcca ttgttcttgc atccatccag catcgtacgt acgtcctcta 2340
tacatctgcg gatcatcatg tgcgcatgtt tgtggcatgc atgcatgcat gtgagcagga 2400
gcaggcttgc gaaaacccca tggacaagaa gtacagcatc ggcctggaca tcggcaccaa 2460
cagcgtgggc tgggccgtga tcaccgacga gtacaaggtg ccgagcaaga agttcaaggt 2520
gctgggcaac accgacaggc acagcatcaa gaagaacctg atcggcgccc tgctgttcga 2580
cagcggcgag accgccgagg ccaccaggct gaagaggacc gccaggagga ggtacaccag 2640
gaggaagaac aggatctgct acctgcagga gatcttcagc aacgagatgg ccaaggtgga 2700
cgacagcttc ttccacaggc tggaggagag cttcctggtg gaggaggaca agaagcacga 2760
gaggcacccg atcttcggca acatcgtgga cgaggtggcc taccacgaga agtacccgac 2820
catctaccac ctgaggaaga agctggtgga cagcaccgac aaggccgacc tgaggctgat 2880
ctacctggcc ctggcccaca tgatcaagtt caggggccac ttcctgatcg agggcgacct 2940
gaacccggac aacagcgacg tggacaagct gttcatccag ctggtgcaga cctacaacca 3000
gctgttcgag gagaacccga tcaacgccag cggcgtggac gccaaggcca tcctgagcgc 3060
caggctgagc aagagcagga ggctggagaa cctgatcgcc cagctgccgg gcgagaagaa 3120
gaacggcctg ttcggcaacc tgatcgccct gagcctgggc ctgaccccga acttcaagag 3180
caacttcgac ctggccgagg acgccaagct gcagctgagc aaggacacct acgacgacga 3240
cctggacaac ctgctggccc agatcggcga ccagtacgcc gacctgttcc tggccgccaa 3300
gaacctgagc gacgccatcc tgctgagcga catcctgagg gtgaacaccg agatcaccaa 3360
ggccccgctg agcgccagca tgatcaagag gtacgacgag caccaccagg acctgaccct 3420
gctgaaggcc ctggtgaggc agcagctgcc ggagaagtac aaggagatct tcttcgacca 3480
gagcaagaac ggctacgccg gctacatcga cggcggcgcc agccaggagg agttctacaa 3540
gttcatcaag ccgatcctgg agaagatgga cggcaccgag gagctgctgg tgaagctgaa 3600
cagggaggac ctgctgagga agcagaggac cttcgacaac ggcagcatcc cgcaccagat 3660
ccacctgggc gagctgcacg ccatcctgag gaggcaggag gacttctacc cgttcctgaa 3720
ggacaacagg gagaagatcg agaagatcct gaccttccgc atcccgtact acgtgggccc 3780
gctggccagg ggcaacagca ggttcgcctg gatgaccagg aagagcgagg agaccatcac 3840
cccgtggaac ttcgaggagg tggtggacaa gggcgccagc gcccagagct tcatcgagag 3900
gatgaccaac ttcgacaaga acctgccgaa cgagaaggtg ctgccgaagc acagcctgct 3960
gtacgagtac ttcaccgtgt acaacgagct gaccaaggtg aagtacgtga ccgagggcat 4020
gaggaagccg gccttcctga gcggcgagca gaagaaggcc atcgtggacc tgctgttcaa 4080
gaccaacagg aaggtgaccg tgaagcagct gaaggaggac tacttcaaga agatcgagtg 4140
cttcgacagc gtggagatca gcggcgtgga ggacaggttc aacgccagcc tgggcaccta 4200

ccacgacctg ctgaagatca tcaaggacaa ggacttcctg gacaacgagg agaacgagga 4260
catcctggag gacatcgtgc tgaccctgac cctgttcgag gacagggaga tgatcgagga 4320
gaggctgaag acctacgccc acctgttcga cgacaaggtg atgaagcagc tgaagaggag 4380
gaggtacacc ggctggggca ggctgagcag gaagctgatc aacggcatca gggacaagca 4440
gagcggcaag accatcctgg acttcctgaa gagcgacggc ttcgccaaca ggaacttcat 4500
gcagctgatc cacgacgaca gcctgacctt caaggaggac atccagaagg cccaggtgag 4560
cggccagggc gacagcctgc acgagcacat cgccaacctg gccggcagcc cggccatcaa 4620
gaagggcatc ctgcagaccg tgaaggtggt ggacgagctg gtgaaggtga tgggcaggca 4680
caagccggag aacatcgtga tcgagatggc cagggagaac cagaccaccc agaagggcca 4740
gaagaacagc agggagagga tgaagaggat cgaggagggc atcaaggagc tgggcagcca 4800
gatcctgaag gagcacccgg tggagaacac ccagctgcag aacgagaagc tgtacctgta 4860
ctacctgcag aacggcaggg acatgtacgt ggaccaggag ctggacatca acaggctgag 4920
cgactacgac gtggaccaca tcgtgccgca gagcttcctg aaggacgaca gcatcgacaa 4980
caaggtgctg accaggagcg acaagaacag gggcaagagc gacaacgtgc cgagcgagga 5040
ggtggtgaag aagatgaaaa actactggag gcagctgctg aacgccaagc tgatcaccca 5100
gaggaagttc gacaacctga ccaaggccga gaggggcggc ctgagcgagc tggacaaggc 5160
cggcttcatt aaaaggcagc tggtggagac caggcagatc accaagcacg tggcccagat 5220
cctggacagc aggatgaaca ccaagtacga cgagaacgac aagctgatca gggaggtgaa 5280
ggtgatcacc ctgaagagca agctggtgag cgacttcagg aaggacttcc agttctacaa 5340
ggtgagggag atcaataatt accaccacgc ccacgacgcc tacctgaacg ccgtggtggg 5400
caccgccctg attaaaaagt acccgaagct ggagagcgag ttcgtgtacg gcgactacaa 5460
ggtgtacgac gtgaggaaga tgatcgccaa gagcgagcag gagatcggca aggccaccgc 5520
caagtacttc ttctacagca acatcatgaa cttcttcaag accgagatca ccctggccaa 5580
cggcgagatc aggaagaggc cgctgatcga gaccaacggc gagaccggcg agatcgtgtg 5640
ggacaagggc agggacttcg ccaccgtgag gaaggtgctg tccatgccgc aggtgaacat 5700
cgtgaagaag accgaggtgc agaccggcgg cttcagcaag gagagcatcc tgccgaagag 5760
gaacagcgac aagctgatcg ccaggaagaa ggactgggac ccgaagaagt acggcggctt 5820
cgacagcccg accgtggcct acagcgtgct ggtggtggcc aaggtggaga agggcaagag 5880
caagaagctg aagagcgtga aggagctggt gggcatcacc atcatggaga ggagcagctt 5940
cgagaagaac ccagtggact tcctggaggc caagggctac aaggaggtga agaaggacct 6000
gatcattaaa ctgccgaagt acagcctgtt cgagctggag aacggcagga agaggatgct 6060
ggccagcgcc ggcgagctgc agaagggcaa cgagctggcc ctgccgagca agtacgtgaa 6120
cttcctgtac ctggccagcc actacgagaa gctgaagggc agcccggagg acaacgagca 6180
gaagcagctg ttcgtggagc agcacaagca ctacctggac gagatcatcg agcagatcag 6240

cgagttcagc aagagggtga tcctggccga cgccaacctg gacaaggtgc tgagcgccta 6300
caacaagcac agggacaagc cgatcaggga gcaggccgag aacatcatcc acctgttcac 6360
cctgaccaac ctgggcgccc cggccgcctt caagtacttc gacaccacca tcgacaggaa 6420
gaggtacacc agcaccaagg aggtgctgga cgccaccctg atccaccaga gcatcaccgg 6480
cctgtacgag accaggatcg acctgagcca gctgggcggc gacagcagcc cgccgaagaa 6540
gaagaggaag gtgagctgga aggacgccag cggctggagc aggatgtgac catgggacaa 6600
gtggctttac tgtcagtcac atgcttgtaa ataagtagac tttattttaa taaaacataa 6660
aaatatatat atgttcttga atataaaatt gataaccaaa ttaaaattcg aaccatcact 6720
tatacataat tttactttat tttttataaa acgtgaacgg gaaggactac cgtgaatgac 6780
tatagaacca atcatactag tataaaatat atgatgacac tacgggagag acaaactttg 6840
tctggcgcta aatattttgc cgagtgtgaa ttcacgggca ctaggcaaag atcttctttg 6900
ccgagtgtta cgctgggcaa agtaagacac taggtaaatc agtcatttgc cgagtgtccg 6960
ccactaggca aagcaaaaca ctggcaaatc aaaagtttac ctagtgccag acactaggca 7020
aaaaaaaaac gctcggcaaa tcggaagttt ccctagtgcc agacactaga caaagaaaaa 7080
cacttgataa actagcgtcg tcagctaaca ccatccacca accgttaacg ttgccgagta 7140
tctgacttcg acactcggca aagaaggtct ctttgcctag tgtcggtctg gaacactagg 7200
caaagaggca ctttacctag tgtcgtattt tgacactcag taaaataatt ttttttcttt 7260
ctgcttccaa acttttttatg atgtgttcct atagcaccta gaactacatg tcaagttttg 7320
gtaaaatttt tgaagttttt gctatattta cttaatttat tttatttaat tgaatttctt 7380
ttgataattc aaatttgaac tcggcaaggt aagaagcgag ggtagcctgg aaacacactt 7440
tgcctagtgt tacactcggt acaggagcct cccctgccta gtgctgcact cgacaaaaga 7500
ttcgcctttg cctagcgctg cactcggcac aggagtcgcc tttgcctagt gctgcactag 7560
gcaaagcctc cgttaccgtg ccttccatcg tcggaccctt cgaagggatc tttaaacata 7620
cgaacagatc acttaaagtt cttctgaagc aacttaaagt tatcaggcat gcatggatct 7680
tggaggaatc agatgtgcag tcagggacca tagcacagga caggcgtctt ctactggtgc 7740
taccagcaaa tgctggaagc cgggaacact gggtacgttg gaaaccacgt gatgtggagt 7800
aagataaact gtaggagaaa agcatttcgt agtgggccat gaagcctttc aggacatgta 7860
ttgcagtatg ggccggccca ttacgcaatt ggacgacaac aaagactagt attagtacca 7920
cctcggctat ccacatagat caaagctggt ttaaaagagt tgtgcagatg atccgtggca 7980
gctggagctg agcttccggg gttttagagc tagaaatagc aagttaaaat aaggctagtc 8040
cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tttttcggac cgcgcctgca 8100
gtgcagcgtg acccggtcgt gcccctctct agagataatg agcattgcat gtctaagtta 8160
taaaaaatta ccacatattt tttttgtcac acttgtttga agtgcagttt atctatcttt 8220
atacatatat ttaaacttta ctctacgaat aatataatct atagtactac aataatatca 8280

gtgttttaga gaatcatata aatgaacagt tagacatggt ctaaaggaca attgagtatt 8340
ttgacaacag gactctacag ttttatcttt ttagtgtgca tgtgttctcc tttttttttg 8400
caaatagctt cacctatata atacttcatc cattttatta gtacatccat ttagggttta 8460
gggttaatgg tttttataga ctaatttttt tagtacatct attttattct attttagcct 8520
ctaaattaag aaaactaaaa ctctatttta gttttttttat ttaataattt agatataaaa 8580
tagaataaaa taaagtgact aaaaattaaa caaataccct ttaagaaatt aaaaaaacta 8640
aggaaacatt tttcttgttt cgagtagata atgccagcct gttaaacgcc gtcgacgagt 8700
ctaacggaca ccaaccagcg aaccagcagc gtcgcgtcgg gccaagcgaa gcagacggca 8760
cggcatctct gtcgctgcct ctggacccct ctcgagagtt ccgctccacc gttggacttg 8820
ctccgctgtc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc ggcacggcag 8880
gcggcctcct cctcctctca cggcaccggc agctacgggg gattcctttc ccaccgctcc 8940
ttcgctttcc cttcctcgcc cgccgtaata aatagacacc ccctccacac cctctttccc 9000
caacctcgtg ttgttcggag cgcacacaca cacaaccaga tctcccccaa atccacccgt 9060
cggcacctcc gcttcaaggt acgccgctcg tcctcccccc ccccctctc taccttctct 9120
agatcggcgt tccggtccat ggttagggcc cggtagttct acttctgttc atgtttgtgt 9180
tagatccgtg tttgtgttag atccgtgctg ctagcgttcg tacacggatg cgacctgtac 9240
gtcagacacg ttctgattgc taacttgcca gtgtttctct ttggggaatc ctgggatggc 9300
tctagccgtt ccgcagacgg gatcgatttc atgatttttt ttgtttcgtt gcatagggtt 9360
tggtttgccc ttttccttta tttcaatata tgccgtgcac ttgtttgtcg ggtcatcttt 9420
tcatgctttt ttttgtcttg gttgtgatga tgtggtctgg ttgggcggtc gttctagatc 9480
ggagtagaat tctgtttcaa actacctggt ggatttatta attttggatc tgtatgtgtg 9540
tgccatacat attcatagtt acgaattgaa gatgatggat ggaaatatcg atctaggata 9600
ggtatacatg ttgatgcggg ttttactgat gcatatacag agatgctttt tgttcgcttg 9660
gttgtgatga tgtggtgtgg ttgggcggtc gttcattcgt tctagatcgg agtagaatac 9720
tgtttcaaac tacctggtgt atttattaat tttggaactg tatgtgtgtg tcatacatct 9780
tcatagttac gagtttaaga tggatggaaa tatcgatcta ggataggtat acatgttgat 9840
gtgggtttta ctgatgcata tacatgatgg catatgcagc atctattcat atgctctaac 9900
cttgagtacc tatctattat aataaacaag tatgttttat aattattttg atcttgatat 9960
acttggatga tggcatatgc agcagctata tgtggatttt tttagccctg ccttcatacg 10020
ctatttattt gcttggtact gtttcttttg tcgatgctca ccctgttgtt tggtgttact 10080
tctgcaggga tccggcagca gccatgcaga agctgatcaa cagcgtgcag aactacgcct 10140
ggggcagcaa gaccgccctg accgagctgt acggcatgga gaaccccagc agccagccca 10200
tggccgagct gtggatgggc gcccacccca agagcagcag ccgcgtgcag aacgccgccg 10260
gcgacatcgt gagcctgcgc gacgtgatcg agagcgacaa gagcaccctg ctgggcgagg 10320

ccgtggccaa gcgcttcggc gagctgccct tcctgttcaa ggtgctgtgc gccgcccagc 10380
ccctgagcat ccaggtgcac cccaacaagc acaacagcga gatcggcttc gccaaggaga 10440
acgccgccgg catccccatg gacgccgccg agcgcaacta caaggacccc aaccacaagc 10500
ccgagctggt gttcgccctg acccccttcc tggccatgaa cgccttccgc gagttcagcg 10560
agatcgtgag cctgctgcag cccgtggccg gcgcccaccc cgccatcgcc cacttcctgc 10620
agcagcccga cgccgagcgc ctgagcgagc tgttcgccag cctgctgaac atgcagggcg 10680
aggagaagag ccgcgccctg gccatcctga agagcgccct ggacagccag cagggcgagc 10740
cctggcagac catccgcctg atcagcgagt tctaccccga ggacagcggc ctgttcagcc 10800
ccctgctgct gaacgtggtg aagctgaacc ccggcgaggc catgttcctg ttcgccgaga 10860
ccccccacgc ctacctgcag ggcgtggccc tggaggtgat ggccaacagc gacaacgtgc 10920
tgcgcgccgg cctgaccccc aagtacatcg acatccccga gctggtggcc aacgtgaagt 10980
tcgaggccaa gcccgccaac cagctgctga cccagcccgt gaagcagggc gccgagctgg 11040
acttccccat ccccgtggac gacttcgcct tcagcctgca cgacctgagc gacaaggaga 11100
ccaccatcag ccagcagagc gccgccatcc tgttctgcgt ggagggcgac gccaccctgt 11160
ggaagggcag ccagcagctg cagctgaagc ccggcgagag cgccttcatc gccgccaacg 11220
agagccccgt gaccgtgaag ggccacggcc gcctggcccg cgtgtacaac aagctgtgat 11280
aggagctcga tccgtcgacc tgcagatcgt tcaaacattt ggcaataaag tttcttaaga 11340
ttgaatcctg ttgccggtct tgcgatgatt atcatataat ttctgttgaa ttacgttaag 11400
catgtaataa ttaacatgta atgcatgacg ttatttatga gatgggtttt tatgattaga 11460
gtcccgcaat tatacattta atacgcgata gaaaacaaaa tatagcgcgc aaactaggat 11520
aaattatcgc gcgcggtgtc atctatgtta ctagatcggc gcgccgcaat tgaagtttgg 11580
gcggccagca tggccgtatc cgcaatgtgt tattaagttg tctaagcgtc aatttgttta 11640
caccacaata tatcctgcca ccagccagcc aacagctccc cgaccggcag ctcggcacaa 11700
aatcaccact cgatacaggc agcccatcag aattaattct catgtttgac agcttatcat 11760
cgactgcacg gtgcaccaat gcttctggcg tcaggcagcc atcggaagct gtggtatggc 11820
tgtgcaggtc gtaaatcact gcataattcg tgtcgctcaa ggcgcactcc cgttctggat 11880
aatgtttttt gcgccgacat cataacggtt ctggcaaata ttctgaaatg agctgttgac 11940
aattaatcat ccggctcgta taatgtgtgg aattgtgagc ggataacaat ttcacacagg 12000
aaacagacca tgagggaagc gttgatcgcc gaagtatcga ctcaactatc agaggtagtt 12060
ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta cggctccgca 12120
gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt gaccgtaagg 12180
cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc ggcttccccc 12240
ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga cgacatcatt 12300
ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg caatgacatt 12360

cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt gctgacaaaa 12420
gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt tgatccggtt 12480
cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa ctcgccgccc 12540
gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg gtacagcgca 12600
gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga gcgcctgccg 12660
gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca agaagatcgc 12720
ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg cgagatcacc 12780
aaagtagtcg gcaaataaag ctctagtgga tctccgtacc cggggatctg gctcgcggcg 12840
gacgcacgac gccggggcga gaccataggc gatctcctaa atcaatagta gctgtaacct 12900
cgaagcgttt cacttgtaac aacgattgag aatttttgtc ataaaattga aatacttggt 12960
tcgcattttt gtcatccgcg gtcagccgca attctgacga actgcccatt tagctggaga 13020
tgattgtaca tccttcacgt gaaaatttct caagcgctgt gaacaagggt tcagatttta 13080
gattgaaagg tgagccgttg aaacacgttc ttcttgtcga tgacgacgtc gctatgcggc 13140
atcttattat tgaatacctt acgatccacg ccttcaaagt gaccgcggta gccgacagca 13200
cccagttcac aagagtactc tcttccgcga cggtcgatgt cgtggttgtt gatctagatt 13260
taggtcgtga agatgggctc gagctaggag caagtgattt tatcgctaag ccgttcagta 13320
tcagagagtt tctagcacgc attcgggttg ccttgcgcgt gcgccccaac gttgtccgct 13380
ccaaagaccg acggtctttt tgttttactg actggacact taatctcagg caacgtcgct 13440
tgatgtccga agctggcggt gaggtgaaac ttacggcagg tgagttcaat cttctcctcg 13500
cgtttttaga gaaaccccgc gacgttctat cgcgcgagca acttctcatt gccagtcgag 13560
tacgcgacga ggaggtttat gacaggagta tagatgttct cattttgagg ctgcgccgca 13620
aacttgaggc agatccgtca agccctcaac tgataaaaac agcaagaggt gccggttatt 13680
tctttgacgc ggacgtgcag gtttcgcacg gggggacgat ggcagcctga gccaattccc 13740
agatccccga ggaatcggcg tgagcggtcg caaaccatcc ggcccggtac aaatcggcgc 13800
ggcgctgggt gatgacctgg tggagaagtt gaaggccgcg caggccgccc agcggcaacg 13860
catcgaggca gaagcacgcc ccggtgaatc gtggcaagcg gccgctgatc gaatccgcaa 13920
agaatcccgg caaccgccgg cagccggtgc gccgtcgatt aggaagccgc ccaagggcga 13980
cgagcaacca gatttttttcg ttccgatgct ctatgacgtg ggcacccgcg atagtcgcag 14040
catcatggac gtggccgttt tccgtctgtc gaagcgtgac cgacgagctg gcgaggtgat 14100
ccgctacgag cttccagacg ggcacgtaga ggtttccgca gggccggccg gcatggccag 14160
tgtgtgggat tacgacctgg tactgatggc ggtttcccat ctaaccgaat ccatgaaccg 14220
ataccgggaa gggaagggag acaagcccgg ccgcgtgttc cgtccacacg ttgcggacgt 14280
actcaagttc tgccggcgag ccgatggcgg aaagcagaaa gacgacctgg tagaaacctg 14340
cattcggtta aacaccacgc acgttgccat gcagcgtacg aagaaggcca agaacggccg 14400

cctggtgacg gtatccgagg gtgaagcctt gattagccgc tacaagatcg taaagagcga 14460
aaccgggcgg ccggagtaca tcgagatcga gctagctgat tggatgtacc gcgagatcac 14520
agaaggcaag aacccggacg tgctgacggt tcaccccgat tacttttga tcgatcccgg 14580
catcggccgt tttctctacc gcctggcacg ccgcgccgca ggcaaggcag aagccagatg 14640
gttgttcaag acgatctacg aacgcagtgg cagcgccgga gagttcaaga agttctgttt 14700
caccgtgcgc aagctgatcg ggtcaaatga cctgccggag tacgatttga aggaggaggc 14760
ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac ctgatcgagg gcgaagcatc 14820
cgccggttcc taatgtacgg agcagatgct agggcaaatt gccctagcag gggaaaaagg 14880
tcgaaaaggt ctctttcctg tggatagcac gtacattggg aacccaaagc cgtacattgg 14940
gaaccggaac ccgtacattg ggaacccaaa gccgtacatt gggaaccggt cacacatgta 15000
agtgactgat ataaaagaga aaaaaggcga tttttccgcc taaaactctt taaaacttat 15060
taaaactctt aaaacccgcc tggcctgtgc ataactgtct ggccagcgca cagccgaaga 15120
gctgcaaaaa gcgcctaccc ttcggtcgct gcgctcccta cgccccgccg cttcgcgtcg 15180
gcctatcgcg gccgctggcc gctcaaaaat ggctggccta cggccaggca atctaccagg 15240
gcgcggacaa gccgcgccgt cgccactcga ccgccggcgc tgaggtctgc ctcgtgaaga 15300
aggtgttgct gactcatacc aggcctgaat cgccccatca tccagccaga aagtgaggga 15360
gccacggttg atgagagctt tgttgtaggt ggaccagttg gtgattttga acttttgctt 15420
tgccacggaa cggtctgcgt tgtcgggaag atgcgtgatc tgatccttca actcagcaaa 15480
agttcgattt attcaacaaa gccgccgtcc cgtcaagtca gcgtaatgct ctgccagtgt 15540
tacaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg aaactgcaat 15600
ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 15660
gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 15720
actcgtccaa catcaataca acctattaat ttcccctcgt caaaaataag gttatcaagt 15780
gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctc tgcattaatg 15840
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 15900
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 15960
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 16020
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 16080
cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 16140
actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc ctgttccgac 16200
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 16260
tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 16320
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 16380
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 16440

agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 16500
tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 16560
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa 16620
gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg 16680
gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga gattatcaaa 16740
aaggatcttc acctagatcc ttttgatccg gaatta 16776

<210> 38
<211> 17475
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> vecto 24094

<220>
<221> đặc điểm_hỗn tạp
<222> (4)..(259)
<223> bNRB-05
<220>
<221> vùng khởi động
<222> (330)..(2417)
<223> prZmGRMZM2G471240-01
<220>
<221> gen
<222> (2420)..(7288)
<223> cAmCyanCas9-01
<220>
<221> trình tự kết thúc
<222> (7295)..(8290)
<223> tZmGRMZM2G471240-01
<220>
<221> vùng khởi động
<222> (8304)..(8678)
<223> prOsU3-01

<220>
<221> đặc điểm_hỗn tạp
<222> (8679)..(8784)
<223> rsgRNAZmVLHP-02
<220>
<221> đặc điểm_hỗn tạp
<222> (8680)..(8699)
<223> đích ZmVLHP2
<220>
<221> đặc điểm_hỗn tạp
<222> (8700)..(8711)
<223> rCrRNA-01
<220>
<221> đặc điểm_hỗn tạp
<222> (8716)..(8784)
<223> rTracrRNA-01
<220>
<221> vùng khởi động
<222> (8795)..(10786)
<223> prUbi1-04
<220>
<221> gen
<222> (10803)..(11981)
<223> cPMI-09
<220>
<221> trình tự kết thúc
<222> (12004)..(12256)
<223> tNOS-05-01
<220>
<221> đặc điểm_hỗn tạp
<222> (12300)..(12429)
<223> bNLB-03
<220>
<221> gen

<222> (12709)..(13497)

<223> cSpec-03

<220>

<221> vùng khởi động

<222> (13592)..(13722)

<223> prVirG-01

<220>

<221> gen

<222> (13797)..(14429)

<223> cVirG-09

<220>

<221> gen

<222> (14459)..(15532)

<223> cRepA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (15575)..(15979)

<223> oVS1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (16657)..(17463)

<223> oCOLE-06

<400> 38

attcctgtgg ttggcatgca catacaaatg gacgaacgga taaacctttt cacgcccttt 60

taaatatccg attattctaa taaacgctct tttctcttag gtttacccgc caatatatcc 120

tgtcaaacac tgatagttta aactgaaggc gggaaacgac aatctgatca tgagcggaga 180

attaagggag tcacgttatg acccccgccg atgacgcggg acaagccgtt ttacgtttgg 240

aactgacaga accgcaacgc tgcaggaatt ggccgcagcg gccatttaaa caaagcttgg 300

tacctcgcga atgcatctag atgggaccct atttgtactc attccatgtc tcataaactt 360

tgggcaccat ccatccaaca catccaatct aaacacacca aacgatgggg aatggaaaga 420

gcagtattcg attcaacaat ggcaaacaaa tatcactgaa ttagaccaag aataaaccta 480

attagacaac gacctcccaa ccatcattcg tcaggctgta aagaagataa agctgccttg 540

gggcatggat caagcagaac accagagatg aatccaaaca cacagaaaat cacgcgcgct 600

gtctacaatg acaacaagcc ccacatttca ttgcagtaca ctgggctaca aaggcacgta 660

caacaaagag ctagggaaac attgcggagg gcacgagaga gcagctaact tgacaatata 720
gcagactgag cttgcactgt tagcaggcga ggaagggaat catggggacg gagaatgggg 780
tccatgcccg cgaaggagaa ggcggacgcc gccacggtgg caccggcgca cgcgcacaca 840
gggaacccgc acaggcagcc aaggatgctg cctcgccatt gcgccggtcg tctctgccac 900
gctcctctct ctctcccgct gcatcgccgt ggatggggca agcagagagc agggactgcg 960
acgatctggg cggaggactc gccttggaga gcgcggacgc agacgggatt ctagggagag 1020
agcgaagacg gggcgcgcgc ggcgctcgcg cggcgtggtg gcggcgagat tagcgggggt 1080
gggggggaggg cggagccgtg gtgagggtgt ggacgccctc cttaccctct taagtagtag 1140
tagagatata atccgttcca aaatatccat ccgttcaatt tatatttcgt ttgatctttt 1200
taccctaaat ttgattgact catcttatta aaaaagttca taactattat taatctttat 1260
tgagatatca tttagcatat aatatacttt aagtgtggtt ttagattttt tttaaaaaaa 1320
aaaattcgca aaaattaaat gaaacgaccc aatcaaactt gaaaagtaaa actaattata 1380
aatttgaacg gaaggagtaa gaggatgttt gaatgtacta gagctaatag ttggttgctt 1440
taaaatttgc tagtagaatt agctagctaa taaatatcta gataactatt agctaatttg 1500
ctaaaacagc taatagttga actattagct agattgtttg gatgtattcg gctaatttta 1560
atggctaact attagctata gtacaatatt caaacacctc ctaattaaaa tggacaaata 1620
tctcttcttt tggtcccttg cgttagattt ttcatatctc cttatttagt ataaaagaat 1680
catcaaaaag tggacaaccc ctagtggaac accattttag tagtggttgc atgaaacctt 1740
tcgcgcacca gtttctatgt gtcactctaa aaatgggaca gcatgtacgt agtgcctata 1800
tatatacaag tcatctatcg ttgcctcctc agttcatcac taatcacact tattgtgccc 1860
tcgacgagta tctatagcta gctcattaat cgattcgggg gtgtgttgtc gaaggcggca 1920
ttggcgagct actcgtcgcg gcgtccaagc aatacctgta gcacgaaggc gatcgccggg 1980
agcgtggtcg gcgagcccgt cgtgctgggg cagagggtga cggtgctgac ggtggacggc 2040
ggcggcgtcc ggggtctcat cccgggaacc atcctcgcct tcctggaggc caggctgcag 2100
gagctggacg caccggaggc gaggctggcg gactacttcg actacatcgc cggaaccagc 2160
accggcggtc tcatcaccgc cttgctgacc gcgcccggca aggacaagcg gcctctctag 2220
gctgccaagg acatcaacca cttttacatc cataactgcc cgcgcatctt tcctcagaag 2280
tgagtccgat gctgccgcca ttgttcttgc atccatccag catcgtacgt acgtcctcta 2340
tacatctgcg gatcatcatg tgcgcatgtt tgtggcatgc atgcatgcat gtgagcagga 2400
gcaggcttgc gaaaacccca tggccctgtc caacaagttc atcggcgacg acatgaagat 2460
gacctaccac atggacggct gcgtgaacgg ccactacttc accgtgaagg gcgagggcag 2520
cggcaagccc tacgagggca cccagacctc caccttcaag gtgaccatgg ccaacggcgg 2580
ccccctggcc ttctccttcg acatcctgtc caccgtgttc atgtacggca accgctgctt 2640
caccgcctac cccaccagca tgcccgacta cttcaagcag gccttccccg acggcatgtc 2700

ctacgagaga accttcacct acgaggacgg cggcgtggcc accgccagct gggagatcag 2760
cctgaagggc aactgcttcg agcacaagtc caccttccac ggcgtgaact tccccgccga 2820
cggccccgtg atggccaaga agaccaccgg ctgggacccc tccttcgaga agatgaccgt 2880
gtgcgacggc atcttgaagg gcgacgtgac cgccttcctg atgctgcagg gcggcggcaa 2940
ctacagatgc cagttccaca cctcctacaa gaccaagaag cccgtgacca tgccccccaa 3000
ccacgtggtg gagcaccgca tcgccagaac cgacctggac aagggcggca acagcgtgca 3060
gctgaccgag cacgccgtgg cccacatcac ctccgtggtg cccttcggcg gcggcggatc 3120
cgacaagaag tacagcatcg gcctggacat cggcaccaac agcgtgggct gggccgtgat 3180
caccgacgag tacaaggtgc cgagcaagaa gttcaaggtg ctgggcaaca ccgacaggca 3240
cagcatcaag aagaacctga tcggcgccct gctgttcgac agcggcgaga ccgccgaggc 3300
caccaggctg aagaggaccg ccaggaggag gtacaccagg aggaagaaca ggatctgcta 3360
cctgcaggag atcttcagca acgagatggc caaggtggac gacagcttct tccacaggct 3420
ggaggagagc ttcctggtgg aggaggacaa gaagcacgag aggcacccga tcttcggcaa 3480
catcgtggac gaggtggcct accacgagaa gtacccgacc atctaccacc tgaggaagaa 3540
gctggtggac agcaccgaca aggccgacct gaggctgatc tacctggccc tggcccacat 3600
gatcaagttc aggggccact tcctgatcga gggcgacctg aacccggaca acagcgacgt 3660
ggacaagctg ttcatccagc tggtgcagac ctacaaccag ctgttcgagg agaacccgat 3720
caacgccagc ggcgtggacg ccaaggccat cctgagcgcc aggctgagca agagcaggag 3780
gctggagaac ctgatcgccc agctgccggg cgagaagaag aacggcctgt tcggcaacct 3840
gatcgccctg agcctgggcc tgaccccgaa cttcaagagc aacttcgacc tggccgagga 3900
cgccaagctg cagctgagca aggacaccta cgacgacgac ctggacaacc tgctggccca 3960
gatcggcgac cagtacgccg acctgttcct ggccgccaag aacctgagcg acgccatcct 4020
gctgagcgac atcctgaggg tgaacaccga gatcaccaag gccccgctga gcgccagcat 4080
gatcaagagg tacgacgagc accaccagga cctgaccctg ctgaaggccc tggtgaggca 4140
gcagctgccg gagaagtaca aggagatctt cttcgaccag agcaagaacg gctacgccgg 4200
ctacatcgac ggcggcgcca gccaggagga gttctacaag ttcatcaagc cgatcctgga 4260
gaagatggac ggcaccgagg agctgctggt gaagctgaac agggaggacc tgctgaggaa 4320
gcagaggacc ttcgacaacg gcagcatccc gcaccagatc cacctgggcg agctgcacgc 4380
catcctgagg aggcaggagg acttctaccc gttcctgaag gacaacaggg agaagatcga 4440
gaagatcctg accttccgca tcccgtacta cgtgggcccg ctggccaggg gcaacagcag 4500
gttcgcctgg atgaccagga agagcgagga gaccatcacc ccgtggaact tcgaggaggt 4560
ggtggacaag ggcgccagcg cccagagctt catcgagagg atgaccaact tcgacaagaa 4620
cctgccgaac gagaaggtgc tgccgaagca cagcctgctg tacgagtact tcaccgtgta 4680
caacgagctg accaaggtga agtacgtgac cgagggcatg aggaagccgg ccttcctgag 4740

cggcgagcag aagaaggcca tcgtggacct gctgttcaag accaacagga aggtgaccgt 4800
gaagcagctg aaggaggact acttcaagaa gatcgagtgc ttcgacagcg tggagatcag 4860
cggcgtggag gacaggttca acgccagcct gggcacctac cacgacctgc tgaagatcat 4920
caaggacaag gacttcctgg acaacgagga gaacgaggac atcctggagg acatcgtgct 4980
gaccctgacc ctgttcgagg acagggagat gatcgaggag aggctgaaga cctacgccca 5040
cctgttcgac gacaaggtga tgaagcagct gaagaggagg aggtacaccg gctggggcag 5100
gctgagcagg aagctgatca acggcatcag ggacaagcag agcggcaaga ccatcctgga 5160
cttcctgaag agcgacggct tcgccaacag gaacttcatg cagctgatcc acgacgacag 5220
cctgaccttc aaggaggaca tccagaaggc ccaggtgagc ggccagggcg acagcctgca 5280
cgagcacatc gccaacctgg ccggcagccc ggccatcaag aagggcatcc tgcagaccgt 5340
gaaggtggtg gacgagctgg tgaaggtgat gggcaggcac aagccggaga acatcgtgat 5400
cgagatggcc agggagaacc agaccaccca gaagggccag aagaacagca gggagaggat 5460
gaagaggatc gaggagggca tcaaggagct gggcagccag atcctgaagg agcacccggt 5520
ggagaacacc cagctgcaga acgagaagct gtacctgtac tacctgcaga acggcaggga 5580
catgtacgtg gaccaggagc tggacatcaa caggctgagc gactacgacg tggaccacat 5640
cgtgccgcag agcttcctga aggacgacag catcgacaac aaggtgctga ccaggagcga 5700
caagaacagg ggcaagagcg acaacgtgcc gagcgaggag gtggtgaaga agatgaaaaa 5760
ctactggagg cagctgctga acgccaagct gatcacccag aggaagttcg acaacctgac 5820
caaggccgag aggggcggcc tgagcgagct ggacaaggcc ggcttcatta aaaggcagct 5880
ggtggagacc aggcagatca ccaagcacgt ggcccagatc ctggacagca ggatgaacac 5940
caagtacgac gagaacgaca agctgatcag ggaggtgaag gtgatcaccc tgaagagcaa 6000
gctggtgagc gacttcagga aggacttcca gttctacaag gtgagggaga tcaataatta 6060
ccaccacgcc cacgacgcct acctgaacgc cgtggtgggc accgccctga ttaaaaagta 6120
cccgaagctg gagagcgagt tcgtgtacgg cgactacaag gtgtacgacg tgaggaagat 6180
gatcgccaag agcgagcagg agatcggcaa ggccaccgcc aagtacttct tctacagcaa 6240
catcatgaac ttcttcaaga ccgagatcac cctggccaac ggcgagatca ggaagaggcc 6300
gctgatcgag accaacggcg agaccggcga gatcgtgtgg gacaagggca gggacttcgc 6360
caccgtgagg aaggtgctgt ccatgccgca ggtgaacatc gtgaagaaga ccgaggtgca 6420
gaccggcggc ttcagcaagg agagcatcct gccgaagagg aacagcgaca agctgatcgc 6480
caggaagaag gactgggacc cgaagaagta cggcggcttc gacagcccga ccgtggccta 6540
cagcgtgctg gtggtggcca aggtggagaa gggcaagagc aagaagctga agagcgtgaa 6600
ggagctggtg ggcatcacca tcatggagag gagcagcttc gagaagaacc cagtggactt 6660
cctggaggcc aagggctaca aggaggtgaa gaaggacctg atcattaaac tgccgaagta 6720
cagcctgttc gagctggaga acggcaggaa gaggatgctg gccagcgccg gcgagctgca 6780

gaagggcaac gagctggccc tgccgagcaa gtacgtgaac ttcctgtacc tggccagcca 6840
ctacgagaag ctgaagggca gcccggagga caacgagcag aagcagctgt tcgtggagca 6900
gcacaagcac tacctggacg agatcatcga gcagatcagc gagttcagca agagggtgat 6960
cctggccgac gccaacctgg acaaggtgct gagcgcctac aacaagcaca gggacaagcc 7020
gatcagggag caggccgaga acatcatcca cctgttcacc ctgaccaacc tgggcgcccc 7080
ggccgccttc aagtacttcg acaccaccat cgacaggaag aggtacacca gcaccaagga 7140
ggtgctggac gccaccctga tccaccagag catcaccggc ctgtacgaga ccaggatcga 7200
cctgagccag ctgggcggcg acagcagccc gccgaagaag aagaggaagg tgagctggaa 7260
ggacgccagc ggctggagca ggatgtgacc atgggacaag tggctttact gtcagtcaca 7320
tgcttgtaaa taagtagact ttattttaat aaaacataaa aatatatata tgttcttgaa 7380
tataaaattg ataaccaaat taaaattcga accatcactt atacataatt ttactttatt 7440
ttttataaaa cgtgaacggg aaggactacc gtgaatgact atagaaccaa tcatactagt 7500
ataaaatata tgatgacact acgggagaga caaactttgt ctggcgctaa atattttgcc 7560
gagtgtgaat tcacgggcac taggcaaaga tcttctttgc cgagtgttac gctgggcaaa 7620
gtaagacact aggtaaatca gtcatttgcc gagtgtccgc cactaggcaa agcaaaacac 7680
tggcaaatca aaagtttacc tagtgccaga cactaggcaa aaaaaaaacg ctcggcaaat 7740
cggaagtttc cctagtgcca gacactagac aaagaaaaac acttgataaa ctagcgtcgt 7800
cagctaacac catccaccaa ccgttaacgt tgccgagtat ctgacttcga cactcggcaa 7860
agaaggtctc tttgcctagt gtcggtctgg aacactaggc aaagaggcac tttacctagt 7920
gtcgtatttt gacactcagt aaaataattt tttttctttc tgcttccaaa ctttttatga 7980
tgtgttccta tagcacctag aactacatgt caagttttgg taaaattttt gaagtttttg 8040
ctatatttac ttaatttatt ttatttaatt gaatttcttt tgataattca aatttgaact 8100
cggcaaggta agaagcgagg gtagcctgga aacacacttt gcctagtgtt acactcggta 8160
caggagcctc ccctgcctag tgctgcactc gacaaaagat tcgcctttgc ctagcgctgc 8220
actcggcaca ggagtcgcct ttgcctagtg ctgcactagg caaagcctcc gttaccgtgc 8280
cttccatcgt cggacccttc gaagggatct ttaaacatac gaacagatca cttaaagttc 8340
ttctgaagca acttaaagtt atcaggcatg catggatctt ggaggaatca gatgtgcagt 8400
cagggaccat agcacaggac aggcgtcttc tactggtgct accagcaaat gctggaagcc 8460
gggaacactg ggtacgttgg aaaccacgtg atgtggagta agataaactg taggagaaaa 8520
gcatttcgta gtgggccatg aagcctttca ggacatgtat tgcagtatgg gccggcccat 8580
tacgcaattg gacgacaaca aagactagta ttagtaccac ctcggctatc cacatagatc 8640
aaagctggtt taaaagagtt gtgcagatga tccgtggcag ctggagctga gcttccgggg 8700
ttttagagct agaaatagca agttaaaata aggctagtcc gttatcaact tgaaaaagtg 8760
gcaccgagtc ggtgcttttt ttttcggacc gcgcctgcag tgcagcgtga cccggtcgtg 8820

ccccctctcta gagataatga gcattgcatg tctaagttat aaaaaattac cacatatttt 8880
ttttgtcaca cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac 8940
tctacgaata atataatcta tagtactaca ataatatcag tgttttagag aatcatataa 9000
atgaacagtt agacatggtc taaaggacaa ttgagtattt tgacaacagg actctacagt 9060
tttatctttt tagtgtgcat gtgttctcct tttttttgc aaatagcttc acctatataa 9120
tacttcatcc attttattag tacatccatt tagggtttag ggttaatggt ttttatagac 9180
taatttttt agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac 9240
tctattttag tttttttatt taataattta gatataaaat agaataaaat aaagtgacta 9300
aaaattaaac aaataccctt taagaaatta aaaaaactaa ggaaacattt ttcttgtttc 9360
gagtagataa tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga 9420
accagcagcg tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg tcgctgcctc 9480
tggacccctc tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcatccagaa 9540
attgcgtggc ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctcctctcac 9600
ggcaccggca gctacggggg attcctttcc caccgctcct tcgctttccc ttcctcgccc 9660
gccgtaataa atagacaccc cctccacacc ctctttcccc aacctcgtgt tgttcggagc 9720
gcacacacac acaaccagat ctcccccaaa tccacccgtc ggcacctccg cttcaaggta 9780
cgccgctcgt cctcccccc ccccctctct accttctcta gatcggcgtt ccggtccatg 9840
gttagggccc ggtagttcta cttctgttca tgtttgtgtt agatccgtgt ttgtgttaga 9900
tccgtgctgc tagcgttcgt acacggatgc gacctgtacg tcagacacgt tctgattgct 9960
aacttgccag tgtttctctt tggggaatcc tgggatggct ctagccgttc cgcagacggg 10020
atcgatttca tgatttttt tgtttcgttg catagggttt ggtttgccct tttcctttat 10080
ttcaatatat gccgtgcact tgtttgtcgg gtcatctttt catgcttttt tttgtcttgg 10140
ttgtgatgat gtggtctggt tgggcggtcg ttctagatcg gagtagaatt ctgtttcaaa 10200
ctacctggtg gatttattaa ttttggatct gtatgtgtgt gccatacata ttcatagtta 10260
cgaattgaag atgatggatg gaaatatcga tctaggatag gtatacatgt tgatgcgggt 10320
tttactgatg catatacaga gatgcttttt gttcgcttgg ttgtgatgat gtggtgtggt 10380
tgggcggtcg ttcattcgtt ctagatcgga gtagaatact gtttcaaact acctggtgta 10440
tttattaatt ttggaactgt atgtgtgtgt catacatctt catagttacg agtttaagat 10500
ggatggaaat atcgatctag gataggtata catgttgatg tgggttttac tgatgcatat 10560
acatgatggc atatgcagca tctattcata tgctctaacc ttgagtacct atctattata 10620
ataaacaagt atgttttata attattttga tcttgatata cttggatgat ggcatatgca 10680
gcagctatat gtggattttt ttagccctgc cttcatacgc tatttatttg cttggtactg 10740
tttcttttgt cgatgctcac cctgttgttt ggtgttactt ctgcagggat ccggcagcag 10800
ccatgcagaa gctgatcaac agcgtgcaga actacgcctg gggcagcaag accgccctga 10860

ccgagctgta cggcatggag aaccccagca gccagcccat ggccgagctg tggatgggcg 10920
cccaccccaa gagcagcagc cgcgtgcaga acgccgccgg cgacatcgtg agcctgcgcg 10980
acgtgatcga gagcgacaag agcaccctgc tgggcgaggc cgtggccaag cgcttcggcg 11040
agctgccctt cctgttcaag gtgctgtgcg ccgcccagcc cctgagcatc caggtgcacc 11100
ccaacaagca caacagcgag atcggcttcg ccaaggagaa cgccgccggc atccccatgg 11160
acgccgccga gcgcaactac aaggacccca accacaagcc cgagctggtg ttcgccctga 11220
ccccccttcct ggccatgaac gccttccgcg agttcagcga gatcgtgagc ctgctgcagc 11280
ccgtggccgg cgcccacccc gccatcgccc acttcctgca gcagcccgac gccgagcgcc 11340
tgagcgagct gttcgccagc ctgctgaaca tgcagggcga ggagaagagc cgcgccctgg 11400
ccatcctgaa gagcgccctg gacagccagc agggcgagcc ctggcagacc atccgcctga 11460
tcagcgagtt ctaccccgag gacagcggcc tgttcagccc cctgctgctg aacgtggtga 11520
agctgaaccc cggcgaggcc atgttcctgt tcgccgagac cccccacgcc tacctgcagg 11580
gcgtggccct ggaggtgatg gccaacagcg acaacgtgct gcgcgccggc ctgaccccca 11640
agtacatcga catccccgag ctggtggcca acgtgaagtt cgaggccaag cccgccaacc 11700
agctgctgac ccagcccgtg aagcagggcg ccgagctgga cttccccatc cccgtggacg 11760
acttcgcctt cagcctgcac gacctgagcg acaaggagac caccatcagc cagcagagcg 11820
ccgccatcct gttctgcgtg gagggcgacg ccaccctgtg gaagggcagc cagcagctgc 11880
agctgaagcc cggcgagagc gccttcatcg ccgccaacga gagccccgtg accgtgaagg 11940
gccacggccg cctggcccgc gtgtacaaca agctgtgata ggagctcgat ccgtcgacct 12000
gcagatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt tgccggtctt 12060
gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat taacatgtaa 12120
tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt atacatttaa 12180
tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg cgcggtgtca 12240
tctatgttac tagatcggcg cgccgcaatt gaagtttggg cggccagcat ggccgtatcc 12300
gcaatgtgtt attaagttgt ctaagcgtca atttgtttac accacaatat atcctgccac 12360
cagccagcca acagctcccc gaccggcagc tcggcacaaa atcaccactc gatacaggca 12420
gcccatcaga attaattctc atgtttgaca gcttatcatc gactgcacgg tgcaccaatg 12480
cttctggcgt caggcagcca tcggaagctg tggtatggct gtgcaggtcg taaatcactg 12540
cataattcgt gtcgctcaag gcgcactccc gttctggata atgtttttg cgccgacatc 12600
ataacggttc tggcaaatat tctgaaatga gctgttgaca attaatcatc cggctcgtat 12660
aatgtgtgga attgtgagcg gataacaatt tcacacagga aacagaccat gagggaagcg 12720
ttgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga gcgccatctc 12780
gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg cctgaagcca 12840
cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac aacgcggcga 12900

gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga gattctccgc 12960
gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta tccagctaag 13020
cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat cttcgagcca 13080
gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca tagcgttgcc 13140
ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga tctatttgag 13200
gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg cgatgagcga 13260
aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa aatcgcgccg 13320
aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca gcccgtcata 13380
cttgaagcta ggcaggctta tcttggacaa gaagatcgct tggcctcgcg cgcagatcag 13440
ttggaagaat ttgttcacta cgtgaaaggc gagatcacca aagtagtcgg caaataaagc 13500
tctagtggat ctccgtaccc ggggatctgg ctcgcggcgg acgcacgacg ccggggcgag 13560
accataggcg atctcctaaa tcaatagtag ctgtaacctc gaagcgtttc acttgtaaca 13620
acgattgaga atttttgtca taaaattgaa atacttggtt cgcatttttg tcatccgcgg 13680
tcagccgcaa ttctgacgaa ctgcccattt agctggagat gattgtacat ccttcacgtg 13740
aaaatttctc aagcgctgtg aacaagggtt cagattttag attgaaaggt gagccgttga 13800
aacacgttct tcttgtcgat gacgacgtcg ctatgcggca tcttattatt gaatacctta 13860
cgatccacgc cttcaaagtg accgcggtag ccgacagcac ccagttcaca agagtactct 13920
cttccgcgac ggtcgatgtc gtggttgttg atctagattt aggtcgtgaa gatgggctcg 13980
agctaggagc aagtgatttt atcgctaagc cgttcagtat cagagagttt ctagcacgca 14040
ttcgggttgc cttgcgcgtg cgccccaacg ttgtccgctc caaagaccga cggtcttttt 14100
gttttactga ctggacactt aatctcaggc aacgtcgctt gatgtccgaa gctggcggtg 14160
aggtgaaact tacggcaggt gagttcaatc ttctcctcgc gtttttagag aaaccccgcg 14220
acgttctatc gcgcgagcaa cttctcattg ccagtcgagt acgcgacgag gaggtttatg 14280
acaggagtat agatgttctc attttgaggc tgcgccgcaa acttgaggca gatccgtcaa 14340
gccctcaact gataaaaaca gcaagaggtg ccggttattt ctttgacgcg gacgtgcagg 14400
tttcgcacgg ggggacgatg gcagcctgag ccaattccca gatccccgag gaatcggcgt 14460
gagcggtcgc aaaccatccg gcccggtaca aatcggcgcg gcgctgggtg atgacctggt 14520
ggagaagttg aaggccgcgc aggccgccca gcggcaacgc atcgaggcag aagcacgccc 14580
cggtgaatcg tggcaagcgg ccgctgatcg aatccgcaaa gaatcccggc aaccgccggc 14640
agccggtgcg ccgtcgatta ggaagccgcc caagggcgac gagcaaccag atttttcgt 14700
tccgatgctc tatgacgtgg gcacccgcga tagtcgcagc atcatggacg tggccgtttt 14760
ccgtctgtcg aagcgtgacc gacgagctgg cgaggtgatc cgctacgagc ttccagacgg 14820
gcacgtagag gtttccgcag ggccggccgg catggccagt gtgtgggatt acgacctggt 14880
actgatggcg gtttcccatc taaccgaatc catgaaccga taccgggaag ggaagggaga 14940

caagcccggc cgcgtgttcc gtccacacgt tgcggacgta ctcaagttct gccggcgagc 15000
cgatggcgga aagcagaaag acgacctggt agaaacctgc attcggttaa acaccacgca 15060
cgttgccatg cagcgtacga agaaggccaa gaacggccgc ctggtgacgg tatccgaggg 15120
tgaagccttg attagccgct acaagatcgt aaagagcgaa accgggcggc cggagtacat 15180
cgagatcgag ctagctgatt ggatgtaccg cgagatcaca gaaggcaaga acccggacgt 15240
gctgacggtt caccccgatt acttttttgat cgatcccggc atcggccgtt ttctctaccg 15300
cctggcacgc cgcgccgcag gcaaggcaga agccagatgg ttgttcaaga cgatctacga 15360
acgcagtggc agcgccggag agttcaagaa gttctgtttc accgtgcgca agctgatcgg 15420
gtcaaatgac ctgccggagt acgatttgaa ggaggaggcg gggcaggctg gcccgatcct 15480
agtcatgcgc taccgcaacc tgatcgaggg cgaagcatcc gccggttcct aatgtacgga 15540
gcagatgcta gggcaaattg ccctagcagg ggaaaaaggt cgaaaaggtc tctttcctgt 15600
ggatagcacg tacattggga acccaaagcc gtacattggg aaccggaacc cgtacattgg 15660
gaacccaaag ccgtacattg ggaaccggtc acacatgtaa gtgactgata taaaagagaa 15720
aaaaggcgat ttttccgcct aaaactcttt aaaacttatt aaaactctta aaacccgcct 15780
ggcctgtgca taactgtctg gccagcgcac agccgaagag ctgcaaaaag cgcctaccct 15840
tcggtcgctg cgctccctac gccccgccgc ttcgcgtcgg cctatcgcgg ccgctggccg 15900
ctcaaaaatg gctggcctac ggccaggcaa tctaccaggg cgcggacaag ccgcgccgtc 15960
gccactcgac cgccggcgct gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca 16020
ggcctgaatc gccccatcat ccagccagaa agtgagggag ccacggttga tgagagcttt 16080
gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac ggtctgcgtt 16140
gtcgggaaga tgcgtgatct gatccttcaa ctcagcaaaa gttcgattta ttcaacaaag 16200
ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt aaccaattct 16260
gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc aggattatca 16320
ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc gaggcagttc 16380
cataggatgg caagatcctg gtatcggtct gcgattccga ctcgtccaac atcaatacaa 16440
cctattaatt tcccctcgtc aaaaataagg ttatcaagtg agaaatcacc atgagtgacg 16500
actgaatccg gtgagaatgg caaaagctct gcattaatga atcggccaac gcgcggggag 16560
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt 16620
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga 16680
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 16740
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa 16800
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 16860
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 16920
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct 16980

cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 17040
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 17100
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 17160
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat 17220
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 17280
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa 17340
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 17400
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct 17460
tttgatccgg aatta 17475

<210> 39
<211> 19
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 39
ttgtgctgct ccacgaaca 19

<210> 40
<211> 20
<212> ADN
<213> trình tự nhân tạo
<220>
<223> đoạn mồi
<400> 40
gccagccact acgagaagct 20

<210> 41
<211> 24
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn dò

<400> 41

ctgcttctgc tcgttgtcct ccgg 24

<210> 42

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 42

gcggatgctg gcacagc 17

<210> 43

<211> 19

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 43

ggcattgctt ccttctccg 19

<210> 44

<211> 15

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 44

cagggagcga ggtac 15

<210> 45

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 45

ctggtggcca acgtgaagtt 20

<210> 46

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 46

gcttcacggg ctgggtc 17

<210> 47

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 47

aggccaagcc cgccaaccag 20

<210> 48

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 48

gcggatgctg gcacaga 17

<210> 49

<211> 18

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 49

gcattgcttc cttcgcca 18

<210> 50

<211> 16

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 50

cagggaggta cgaacc 16

<210> 51

<211> 16

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 51

gcggcgaaga agcgaa 16

<210> 52

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 52

gcggcgtctc cagcttc 17

<210> 53
<211> 12
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn dò
<400> 53
ccaggaactg cg 12

<210> 54
<211> 18
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 54
aagaaacgcc ggctgagt 18

<210> 55
<211> 15
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 55
accttgcggg gcgtt 15

<210> 56
<211> 12
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn dò
<400> 56

ccaggaactg cg 12

<210> 57
<211> 18
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 57
aagaaacgcc ggctgagt 18

<210> 58
<211> 14
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 58
ccttgcgcgg cgtc 14

<210> 59
<211> 12
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn dò
<400> 59
ccaggaactg cg 12

<210> 60
<211> 19
<212> ADN
<213> Trình tự Nhân tạo
<220>

<223> đoạn mồi

<400> 60

tgatcctcga ggccaagct 19

<210> 61

<211> 18

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 61

aggtcgaggt cccctcca 18

<210> 62

<211> 14

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 62

cctgctaccc gggc 14

<210> 63

<211> 15

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 63

cgcgccctgc taccc 15

<210> 64

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 64

gcgcgtgctt accagga 17

<210> 65

<211> 13

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 65

tcgaggagtg ccc 13

<210> 66

<211> 16

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 66

caccgatgag caggcg 16

<210> 67

<211> 18

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 67

agatacacct tccggccg 18

<210> 68

<211> 14

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 68

ttcctcccgg aagc 14

<210> 69

<211> 16

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 69

caccgatgag caggcg 16

<210> 70

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 70

agatacacct tccggccagt 20

<210> 71

<211> 14

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 71

ctcctcccgg aagc 14

<210> 72
<211> 25
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 72
caagtttctg gacaaggaga ttctc 25

<210> 73
<211> 27
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 73
aagaattccc ttcttaatag ctggaga 27

<210> 74
<211> 26
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn dò
<400> 74
cacgagcaca ttgctaacct tgctgg 26

<210> 75
<211> 17
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi

<400> 75

tcaccgatga gcaggca 17

<210> 76

<211> 18

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 76

atacaccttc cggccagc 18

<210> 77

<211> 14

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 77

ttcctcccgg aagc 14

<210> 78

<211> 24

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 78

gatagggcta aagagatgtg ggaa 24

<210> 79

<211> 26

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 79

ctttgttcac attagggctc aaataa 26

<210> 80

<211> 16

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 80

tagactgaga tggatg 16

<210> 81

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 81

aaaaccaccg gagaagacga 20

<210> 82

<211> 17

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 82

aggtgtggcg gcagtga 17

<210> 83

<211> 14

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 83

caccgtcatt gttc 14

<210> 84

<211> 25

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 84

caagtttctg gacaaggaga ttctc 25

<210> 85

<211> 27

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 85

aagaattccc ttcttaatag ctggaga 27

<210> 86

<211> 26

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 86

cacgagcaca ttgctaacct tgctgg 26

<210> 87

<211> 15

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 87

gcgacgccgg aaagg 15

<210> 88

<211> 21

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 88

tggcgtggtt tcgtcttctt a 21

<210> 89

<211> 25

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn dò

<400> 89

aagagcggcg tctggaggtg actca 25

<210> 90

<211> 20

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 90

aaccgcatcg tcagaaaaac 20

<210> 91
<211> 20
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 91
tcaacttaac cggccaaatc 20

<210> 92
<211> 21
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 92
catcccttct cttcccteet g 21

<210> 93
<211> 23
<212> ADN
<213> Trình tự Nhân tạo
<220>
<223> đoạn mồi
<400> 93
gccagtgtga gtgtgtatga gca 23

<210> 94
<211> 21
<212> ADN
<213> Trình tự Nhân tạo
<220>

<223> đoạn mồi

<400> 94

catcgttttc tcccctcctc a 21

<210> 95

<211> 19

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 95

actgatatgc acggcgcca 19

<210> 96

<211> 22

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 96

tgcagtagct tcattttcac cg 22

<210> 97

<211> 22

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> đoạn mồi

<400> 97

aggaattgat atgtacgccc gt 22

<210> 98

<211> 16279

<212> ADN

<213> Trình tự Nhân tạo

<220>

<223> vecto 24075

<220>

<221> đặc điểm_hỗn tạp

<222> (1)..(517)

<223> bNRB-07

<220>

<221> vùng khởi động

<222> (538)..(1697)

<223> prAtEFaA1-02

<220>

<221> gen

<222> (1716)..(5885)

<223> cCas9-05

<220>

<221> đột biến

<222> (5205)..(5207)

<223> đột biến L thành V

<220>

<221> đột biến

<222> (5250)..(5252)

<223> đột biến I thành V

<220>

<221> trình tự kết thúc

<222> (5894)..(6146)

<223> tNOS-05-01

<220>

<221> vùng khởi động

<222> (6173)..(6620)

<223> prAtU6-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6621)..(6640)

<223> AtGL1 target1

<220>

<221> đặc điểm_hỗn tạp

<222> (6621)..(6725)

<223> rsgRNA AtGL1-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6641)..(6652)

<223> rCrRNA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (6657)..(6725)

<223> rTracrRNA-01

<220>

<221> vùng khởi động

<222> (6726)..(7173)

<223> prAtU6-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7174)..(7193)

<223> AtGl1 target 2

<220>

<221> đặc điểm_hỗn tạp

<222> (7174)..(7278)

<223> rsgRNA AtGL1-02

<220>

<221> đặc điểm_hỗn tạp

<222> (7194)..(7205)

<223> rCrRNA-01

<220>

<221> đặc điểm_hỗn tạp

<222> (7210)..(7278)

<223> rTracrRNA-01

<220>
<221> vùng khởi động
<222> (7295)..(7640)
<223> prCMP-02
<220>
<221> gen
<222> (7653)..(8447)
<223> cNpt2-10
<220>
<221> trình tự kết thúc
<222> (8476)..(8728)
<223> tNOS-05-01
<220>
<221> vùng khởi động
<222> (8755)..(10752)
<223> prGmUBI-01
<220>
<221> gen
<222> (10765)..(11454)
<223> cAmCyan-06
<220>
<221> trình tự kết thúc
<222> (11477)..(12119)
<223> tPsE9-01
<220>
<221> đặc điểm_hỗn tạp
<222> (12193)..(12311)
<223> bNLB-03
<220>
<221> gen
<222> (12928)..(13716)
<223> cSpec-03
<220>
<221> gen

<222> (13751)..(14824)

<223> cRepA-08

<220>

<221> đặc điểm_hỗn tạp

<222> (14867)..(15271)

<223> oVC1-04

<220>

<221> đặc điểm_hỗn tạp

<222> (15441)..(16247)

<223> oCOLE-06

<400> 98

gccaggacgt cggccgaaag agcgacaagc agatcacgct tttcgacagc gtcggatttg 60

cgatcgagga tttttcggcg ctgcgctacg tccgcgaccg cgttgaggga tcaagccaca 120

gcagcccact cgaccttcta gccgacccag acgagccaag ggatcttttt ggaatgctgc 180

tccgtcgtca ggctttccga cgtttgggtg gttgaacaga agtcattatc gcacggaatg 240

ccaagcactc ccgaggggaa ccctgtggtt ggcatgcaca tacaaatgga cgaacggata 300

aacctttca cgccctttta aatatccgat tattctaata aacgctcttt tctcttaggt 360

ttacccgcca atatatcctg tcaaacactg atagtttaaa ctgaaggcgg gaaacgacaa 420

tctgatcatg agcggagaat taagggagtc acgttatgac ccccgccgat gacgcgggac 480

aagccgtttt acgtttggaa ctgacagaac cgcaacgttg aaggcgcgcc actcagcaag 540

cttgatatcg gaagtttctc tcttgaggga ggttgctcgt ggaatgggac acatatggtt 600

gttataataa accatttcca ttgtcatgag attttgaggt taatatatac tttacttgtt 660

cattatttta tttggtgttt gaataaatga tataaatggc tcttgataat ctgcattcat 720

tgagatatca aatatttact ctagagaaga gtgtcatata gattgatggt ccacaatcaa 780

tgaaattttt gggagacgaa catgtataac catttgcttg aataacctta attaaaaggt 840

gtgattaaat gatgtttgta acatgtagta ctaaacattc ataaaacaca accaacccaa 900

gaggtattga gtattcacgg ctaaacaggg gcataatggt aatttaaaga atgatattat 960

tttatgttaa accctaacat tggtttcgga ttcaacgcta taaataaaac cactctcgtt 1020

gctgattcca tttatcgttc ttattgaccc tagccgctac acacttttct gcgatatctc 1080

tgaggtaagc gttaacgtac ccttagatcg ttctttttct ttttcgtctg ctgatcgttg 1140

ctcatattat ttcgatgatt gttggattcg atgctctttg ttgattgatc gttctgaaaa 1200

ttctgatctg ttgtttagat tttatcgatt gttaatatca acgtttcact gcttctaaac 1260

gataatttat tcatgaaact attttcccat tctgatcgat cttgttttga gattttaatt 1320

tgttcgattg attgttggtt ggtggatcta tatacgagtg aacttgttga tttgcgtatt 1380

taagatgtat gtcgatttga attgtgattg ggtaattctg gagtagcata acaaatccag 1440

tgttcccttt ttctaagggt aattctcgga ttgtttgctt tatatctctt gaaattgccg 1500

atttgattga atttagctcg cttagctcag atgatagagc accacaattt ttgtggtaga 1560

aatcggtttg actccgatag cggcttttta ctatgattgt tttgtgttaa agatgatttt 1620

cataatggtt atatatgtct actgtttttа ttgattcaat atttgattgt tctttttttt 1680

gcagatttgt tgaccaggga tccgcggccg ctaaaatgga taagaagtat tctattggac 1740

ttgatattgg aaccaactct gtgggatggg ctgttattac tgacgagtat aaggttccat 1800

ctaagaagtt caaggttctt ggaaacactg atagacactc tattaagaag aaccttattg 1860

gtgctcttct tttcgattct ggagagactg ctgaggctac tagacttaag agaactgcta 1920

gaagaagata tactagaaga aagaacagaa tttgctatct tcaagagatt ttctctaacg 1980

agatggctaa ggttgacgat tctttcttcc acagacttga ggagtctttc cttgttgagg 2040

aggataagaa gcacgagaga cacccaattt tcggaaacat tgttgacgag gttgcttatc 2100

acgagaagta tccaactatt tatcaccttа gaaagaagct cgttgattct actgataagg 2160

ctgatcttag acttatttat cttgctcttg ctcacatgat taagttcaga ggacacttcc 2220

ttattgaggg agatcttaac ccagataact ctgacgttga taagctcttc attcaacttg 2280

ttcaaactta taaccaactt ttcgaggaga acccaattaa cgcttctgga gttgacgcta 2340

aggctattct ttctgctaga ctttctaagt ctagaaggct tgagaacctt attgctcaac 2400

ttccaggaga gaagaagaac ggacttttcg gaaaccttat tgctctttct cttggactta 2460

ctccaaactt caagtctaac ttcgatcttg ctgaggacgc taagctccaa ctttctaagg 2520

atacttacga cgatgatctt gataaccttc ttgctcaaat tggagatcaa tacgctgatc 2580

ttttccttgc tgctaagaac ctttctgacg ctattcttct ttctgatatt cttagagtta 2640

acactgagat tactaaggct ccactttctg cttctatgat taagagatac gacgagcacc 2700

accaagatct tactcttctt aaggctcttg ttagacaaca acttccagag aagtataagg 2760

agattttctt cgatcaatct aagaacggat acgctggata tattgacgga ggagcttctc 2820

aagaggagtt ctataagttc attaagccaa ttcttgagaa gatggacgga actgaggagc 2880

ttcttgttaa gctcaacaga gaggatcttc ttagaaagca aagaactttc gataacggat 2940

ctattccaca ccaaattcac cttggagagc ttcacgctat tcttagaagg caagaggatt 3000

tctatccatt ccttaaggat aacagagaga agattgagaa gattcttact ttccgtattc 3060

catattacgt tggaccactt gctagaggaa actctagatt cgcttggatg actagaaagt 3120

ctgaggagac tattactcct tggaacttcg aggaggttgt tgataaggga gcttctgctc 3180

aatctttcat tgagagaatg actaacttcg ataagaacct tccaaacgag aaggttcttc 3240

caaagcactc tcttctttac gagtatttca ctgtttataa cgagcttact aaggttaagt 3300

acgttactga gggaatgaga aagccagctt tcctttctgg agagcaaaag aaggctattg 3360

ttgatcttct tttcaagact aacagaaagg ttactgttaa gcaacttaag gaggattatt 3420

tcaagaagat tgagtgcttc gattctgttg agatttctgg agttgaggat agattcaacg 3480
cttctcttgg aacttatcac gatcttctta agattattaa ggataaggat ttccttgata 3540
acgaggagaa cgaggatatt cttgaggata ttgttcttac tcttactctt ttcgaggata 3600
gagagatgat tgaggagaga cttaagactt acgctcacct tttcgacgat aaggttatga 3660
agcaacttaa gagaagaaga tatactggat ggggtagact ttctagaaag ctcattaacg 3720
gaattagaga taagcaatct ggaaagacta ttcttgattt ccttaagtct gacggattcg 3780
ctaacagaaa cttcatgcaa cttattcacg acgattctct tactttcaag gaggatattc 3840
aaaaggctca agtttctgga caaggagatt ctcttcacga gcacattgct aaccttgctg 3900
gatctccagc tattaagaag ggaattcttc aaactgttaa ggttgttgac gagcttgtta 3960
aggttatggg tagacacaag ccagagaaca ttgttattga gatggctaga gagaaccaaa 4020
ctactcaaaa gggacaaaag aactctagag agagaatgaa gagaattgag gagggaatta 4080
aggagcttgg atctcaaatt cttaaggagc acccagttga gaacactcaa cttcaaaacg 4140
agaagctcta tctttattat cttcaaaacg gaagagatat gtacgttgat caagagcttg 4200
atattaacag actttctgat tacgacgttg atcacattgt tccacaatct ttccttaagg 4260
acgattctat tgataacaag gttcttacta gatctgataa gaacagagga aagtctgata 4320
acgttccatc tgaggaggtt gttaagaaga tgaagaacta ttggagacaa cttcttaacg 4380
ctaagctcat tactcaaaga aagttcgata accttactaa ggctgagaga ggaggacttt 4440
ctgagcttga taaggctgga ttcattaaga gacaacttgt tgagactaga caaattacta 4500
agcacgttgc tcaaattctt gattctagaa tgaacactaa gtacgacgag aacgataagc 4560
tcattagaga ggttaaggtt attactctta agtctaagct cgtttctgat ttcagaaagg 4620
atttccaatt ctataaggtt agagagatta acaactatca ccacgctcac gacgcttatc 4680
ttaacgctgt tgttggaact gctcttatta agaagtatcc aaaacttgag tctgagttcg 4740
tttacggaga ttataaggtt tacgacgtta gaaagatgat tgctaagtct gagcaagaga 4800
ttggaaaggc tactgctaag tatttcttct attctaacat tatgaacttc ttcaagactg 4860
agattactct tgctaacgga gagattagaa agaggccact tattgagact aacggagaga 4920
ctggagagat tgtttgggat aagggaagag atttcgctac tgttagaaag gttctttcta 4980
tgccacaagt taacattgtt aagaaaactg aggttcaaac tggaggattc tctaaggagt 5040
ctattcttcc aaagagaaac tctgataagc tcattgctag aaagaaggat tgggacccaa 5100
agaagtacgg aggattcgat tctccaactg ttgcttattc tgttcttgtt gttgctaagg 5160
ttgagaaggg aaagtctaag aagctcaagt ctgttaagga gcttgttgga attactatta 5220
tggagagatc ttctttcgag aagaacccag ttgatttcct tgaggctaag ggatataagg 5280
aggttaagaa ggatcttatt attaagctcc caaagtattc tcttttcgag cttgagaacg 5340
gaagaaagag aatgcttgct tctgctggag agcttcaaaa gggaaacgag cttgctcttc 5400
catctaagta cgttaacttc ctttatcttg cttctcacta cgagaagctc aagggatctc 5460

cagaggataa cgagcaaaag caacttttcg ttgagcaaca caagcactat cttgacgaga 5520
ttattgagca aatttctgag ttctctaaga gagttattct tgctgacgct aaccttgata 5580
aggttctttc tgcttataac aagcacagag ataagccaat tagagagcaa gctgagaaca 5640
ttattcacct tttcactctt actaaccttg gtgctccagc tgctttcaag tatttcgata 5700
ctactattga tagaaagaga tatacttcta ctaaggaggt tcttgacgct actcttattc 5760
accaatctat tactggactt tacgagacta gaattgatct ttctcaactt ggaggagatt 5820
cttctccacc aaagaagaag agaaaggttt cttggaagga cgcttctgga tggtctagaa 5880
tgtgacgtcg cgtgatcgtt caaacatttg gcaataaagt ttcttaagat tgaatcctgt 5940
tgccggtctt gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat 6000
taacatgtaa tgcatgacgt tatttatgag atgggttttt atgattagag tcccgcaatt 6060
atacatttaa tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg 6120
cgcggtgtca tctatgttac tagatctgca gatcggaccc ctaattagct aaaagcttcg 6180
ttgaacaacg gaaactcgac ttgccttccg cacaatacat catttcttct tagctttttt 6240
tcttcttctt cgttcataca gtttttttttt gtttatcagc ttacattttc ttgaaccgta 6300
gctttcgttt tcttcttttt aactttccat tcggagtttt tgtatcttgt ttcatagttt 6360
gtcccaggat tagaatgatt aggcatcgaa ccttcaagaa tttgattgaa taaaacatct 6420
tcattcttaa gatatgaaga taatcttcaa aaggcccctg ggaatctgaa agaagagaag 6480
caggcccatt tatatgggaa agaacaatag tatttcttat ataggcccat ttaagttgaa 6540
aacaatcttc aaaagtccca catcgcttag ataagaaaac gaagctgagt ttatatacag 6600
ctagagtcga agtagtgatt ggaaaagttg tagactgaga gttttagagc tagaaatagc 6660
aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt 6720
tttttaagct tcgttgaaca acggaaactc gacttgcctt ccgcacaata catcatttct 6780
tcttagcttt ttttcttctt cttcgttcat acagtttttt tttgtttatc agcttacatt 6840
ttcttgaacc gtagctttcg ttttcttctt tttaactttc cattcggagt ttttgtatct 6900
tgtttcatag tttgtcccag gattagaatg attaggcatc gaaccttcaa gaatttgatt 6960
gaataaaaca tcttcattct taagatatga agataatctt caaaaggccc ctgggaatct 7020
gaaagaagag aagcaggccc atttatatgg gaaagaacaa tagtatttct tatataggcc 7080
catttaagtt gaaaacaatc ttcaaaagtc ccacatcgct tagataagaa aacgaagctg 7140
agtttatata cagctagagt cgaagtagtg attgcagtga tgaacaatga cgggttttag 7200
agctagaaat agcaagttaa aataaggcta gtccgttatc aacttgaaaa agtggcaccg 7260
agtcggtgct ttttttttgg cgcgcctaaa gcttctggca gacaaagtgg cagacatact 7320
gtcccacaaa tgaagatgga atctgtaaaa gaaaacgcgt gaaataatgc gtctgacaaa 7380
ggttaggtcg gctgccttta atcaatacca aagtggtccc taccacgatg gaaaaactgt 7440
gcagtcggtt tggctttttc tgacgaacaa ataagattcg tggccgacag gtgggggtcc 7500

accatgtgaa ggcatcttca gactccaata atggagcaat gacgtaaggg cttacgaaat 7560
aagtaagggt agtttgggaa atgtccactc acccgtcagt ctataaatac ttagcccctc 7620
cctcattgtt aagggagcaa ggatcctaaa ccatgattga acaagatgga ttgcacgcag 7680
gttctccggc cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg 7740
gctgctctga tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca 7800
agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc 7860
tggccacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg 7920
actggctgct attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg 7980
ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta 8040
cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag 8100
ccggtcttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac 8160
tgttcgccag gctcaaggcg cgcatgcccg acggcgagga tctcgtcgtg acccatggcg 8220
atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg 8280
gccggctggg tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg 8340
aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg 8400
attcgcagcg catcgccttc tatcgccttc ttgacgagtt cttctgatga gagctctaga 8460
tccccgaatt tccccgatcg ttcaaacatt tggcaataaa gtttcttaag attgaatcct 8520
gttgccggtc ttgcgatgat tatcatataa tttctgttga attacgttaa gcatgtaata 8580
attaacatgt aatgcatgac gttatttatg agatgggttt ttatgattag agtcccgcaa 8640
ttatacattt aatacgcgat agaaaacaaa atatagcgcg caaactagga taaattatcg 8700
cgcgcggtgt catctatgtt actagatcgg gaattgggta ccctaattag ctaaattcca 8760
aaatttcag ttagtcctta ctaattatta aattatagta ttaatccaat gtgattgcgg 8820
ttacatcatg tacggaaaaa taattctaat ccttgattta aatttgatct tgactattta 8880
tttattcttt atttcatttt gtaaatcatt ttatgtatct cctggcaagc aattttatcc 8940
accttgcacc aacaccttcg ggttccataa tcaaaccacc ttaacttcac accatgctgt 9000
aactcacacc gcccagcatc tccaatgtga aagaagctaa aatttaataa acaatcatac 9060
gaagcagtga caaaatacca gatggtatta atgctttgat aaaattaatt ggaaagtata 9120
aaatggtaga aaataataaa ttataattaa tttaaataag ataaaaaata attaaaaact 9180
aaaatgttaa aattttaaaa aaattatttt aaataatatt taaaaacatt aaaaatcatt 9240
ttaaaaaatt tatttataga acaattaaat aaatatttca gctaataaaa aacaaaagct 9300
tacctagcct tagaagacaa cttgtccaac aattagatga tacccattgc ccttacgttt 9360
tctttaacat caattattgt ttttgtcaac aagctatctt ttagttttat tttattggta 9420
aaaaatatgt cgccttcaag ttgcatcatt taacacatct cgtcattaga aaaataaaac 9480
tcttccctaa acgattagta gaaaaaatca ttcgataata aataagaaag aaaaattaga 9540

aaaaaataac ttcattttaa aaaaatcatt aaggctatat tttttaaatg actaatttta 9600
tatagactgt aactaaaagt atacaattta ttatgctatg tatcttaaag aattacttat 9660
aaaaatctac ggaagaatat cttacaaagt gaaaaacaaa tgagaaagaa tttagtggga 9720
tgattatgat tttatttgaa aattgaaaaa ataattatta aagactttag tggagtaaga 9780
aagctttcct attagtcttt tcttatccat aaaaaaaaaa aaaaatctag cgtgacagct 9840
tttccataga ttttaataat gtaaaatact ggtagcagcc gaccgttcag gtaatggaca 9900
ctgtggtcct aacttgcaac gggtgcgggc ccaatttaat aacgccgtgg taacggataa 9960
agccaagcgt gaagcggtga aggtacatct ctgactccgt caagattacg aaaccgtcaa 10020
ctacgaagga ctccccgaaa tatcatctgt gtcataaaca ccaagtcaca ccatacatgg 10080
gcacgcgtca caatatgatt ggagaacggt tccaccgcat atgctataaa atgcccccac 10140
acccctcgac cctaatcgca cttcaattgc aatcaaatta gttcattctc tttgcgcagt 10200
tccctacctc tcctttcaag gttcgtagat ttcttctgtt tttttttctt cttctttatt 10260
gtttgttcta catcagcatg atgttgattt gattgtgttt tctatcgttt catcgattat 10320
aaattttcat aatcagaaga ttcagctttt attaatgcaa gaacgtcctt aattgatgat 10380
tttataaccg taaattaggt ctaattagag tttttttcat aaagattttc agatccgttt 10440
acaacaagcc ttaattgttg attctgtagt cgtagattaa ggtttttttc atgaactact 10500
tcagatccgt taaacaacag ccttatttgt tgatacttca gtcgtttttc aagaaattgt 10560
tcagatccgt tgataaaagc cttattcgtt gattctgtat ggtatttcaa gagatattgc 10620
tcaggtcctt tagcaactac cttatttgtt gattctgtgg ccatagatta ggatttttt 10680
tcacgaaatt gcttcttgaa attacgtgat ggattttgat tctgatttat cttgtgattg 10740
ttgactctac agagatctaa aaaaatggcc ctgtccaaca agttcatcgg cgacgacatg 10800
aagatgacct accacatgga cggctgcgtg aacggccact acttcaccgt gaagggcgag 10860
ggcagcggca agccctacga gggcacccag acctccacct tcaaggtgac gatggccaac 10920
ggcggccccc tggccttctc cttcgacatc ctgtccaccg tgttcatgta cggcaaccgc 10980
tgcttcaccg cctaccccac cagcatgccc gactacttca agcaggcctt ccccgacggc 11040
atgtcctacg agagaacctt cacctacgag gacggcggcg tggccaccgc cagctgggag 11100
atcagcctga agggcaactg cttcgagcac aagtccacct tccacggcgt gaacttcccc 11160
gccgacggcc ccgtgatggc caagaagacc accggctggg atccctcctt cgagaagatg 11220
accgtgtgcg acggcatctt gaagggcgac gtgaccgcct tcctgatgct gcagggcggc 11280
ggcaactaca gatgccagtt ccacacctcc tacaagacca agaagcccgt gaccatgccc 11340
cccaaccacg tggtggagca ccgcatcgcc agaaccgacc tggacaaggg cggcaacagc 11400
gtgcagctga ccgagcacgc cgtggcccac atcacctccg tggtgccctt ctgatgaact 11460
agtgaattcg agctcaagct ttcgttcgta tcatcggttt cgacaacgtt cgtcaagttc 11520
aatgcatcag tttcattgcg cacacaccag aatcctactg agtttgagta ttatggcatt 11580

gggaaaactg tttttcttgt accatttgtt gtgcttgtaa tttactgtgt tttttattcg 11640
gttttcgcta tcgaactgtg aaatggaaat ggatggagaa gagttaatga atgatatggt 11700
ccttttgttc attctcaaat taatattatt tgtttttttct cttatttgtt gtgtgttgaa 11760
tttgaaatta taagagatat gcaaacattt tgttttgagt aaaaatgtgt caaatcgtgg 11820
cctctaatga ccgaagttaa tatgaggagt aaaacacttg tagttgtacc attatgctta 11880
ttcactaggc aacaaatata ttttcagacc tagaaaagct gcaaatgtta ctgaatacaa 11940
gtatgtcctc ttgtgtttta gacatttatg aactttcctt tatgtaattt tccagaatcc 12000
ttgtcagatt ctaatcattg ctttataatt atagttatac tcatggattt gtagttgagt 12060
atgaaaatat tttttaatgc attttatgac ttgccaattg attgacaaca tgcatcaatc 12120
ggaccgttaa ctagctagac ggccaggatc gccgcgtgag cctttagcaa ctagctagat 12180
taattaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa 12240
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca 12300
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg 12360
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg 12420
aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 12480
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 12540
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 12600
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgtcgactca 12660
tgtttgacag cttatcatcg actgcacggt gcaccaatgc ttctggcgtc aggcagccat 12720
cggaagctgt ggtatggctg tgcaggtcgt aaatcactgc ataattcgtg tcgctcaagg 12780
cgcactcccg ttctggataa tgttttttgc gccgacatca taacggttct ggcaaatatt 12840
ctgaaatgag ctgttgacaa ttaatcatcc ggctcgtata atgtgtggaa ttgtgagcgg 12900
ataacaattt cacacaggaa acagaccatg agggaagcgt tgatcgccga agtatcgact 12960
caactatcag aggtagttgg cgtcatcgag cgccatctcg aaccgacgtt gctggccgta 13020
catttgtacg gctccgcagt ggatggcggc ctgaagccac acagtgatat tgatttgctg 13080
gttacggtga ccgtaaggct tgatgaaaca acgcggcgag ctttgatcaa cgacctttg 13140
gaaacttcgg cttcccctgg agagagcgag attctccgcg ctgtagaagt caccattgtt 13200
gtgcacgacg acatcattcc gtggcgttat ccagctaagc gcgaactgca atttggagaa 13260
tggcagcgca atgacattct tgcaggtatc ttcgagccag ccacgatcga cattgatctg 13320
gctatcttgc tgacaaaagc aagagaacat agcgttgcct tggtaggtcc agcggcggag 13380
gaactctttg atccggttcc tgaacaggat ctatttgagg cgctaaatga aaccttaacg 13440
ctatggaact cgccgcccga ctgggctggc gatgagcgaa atgtagtgct tacgttgtcc 13500
cgcatttggt acagcgcagt aaccggcaaa atcgcgccga aggatgtcgc tgccgactgg 13560
gcaatggagc gcctgccggc ccagtatcag cccgtcatac ttgaagctag gcaggcttat 13620

cttggacaag aagatcgctt ggcctcgcgc gcagatcagt tggaagaatt tgttcactac 13680
gtgaaaggcg agatcaccaa agtagtcggc aaataaagct ctagtggatc tccgtacccg 13740
aggaatcggc gtgagcggtc gcaaaccatc cggcccggta caaatcggcg cggcgctggg 13800
tgatgacctg gtggagaagt tgaaggccgc gcaggccgcc cagcggcaac gcatcgaggc 13860
agaagcacgc cccggtgaat cgtggcaagc ggccgctgat cgaatccgca aagaatcccg 13920
gcaaccgccg gcagccggtg cgccgtcgat taggaagccg cccaagggcg acgagcaacc 13980
agatttttttc gttccgatgc tctatgacgt gggcacccgc gatagtcgca gcatcatgga 14040
cgtggccgtt ttccgtctgt cgaagcgtga ccgacgagct ggcgaggtga tccgctacga 14100
gcttccagac gggcacgtag aggtttccgc agggccggcc ggcatggcca gtgtgtggga 14160
ttacgacctg gtactgatgg cggtttccca tctaaccgaa tccatgaacc gataccggga 14220
agggaaggga gacaagcccg gccgcgtgtt ccgtccacac gttgcggacg tactcaagtt 14280
ctgccggcga gccgatggcg gaaagcagaa agacgacctg gtagaaacct gcattcggtt 14340
aaacaccacg cacgttgcca tgcagcgtac caagaaggcc aagaacggcc gcctggtgac 14400
ggtatccgag ggtgaagcct tgattagccg ctacaagatc gtaaagagcg aaaccgggcg 14460
gccggagtac atcgagatcg agctggctga ttggatgtac cgcgagatca cagaaggcaa 14520
gaacccggac gtgctgacgg ttcaccccga ttacttttttg atcgatcccg gcatcggccg 14580
ttttctctac cgcctggcac gccgcgccgc aggcaaggca gaagccagat ggttgttcaa 14640
gacgatctac gaacgcagtg gcagcgccgg agagttcaag aagttctgtt tcaccgtgcg 14700
caagctgatc gggtcaaatg acctgccgga gtacgatttg aaggaggagg cggggcaggc 14760
tggcccgatc ctagtcatgc gctaccgcaa cctgatcgag ggcgaagcat ccgccggttc 14820
ctaatgtacg gagcagatgc tagggcaaat tgccctagca ggggaaaaag gtcgaaaagg 14880
actctttcct gtggatagca cgtacattgg gaacccaaag ccgtacattg ggaaccggaa 14940
cccgtacatt gggaacccaa agccgtacat tgggaaccgg acacacatgt aagtgactga 15000
tataaaagag aaaaaaggcg atttttccgc ctaaaactct ttaaaactta ttaaaactct 15060
taaaacccgc ctggcctgtg cataactgtc tggccagcgc acagccgaag agctgcaaaa 15120
agcgcctacc cttcggtcgc tgcgctccct acgccccgcc gcttcgcgtc ggcctatcgc 15180
ggccgctggc cgctcaaaaa tggctggcct acggccaggc aatctaccag ggcgcggaca 15240
agccgcgccg tcgccactcg accgccggcg ctgaggtctg cctcgtgaag aaggtgttgc 15300
tgactcatac caggccatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 15360
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 15420
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 15480
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 15540
ctggcgtttt tccataggct ccgccccccct gacgagcatc acaaaaatcg acgctcaagt 15600
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 15660

ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 15720

tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 15780

gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta 15840

tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 15900

gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 15960

tggtggccta actacggcta cactagaaga acagtatttg gtatctgcgc tctgctgaag 16020

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 16080

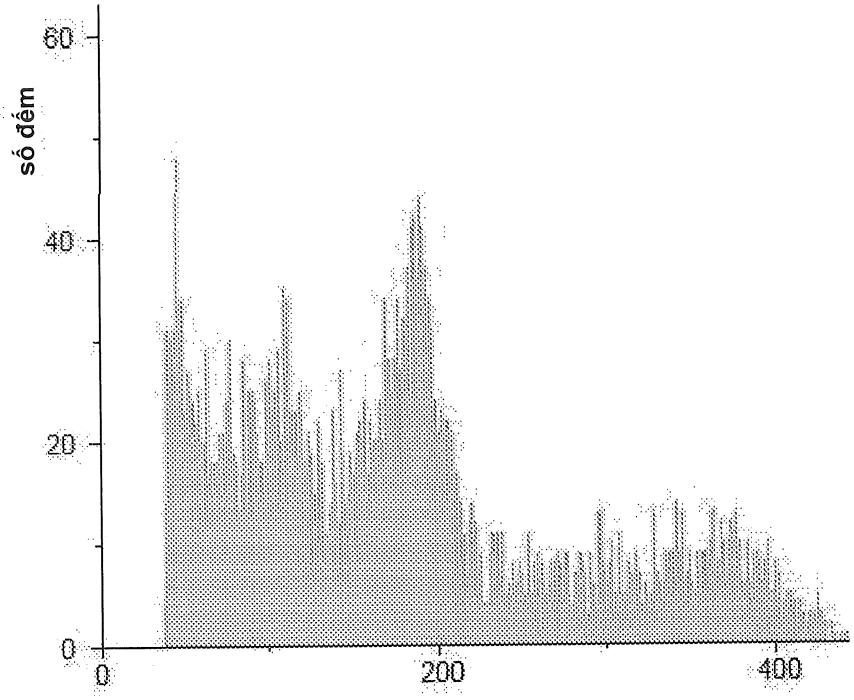
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 16140

gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 16200

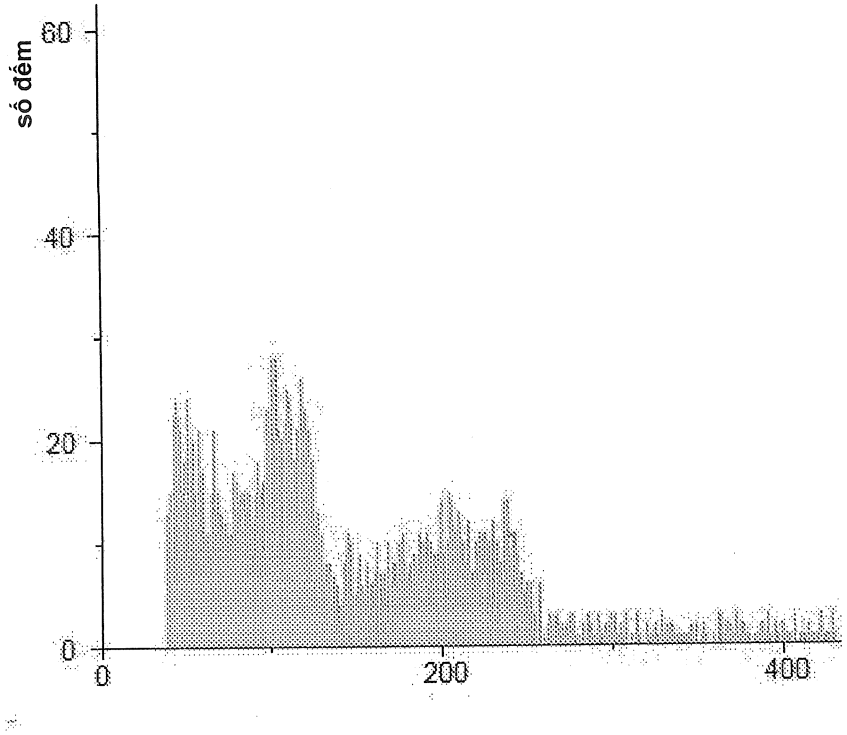
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttgat ccggacaaac 16260

aaacaaatac agtaattta 16279

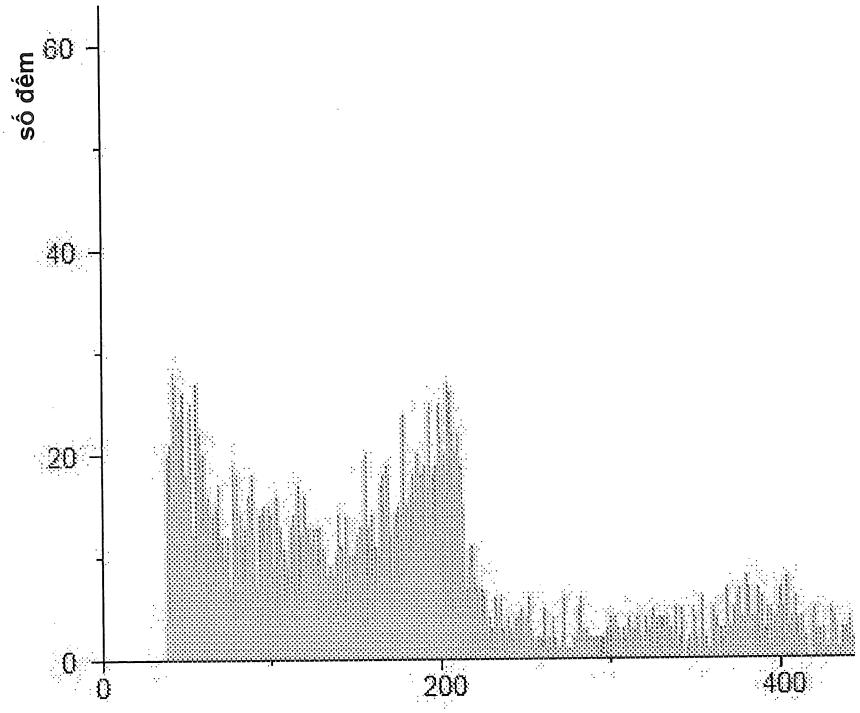
Hình 1.



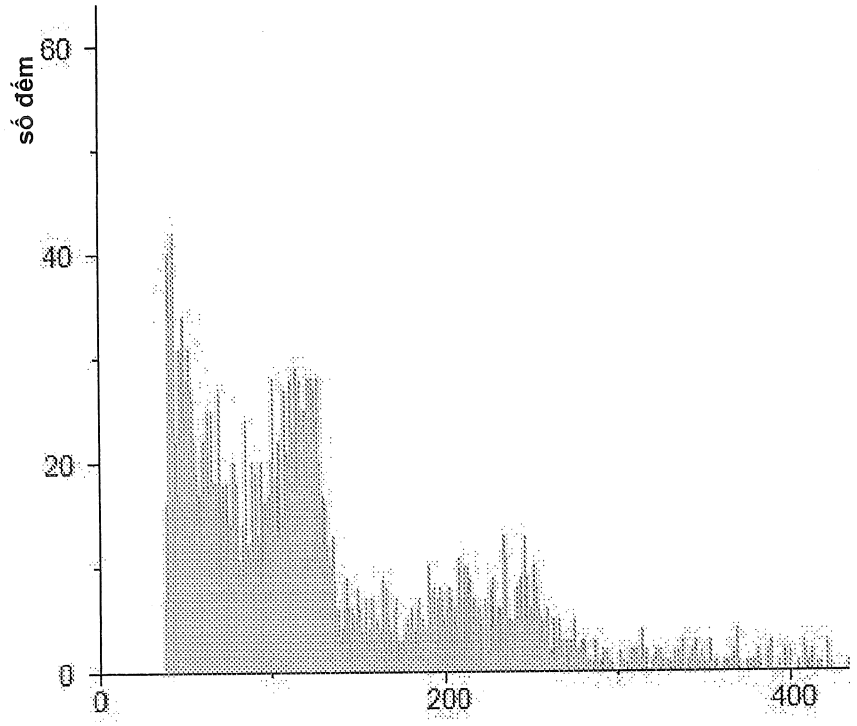
Hình 2.



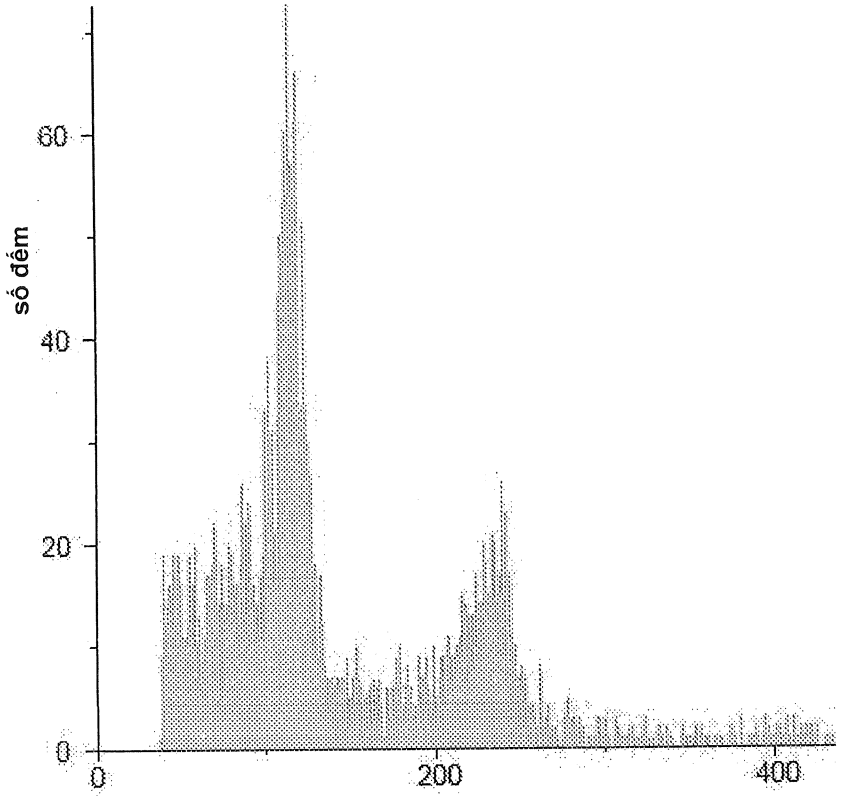
Hình 3.



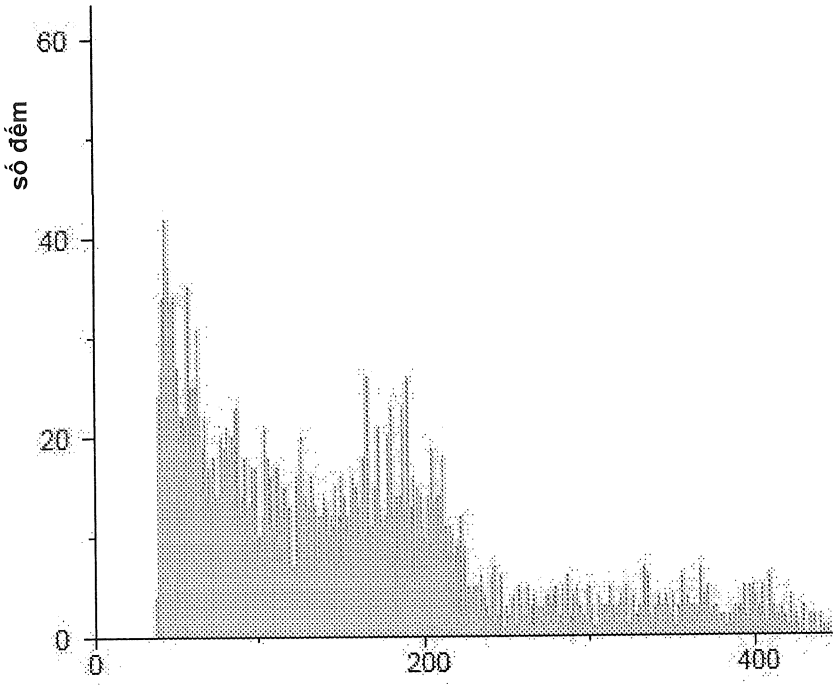
Hình 4.



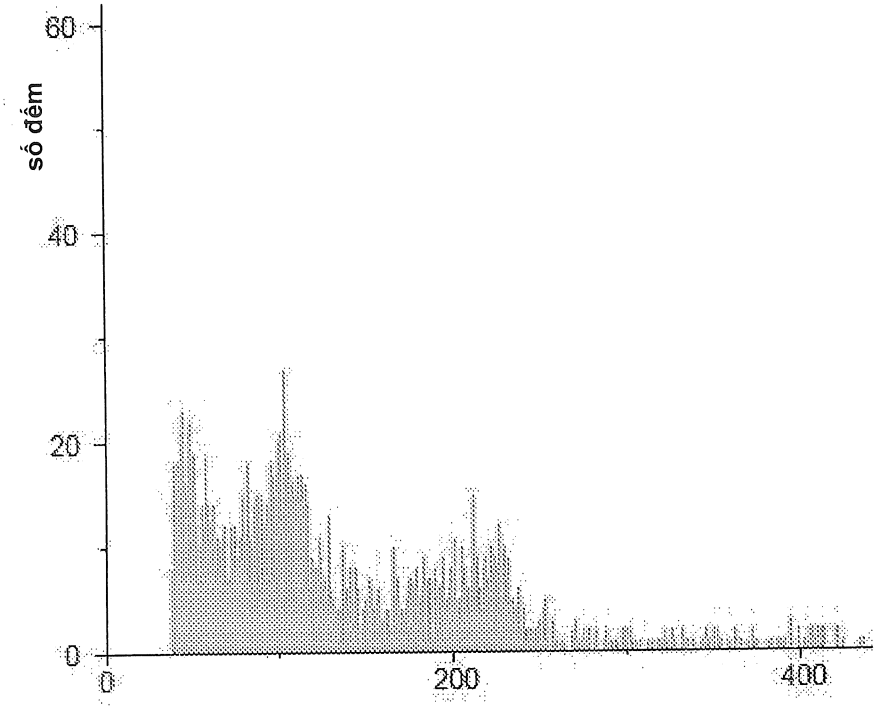
Hình 5.



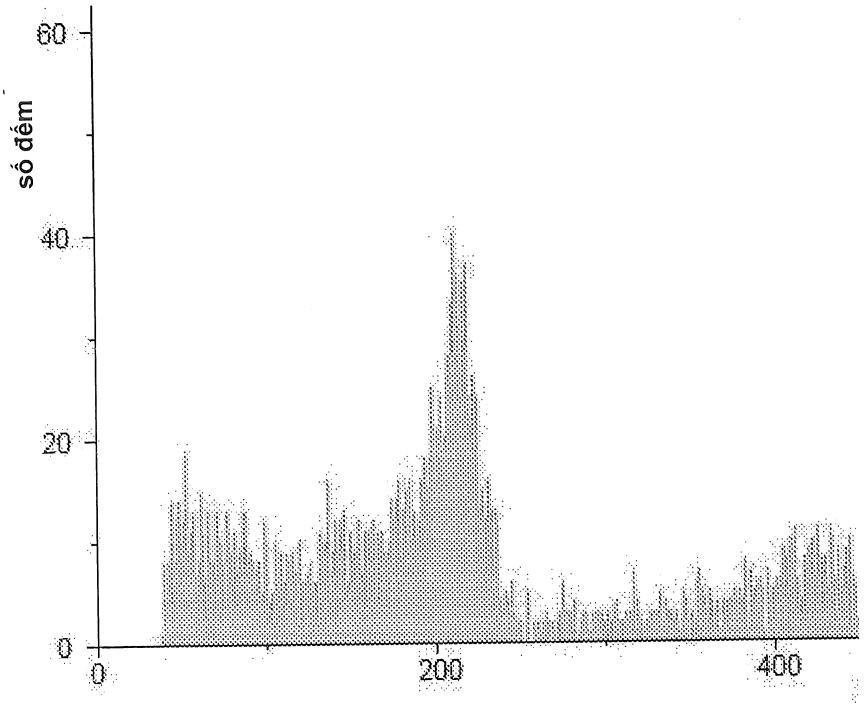
Hình 6.



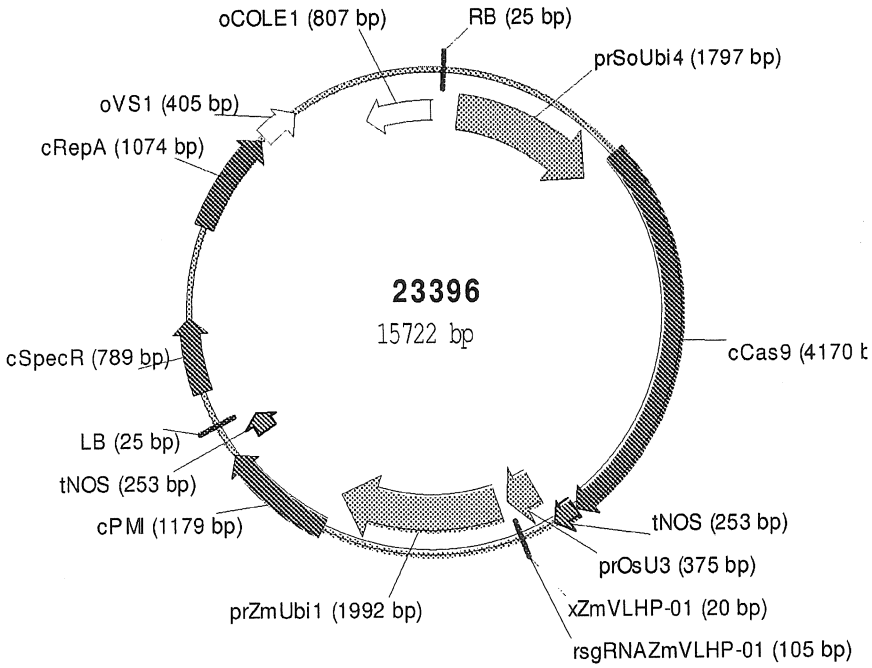
Hình 7.



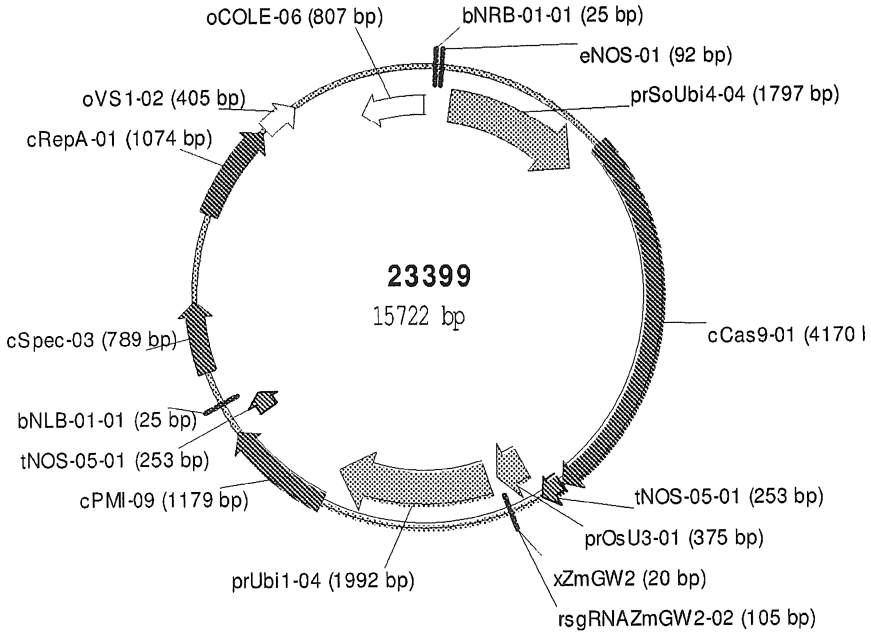
Hình 8.



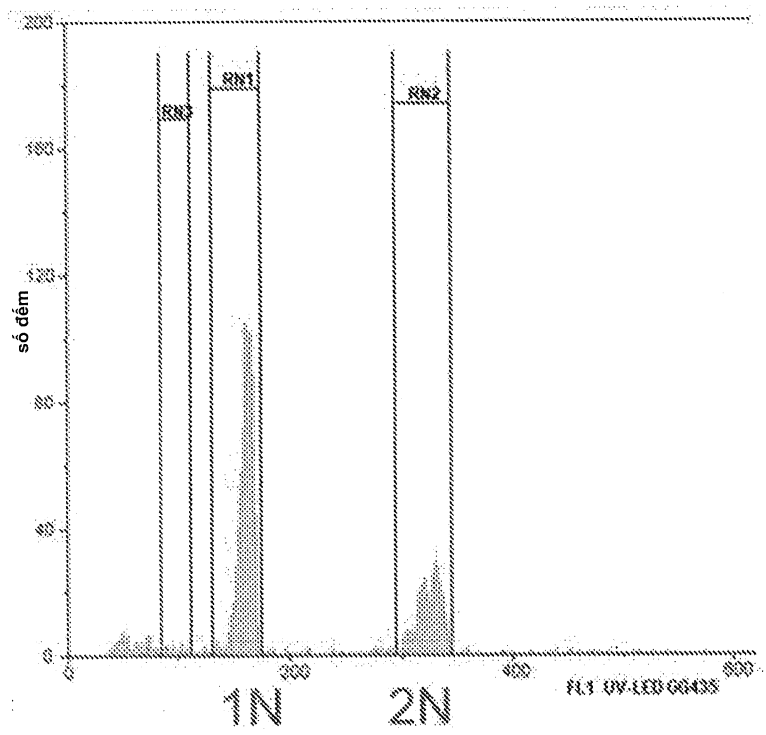
Hình 9.


oCOLE1 (807 bp)
RB (25 bp)
prSoUbi4 (1797 bp)
oVS1 (405 bp)
cRepA (1074 bp)
23396
15722 bp
cSpecR (789 bp)
cCas9 (4170 b
LB (25 bp)
tNOS (253 bp)
cPMI (1179 bp)
tNOS (253 bp)
prOsU3 (375 bp)
prZmUbi1 (1992 bp)
xZmVLHP-01 (20 bp)
rsgRNAZmVLHP-01 (105 bp)

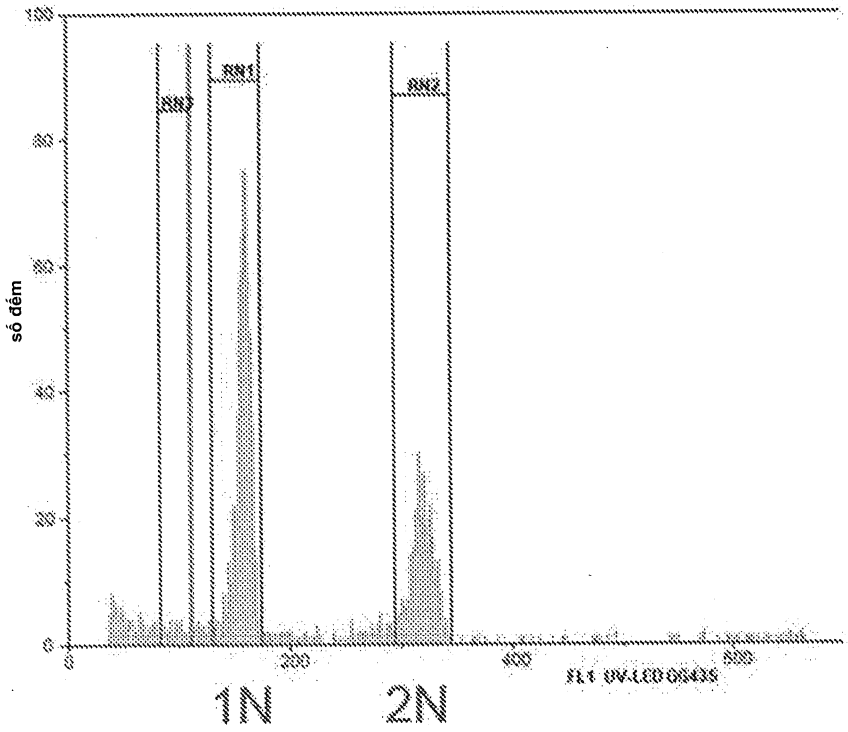
Hình 10.


oCOLE-06 (807 bp)
bNRB-01-01 (25 bp)
eNOS-01 (92 bp)
prSoUbi4-04 (1797 bp)
oVS1-02 (405 bp)
cRepA-01 (1074 bp)
23399
15722 bp
cCas9-01 (4170 l
cSpec-03 (789 bp)
bNLB-01-01 (25 bp)
tNOS-05-01 (253 bp)
cPMI-09 (1179 bp)
tNOS-05-01 (253 bp)
prOsU3-01 (375 bp)
prUbi1-04 (1992 bp)
xZmGW2 (20 bp)
rsgRNAZmGW2-02 (105 bp)

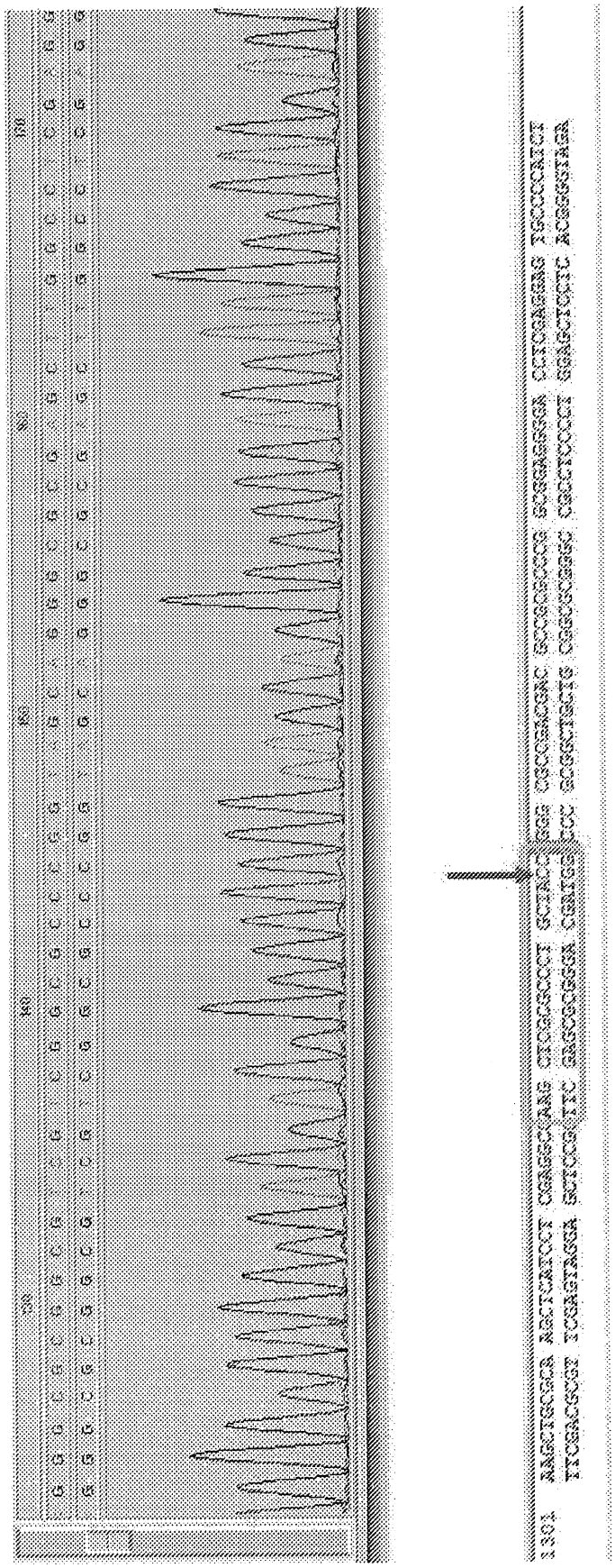
Hình 11.



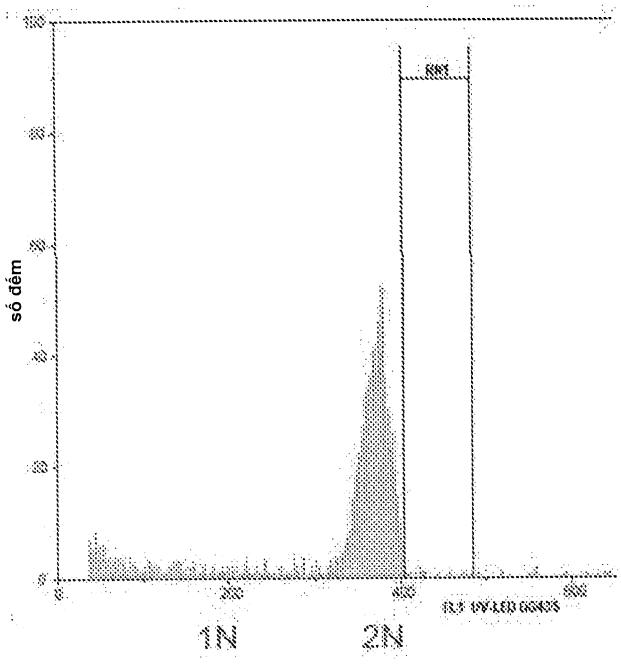
Hình 12.



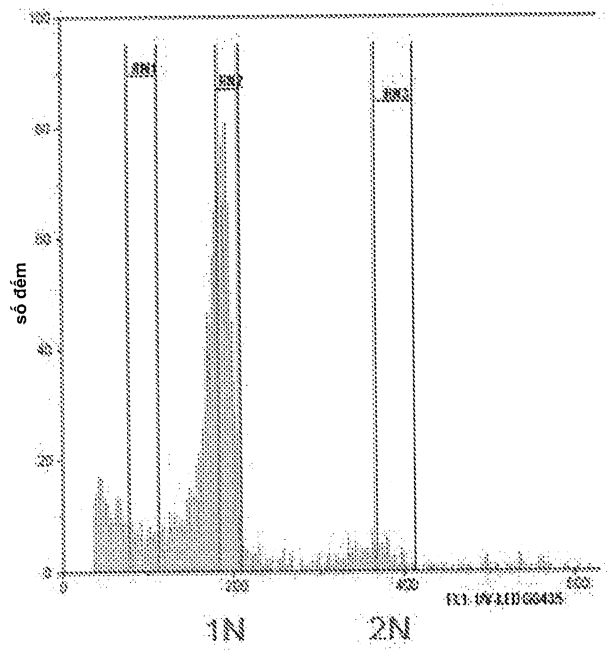
Hình 13.



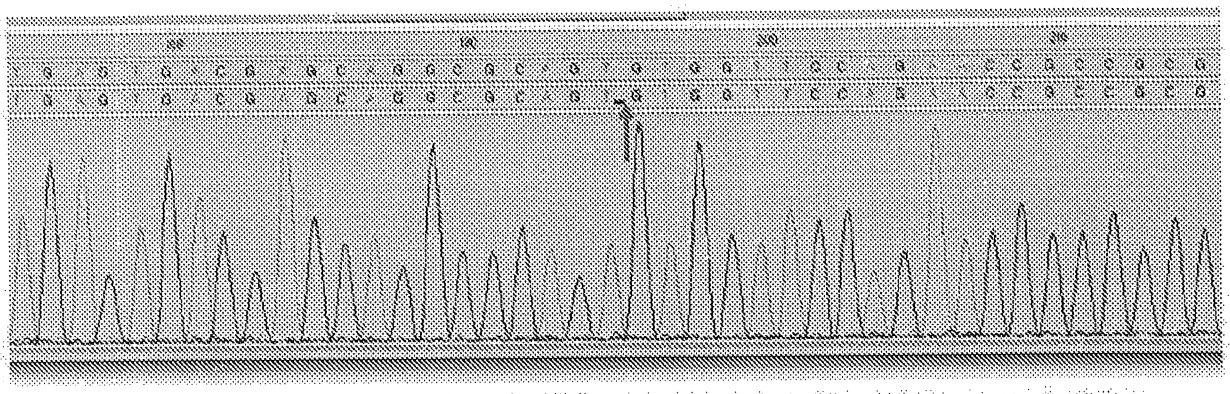
Hình 14.



Hình 15.

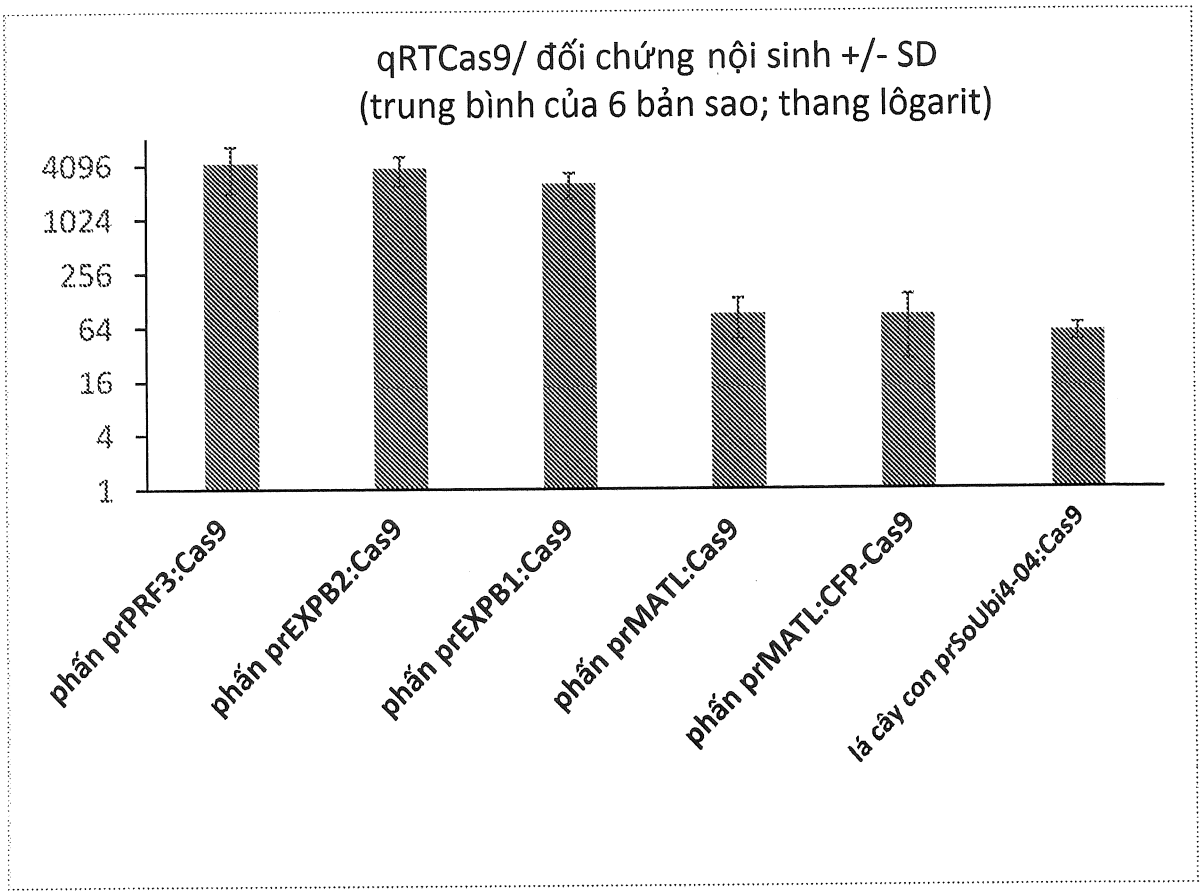


Hình 16.

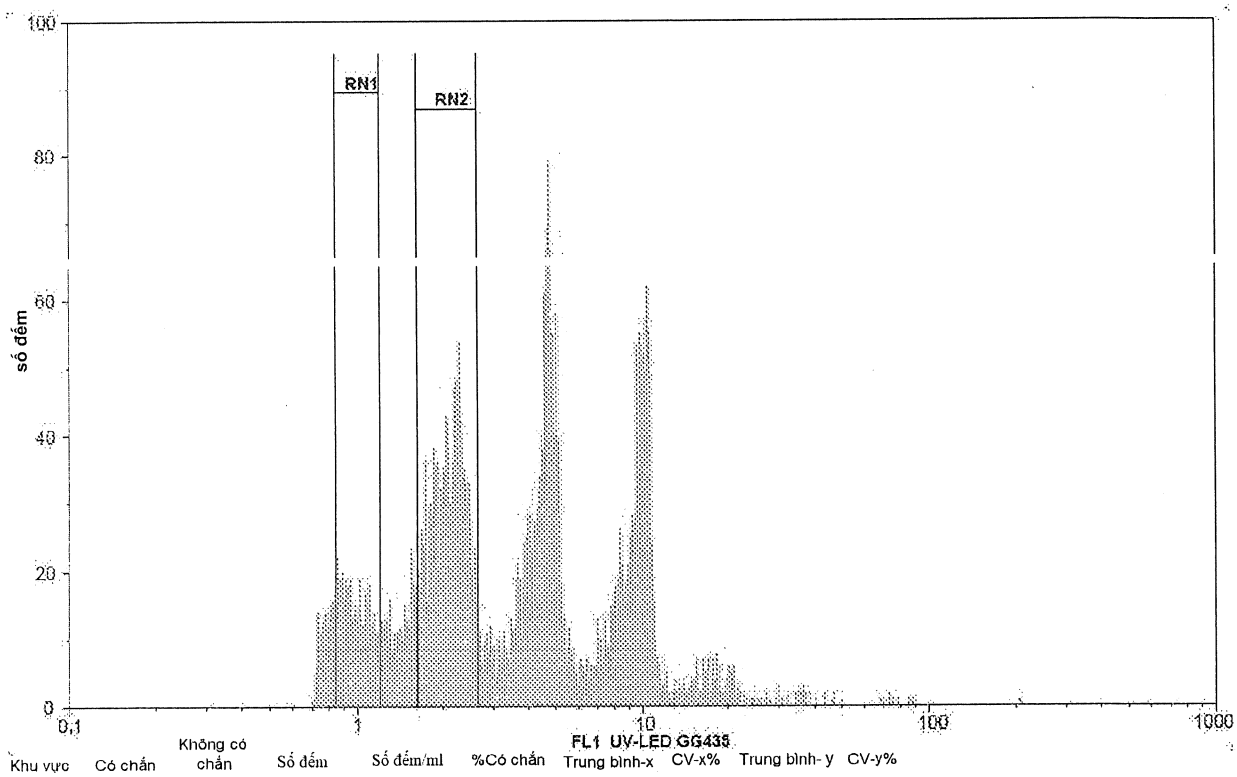


```
1    CTCATCGAGT GTTCNCCGCA ATGCGCTGTT GCTGATTCTC AAGTGCGTGT GGGTGCAGGT
61   GGAGAGCAGA AGAAGGCCGG CCGGCAGCGG CGGAGGAGGA GGGCGAGGCA GGCAGCGGCA
121  GGCGAAGGCG CGGGAGGGGA CGATGCGGCG AAGAAACGCC GGCTGAGTGA CGAGCAGGCG
181  CAGTTCCTGG AGATGAGCTT CAAGAACGAA CGTAAACTGC AAACGCGCCG CAAGGTGCAG
241  CTCGCGGCGG AGCTCGGCCT GGACACCAAG CAGGTCGCGG CGTGGTTCCA GAACCGCCGC
301  GCCCGCTACA AGAGCAAGCT CATCGAGGAG GAGTTCTCCA AGCTCCGCGC GGCACACGAC
361  GCCGTCGTCG TCCACAACTG CCACCTCGAG GCCGAGGTAC AGTGCAACAG TCCGGCTGCC
```

Hình 17.

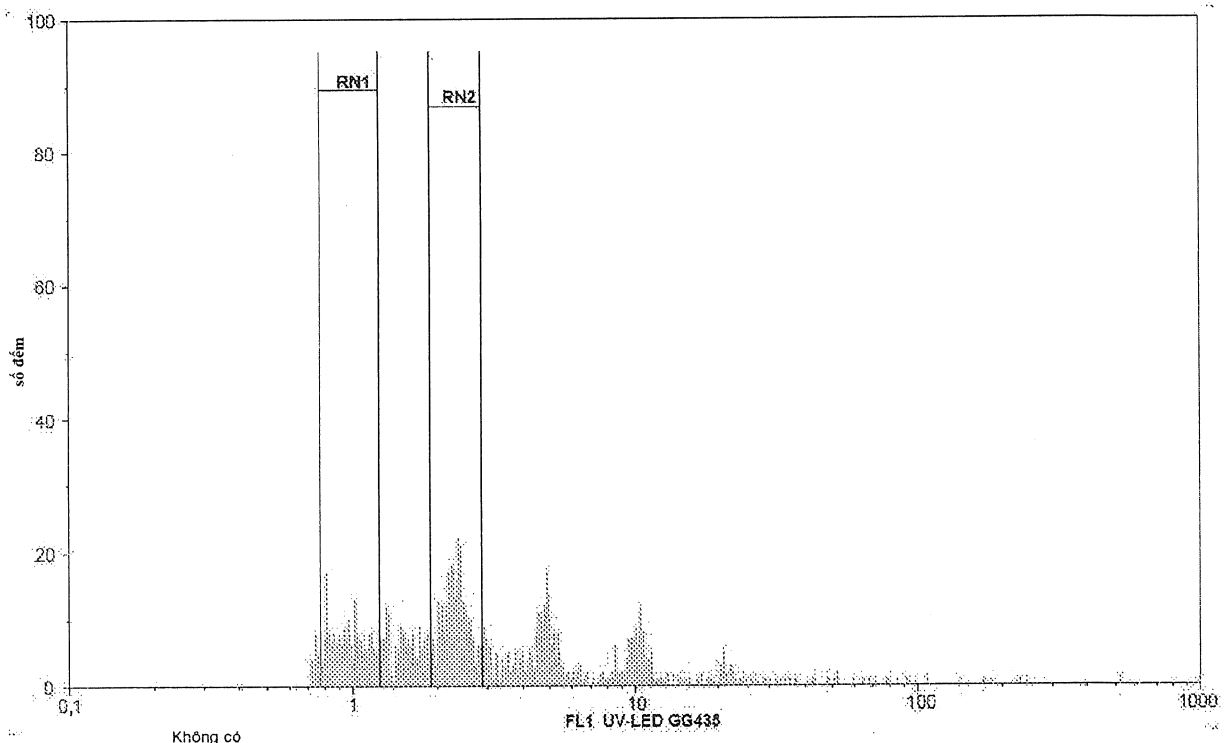

qRTCas9/ đối chứng nội sinh +/- SD
(trung bình của 6 bản sao; thang lôgarit)
4096
1024
256
64
16
4
1
phấn prPRF3:Cas9
phấn prEXPB2:Cas9
phấn prEXPB1:Cas9
phấn prMATL:Cas9
phấn prMATL:CFP-Cas9
lá cây con prSoUbi4-04:Cas9

Hình 18.



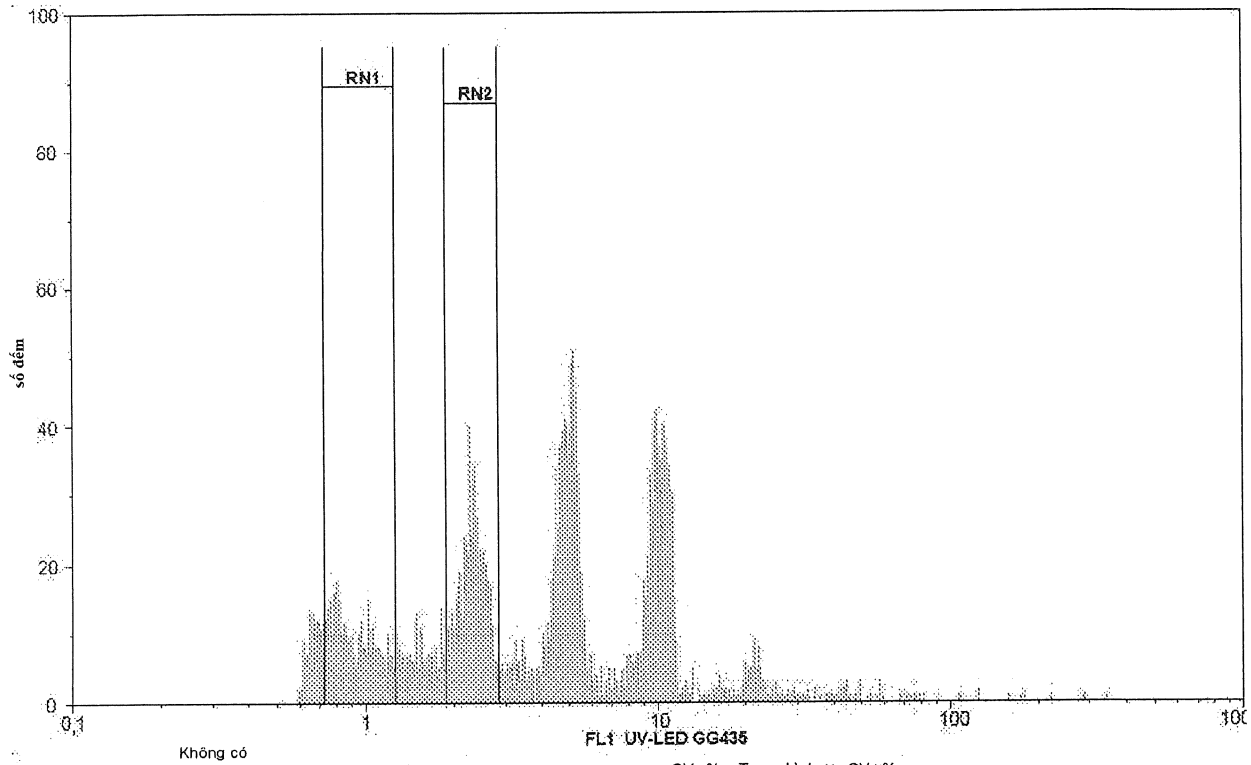
Khu vực	Có chấn	Không có chấn	Số đếm	Số đếm/ml	%Có chấn	Trung bình-x	CV-x%	Trung bình- y	CV-y%
RN1	<Không>	536	536	-	8,70	1,01	10,43	-	-
RN2	<Không>	1618	1618	-	26,27	2,12	12,49	-	-

Hình 19.



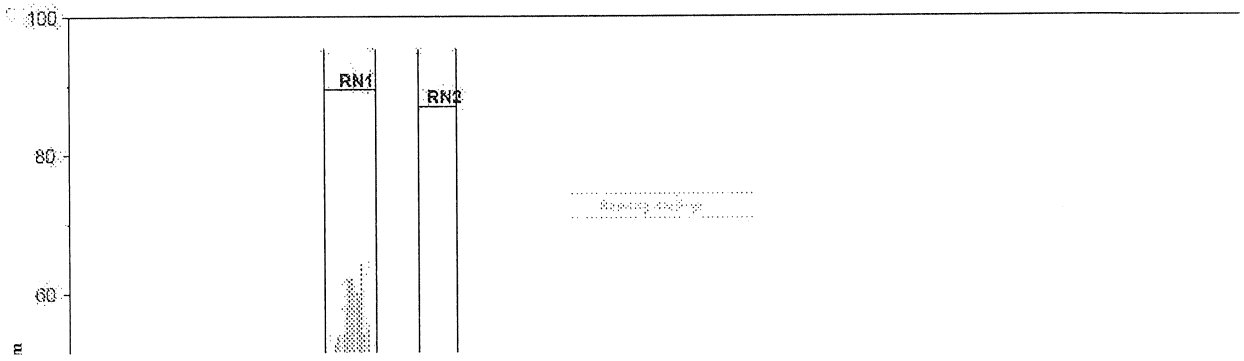
Khu vực	Có chấn	Không có chấn	Số đếm	Số đếm/ml	%Có chấn	Trung bình-x	CV-x%	Trung bình- y	CV-y%
RN1	<Không>	347	347	-	20,39	0,99	14,17	-	-
RN2	<Không>	450	450	-	26,44	2,33	9,45	-	-

Hình 20.

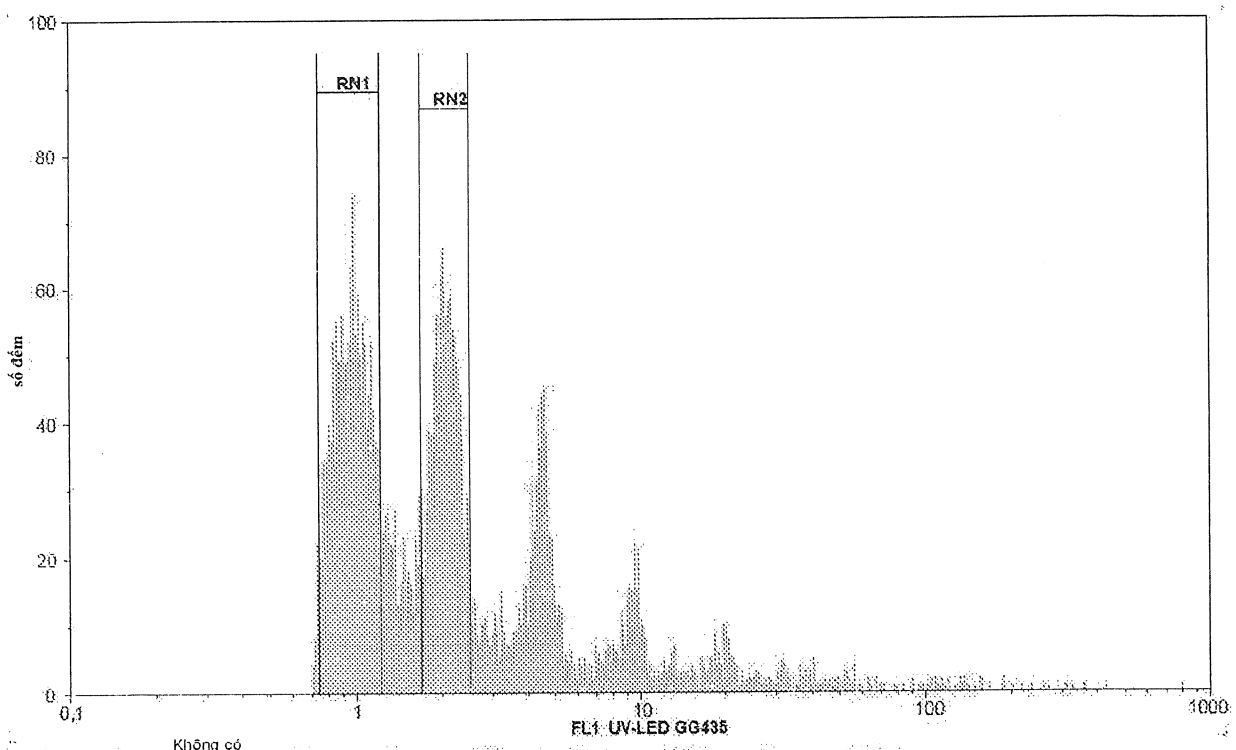


Khu_vực	Có chấn	Không có chấn	Số đếm	Số đếm/ml	%Có chấn	Trung bình-x	CV-x%	Trung bình- y	CV-y%
RN1	<Không>	530	530	-	13,07	0,93	16,32	-	-
RN2	<Không>	732	732	-	18,06	2,33	9,49	-	-

Hình 21.

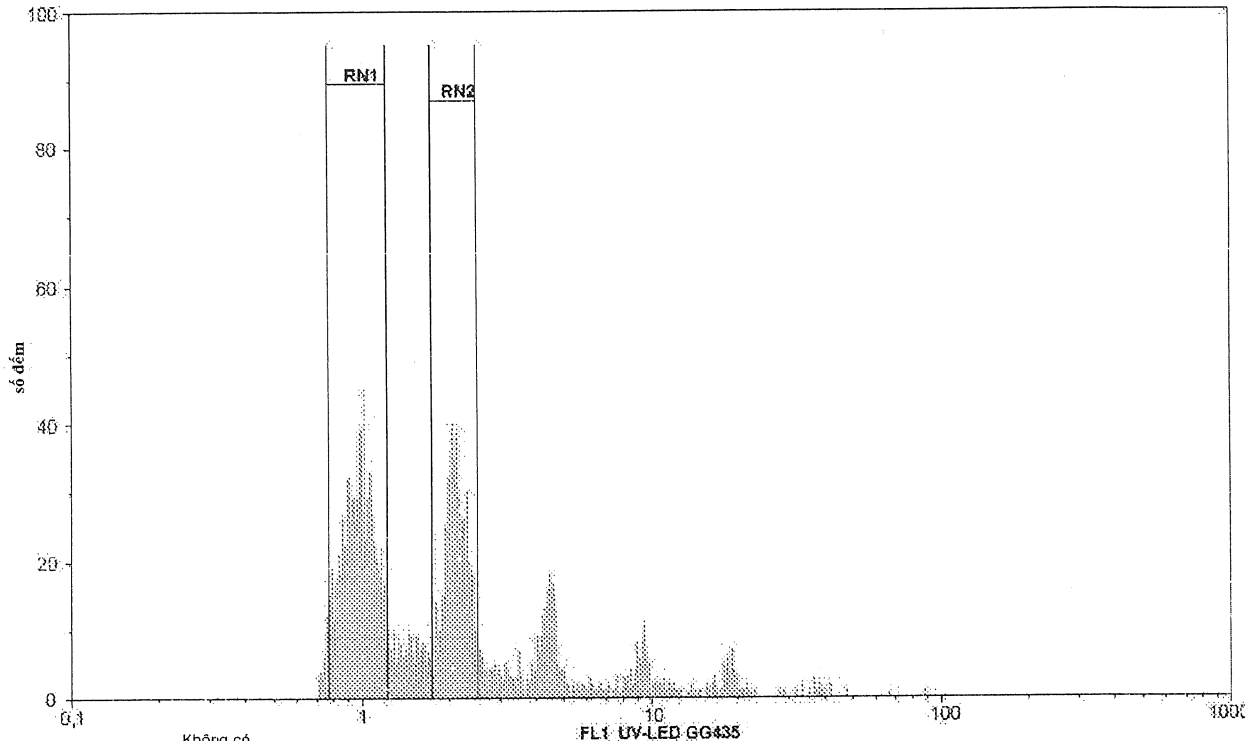


Hình 22.



Khu vực	Có chấn	Không có chấn	Số đếm	Số đếm/ml	%Có chấn	Trung bình-x	CV-x%	Trung bình- y	CV-y%
RN1	<Không>	2350	2350	-	33,51	0,98	12,94	-	-
RN2	<Không>	1740	1740	-	24,81	2,08	9,42	-	-

Hình 23.



Khu vực	Có chấn	Không có chấn	Số đếm	Số đếm/ml	%Có chấn	Trung bình-x	CV-x%	Trung bình- y	CV-y%
RN1	<Không>	1130	1130	-	39,48	0,99	11,67	-	-
RN2	<Không>	789	789	-	27,57	2,12	7,96	-	-

Hình 24.

gRNA
WT
mất 135
cài xen 135
không chỉnh sửa
cài xen A
cài xen T
mất CTG
cài xen A
cài xen G
mất AG
mất TG
cài xen T
cài xen A
cài xen T
mất 16 cài xen 8