



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

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

(11)



1-0039481

(51)⁷ C12N 1/15; C12P 7/40; C12N 15/09

(13) B

(21) 1-2018-03894

(22) 01/02/2017

(86) PCT/JP2017/003648 01/02/2017

(87) WO/2017/135317 10/08/2017

(30) 2016-019676 04/02/2016 JP; 2016-152972 03/08/2016 JP

(45) 25/04/2024 433

(43) 26/11/2018 368A

(73) KAO CORPORATION (JP)

14-10, Nihonbashi-Kayabacho 1-chome, Chuo-ku, Tokyo 1038210 (JP)

(72) TSUBOI, Yuichi (JP).

(74) Công ty Cổ phần Hỗ trợ phát triển công nghệ Detech (DETECH)

(54) PHƯƠNG PHÁP SẢN XUẤT NẤM SỢI ĐỘT BIẾN

(57) Sáng chế đề xuất phương pháp chuẩn bị hiệu quả nấm sợi đột biến. Sáng chế cũng đề xuất phương pháp sản xuất nấm sợi đột biến, trong đó phương pháp này bao gồm việc chuyển ADN nucleaza có thể lập trình và ADN sợi đơn vào nấm sợi chủ, và thay thế vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt cho ADN nucleaza có thể lập trình trong ADN hệ gen của vật chủ bằng ADN sợi đơn thông qua tái tổ hợp tương đồng.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất nấm sợi đột biến.

Tình trạng kỹ thuật của sáng chế

Trong những năm gần đây, công nghệ chỉnh sửa bộ gen sử dụng nucleaza có thể lập trình đã được báo cáo. Chỉnh sửa bộ gen là kỹ thuật thiết kế bộ gen theo cách đột biến điểm trực tiếp, chẳng hạn bằng cách phân cắt đặc hiệu sợi đôi ADN của gen đích trên bộ gen và bao gồm việc xóa bỏ, chèn hoặc thay thế (các) nucleotit trong quá trình sửa chữa ADN bị phân cắt hoặc chèn (các) polynucleotit ngoại lai vào đó. Kỹ thuật như vậy được gọi là TALEN (transcription activator-like effector nuclease - nucleaza giống nhân tố kích hoạt phiên mã), ZFN (zinc-finger nuclease - nucleaza ngón tay kẽm), CRISPR-Cas9 (clustered regularly interspaced short palindromic repeat - nhóm các đoạn ngắn đối xứng lặp lại thường xuyên), CRISPR-Cpf1, endonucleaza tìm nhà (homing endonuclease), bộ thiết kế nhỏ gọn TALEN (compact designer TALEN), v.v.. Đã được báo cáo rằng việc chỉnh sửa bộ gen bằng cách sử dụng nucleaza có thể lập trình có thể cải thiện hiệu quả tái tổ hợp tương đồng của vị trí đích trên ADN hệ gen và có khả năng gây ra tái tổ hợp tương đồng ngay cả ở những sinh vật ban đầu mà có khả năng tái tổ hợp tương đồng thấp (Tài liệu không phải sáng chế 1).

Công nghệ chỉnh sửa bộ gen đã cải thiện hiệu quả tái tổ hợp tương đồng các gen của tế bào động vật có vú mà trước đây tương đối khó được thiết kế. Tài liệu không phải sáng chế 2 báo cáo phương pháp tiếp cận nghiên cứu tính đa

hình nucleotit đơn bằng cách thiết kế một bazơ của vùng hệ gen đích của người bằng cách sử dụng ADN sợi đơn chuyển vào bộ gen bằng công nghệ TALEN. Tài liệu không phải sáng chế 3 báo cáo phương pháp tiếp cận thiết kế gen bằng cách chuyển ADN sợi đơn vào bộ gen của tế bào phôi chuột nhất bằng cách sử dụng công nghệ TALEN. Tuy nhiên, trong phương pháp này, theo báo cáo, vùng hệ gen có từ một đến nhiều bazơ, có chiều dài ngắn hơn 300 bp chỉ được thay thế hoặc xóa bỏ. Tài liệu không phải sáng chế 4 báo cáo phương pháp tiếp cận thiết kế gen bằng cách chuyển ADN sợi đơn vào bộ gen của tế bào phôi chuột công bằng cách sử dụng công nghệ CRISPR-Cas9 và mô tả rằng ADN có kích thước tối đa khoảng 830 bp được chèn vào, và hai ADN sợi đơn cần thiết để chèn hoặc thay thế vùng hệ gen có chuỗi dài hơn. Tuy nhiên, cũng trong phương pháp này, vẫn khó có thể tự chèn một vùng gen vào bộ gen hoặc thay thế vùng hệ gen lớn hiện có bằng vùng gen khác.

Nấm sợi như nấm *Rhizopus* là những vi sinh vật hữu ích mà có thể được sử dụng trong sản xuất vi sinh các chất hữu ích như axit hữu cơ. Có nhu cầu phát triển các thể đột biến của nấm sợi phù hợp hơn để sản xuất chất hữu ích bằng kỹ thuật tái tổ hợp gen.

Nấm sợi *Neurospora crassa* là sinh vật thường được sử dụng trong nghiên cứu tái tổ hợp. Tuy nhiên, *Neurospora crassa* kiểu dại có tỷ lệ tái tổ hợp tương đồng rất thấp và do đó không dễ thiết kế gen bằng phương pháp tái tổ hợp tương đồng.

Sự tái tổ hợp tương đồng ở nấm thuộc chi *Rhizopus* cũng không dễ dàng. Rất hiếm khi nấm thuộc chi *Rhizopus* đưa đoạn ADN được chuyển vào tế bào

thành ADN hệ gen thông qua tái tổ hợp tương đồng. Ngoài ra, trong nấm thuộc chi *Rhizopus*, đoạn ADN được chuyển sẽ được tạo vòng một cách tự nhiên hoặc khuếch đại mà không có điểm khởi đầu phiên mã. Do đó, rất khó để sàng lọc các chủng tái tổ hợp bằng cách sử dụng chỉ thị gen. Hơn nữa, nền tảng di truyền của nấm thuộc chi *Rhizopus* vẫn chưa được nghiên cứu đầy đủ. Do đó, sự phát triển thông thường của các thể đột biến thuộc chi *Rhizopus* vẫn gặp khó khăn về mặt kỹ thuật. Các phương pháp dựa trên việc nhắm vào các phân trình tự tương đồng, chẳng hạn như phương pháp sử dụng *Agrobacterium* (Tài liệu không phải sáng chế 5) và phương pháp biến đổi đầu ADN để chuyển vào tế bào (Tài liệu không phải sáng chế 6) đã được nỗ lực thực hiện nhằm chuyển gen vào bộ gen của nấm thuộc chi *Rhizopus*. Tuy nhiên, các phương pháp này không hiệu quả. Trong khi đó, chỉ có một báo cáo về chủng nấm bị phá vỡ gen thuộc chi *Rhizopus* (Tài liệu không phải sáng chế 7).

Tài liệu không phải sáng chế 1: Science, 2013, 339 (6121): 823-826

Tài liệu không phải sáng chế 2: Int J Mol Sci, 2015, 16: 21128-21137

Tài liệu không phải sáng chế 3: PNAS, 2013, 110 (10): 3782-87

Tài liệu không phải sáng chế 4: Nature Commun, 2016, 20, doi: 10.1038/ncomms10431

Tài liệu không phải sáng chế 5: Mol Genet Genomics, 2004, 271 (4): 499-510

Tài liệu không phải sáng chế 6: Mol Genet Genomics, 2005, 274 (4): 373-383

Tài liệu không phải sáng chế 7: Mol Microbiol, 2010, 77 (3): 587-604

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất phương pháp sản xuất nấm sợi đột biến, trong đó phương pháp này bao gồm bước chuyển ADN nucleaza có thể lập trình và ADN sợi đơn vào nấm sợi chủ, và thay thế vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt cho ADN nucleaza có thể lập trình trong ADN hệ gen của vật chủ bởi ADN sợi đơn thông qua tái tổ hợp tương đồng.

Mô tả chi tiết sáng chế

Tất cả các tài liệu sáng chế, tài liệu không phải sáng chế, và các ấn phẩm khác được trích dẫn ở đây đều được đưa vào đây bằng cách tham khảo toàn bộ nội dung.

Theo sáng chế, các thuật ngữ "ngược chiều" và "xuôi chiều" đối với gen hoặc ADN đề cập đến hướng ngược chiều và xuôi chiều theo hướng phiên mã của gen hoặc ADN trừ khi có quy định khác.

Theo sáng chế, thuật ngữ "liên kết chức năng" giữa gen và vùng kiểm soát đề cập đến liên kết giữa gen và vùng kiểm soát để gen có thể được biểu hiện dưới sự kiểm soát của vùng kiểm soát. Các quy trình "liên kết chức năng" giữa gen và vùng kiểm soát đã được người có hiểu biết trung bình trong lĩnh vực này biết rõ.

Thuật ngữ "chỉnh sửa bộ gen" là kỹ thuật thiết kế bộ gen theo cách đột biến điểm trực tiếp đích bằng cách phân cắt đặc hiệu sợi đôi ADN của locus gen đích trên bộ gen bằng cách sử dụng ADN nucleaza nhân tạo có thể lập trình và gây ra sự xóa bỏ, chèn hoặc thay thế nucleotit trong quá trình sửa chữa ADN bị phân cắt hoặc ví dụ chèn polynucleotit ngoại lai vào đó. Kỹ thuật chỉnh sửa bộ

gen như vậy được biết đến là TALEN (transcription activator-like effector nuclease – nucleaza giống nhân tố kích hoạt phiên mã), ZFN (zinc-finger nuclease - nucleaza ngón tay kẽm), CRISPR-Cas9 (clustered regularly interspaced short palindromic repeat - nhóm các đoạn ngắn đối xứng lặp lại thường xuyên), CRISPR-Cpf1, endonucleaza tìm nhà (homing endonuclease), bộ thiết kế nhỏ gọn TALEN (compact designer TALEN), v.v., thường được đặt tên theo ADN nucleaza nhân tạo có thể lập trình được sử dụng (Nature Reviews Genetics, 2014, 15: 321-334; Nucleic Acids Research, 2011, 39: e82; Nucleic Acids Research, 2006, 34: e149; and Nature communications, 2013, 4: 1762). Kit để chỉnh sửa bộ gen dựa trên các kỹ thuật này có sẵn trên thị trường và có thể được mua từ, chẳng hạn như Life Technologies Corp., Collectis, hoặc Transposagen Biopharmaceuticals, Inc.

Tác giả sáng chế đã thành công trong việc tạo ra nấm sợi đột biến hiệu quả hơn trước đây bằng cách áp dụng công nghệ chỉnh sửa bộ gen sử dụng ADN nucleaza có thể lập trình cho nấm sợi. Tuy nhiên, tỷ lệ tái tổ hợp vẫn không đủ. Đặc biệt, việc thiết kế một vùng lớn của bộ gen, chẳng hạn như sự xóa bỏ hoặc thay thế toàn bộ vùng gen, là rất khó khăn.

Tác giả sáng chế đã tiến hành các nghiên cứu sâu hơn về kỹ thuật tái tổ hợp tương đồng dựa trên việc chỉnh sửa bộ gen của nấm sợi và kết quả là phát hiện ra rằng việc sử dụng ADN sợi đơn như ADN người cho để chuyển vào tế bào cùng với ADN nucleaza có thể lập trình để cải thiện tỷ lệ tái tổ hợp tương đồng và hơn nữa cho phép thiết kế vùng chuỗi dài của bộ gen.

Do đó, sáng chế đề xuất phương pháp sản xuất nấm sợi đột biến. Phương pháp này dựa trên công nghệ chỉnh sửa bộ gen và bao gồm bước chuyển ADN nucleaza có thể lập trình và ADN sợi đơn vào nấm sợi chủ, và thay thế vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt cho ADN nucleaza có thể lập trình trong ADN hệ gen của vật chủ bởi ADN sợi đơn thông qua tái tổ hợp tương đồng.

Theo sáng chế, nấm sợi đột biến có vùng chuỗi dài của bộ gen được thiết kế có thể được sản xuất một cách hiệu quả. Nấm sợi đột biến thu được theo sáng chế có thể là nấm sợi tái tổ hợp thu được bằng cách chuyển gen dị hợp sang vật chủ, có thể là nấm sợi đột biến thu được bằng cách chuyển gen nội sinh sang vật chủ, hoặc có thể là nấm sợi đột biến có đột biến xóa bỏ trong gen của vật chủ.

Theo sáng chế, các ví dụ về nấm sợi ký chủ bao gồm nấm thuộc ngành *Zygomycota*, ngành *Ascomycota*, ngành *Basidiomycota*, và ngành *Chytridiomycota*. Ví dụ về nấm thuộc ngành *Zygomycota* bao gồm nấm thuộc chi *Rhizopus*, chi *Mucor*, và chi *Mortierella*. Ví dụ về nấm thuộc ngành *Ascomycota* bao gồm nấm thuộc chi *Acremonium*, chi *Aspergillus*, chi *Aureobasidium*, chi *Chrysosporium*, chi *Fusarium*, chi *Magnaporthe*, chi *Myceliophthora*, chi *Neurospora*, chi *Paecilomyces*, chi *Penicillium*, chi *Talaromyces*, chi *Thermoascus*, chi *Thielavia*, chi *Tolypocladium*, và chi *Trichoderma*. Ví dụ về nấm thuộc ngành *Basidiomycota* bao gồm nấm thuộc chi *Bjerkandera*, chi *Ceriporiopsis*, chi *Coprinus*, chi *Coriolus*, chi *Cryptococcus*, chi *Folibasidium*, chi *Humicola*, chi *Phanerochaete*, chi *Phlebia*, chi *Pleurotus*, chi *Schizophyllum*, và chi *Trametes*. Ví dụ về nấm thuộc ngành *Chytridiomycota*

bao gồm nấm thuộc chi *Chytridium*, chi *Nowakovskiella*, chi *Olpidium*, chi *Allomyces*, chi *Coelomomyces*, và chi *Monoblepharis*. Trong số đó, nấm thuộc ngành *Zygomycota* hoặc ngành *Ascomycota* được ưu tiên, nấm thuộc ngành *Zygomycota* được ưu tiên hơn, và nấm thuộc chi *Rhizopus* được ưu tiên hơn nữa. Các ví dụ về nấm thuộc chi *Rhizopus* bao gồm *Rhizopus oryzae*, *Rhizopus arrhizus*, *Rhizopus chinensis*, *Rhizopus nigricans*, *Rhizopus tonkinensis*, *Rhizopus tritici*, và *Rhizopus delemar*. Trong số đó, *Rhizopus delemar* và *Rhizopus oryzae* được ưu tiên, và *Rhizopus delemar* được ưu tiên hơn.

ADN nucleaza có thể lập trình được sử dụng trong sáng chế là ADN nucleaza nhận biết và phân cắt một cách đặc hiệu vị trí ADN hệ gen cần phân cắt. Tốt hơn là, ADN nucleaza có thể lập trình được sử dụng trong sáng chế là ADN nucleaza nhân tạo có thể lập trình để sử dụng trong công nghệ chỉnh sửa bộ gen được đề cập ở trên, tốt hơn nữa là từ ADN nucleaza nhân tạo có thể lập trình để sử dụng trong công nghệ TALEN, CRISPR-Cas9, CRISPR-Cpf1, hoặc ZFN. Ví dụ về chúng bao gồm TALEN (transcription activator-like effector nuclease - nucleaza giống nhân tố kích hoạt phiên mã), Cas9 nucleaza hoặc Cas9 (D10A) nucleaza, Cpf1 nucleaza, và ZFN (zinc-finger nuclease - nucleaza ngón tay kẽm). TALEN hoặc ZFN là phức hợp của vùng nhận biết đích liên kết đặc hiệu gần vị trí ADN hệ gen cần phân cắt, và ADN nucleaza, cho phép nhận biết và phân cắt đặc hiệu vị trí ADN hệ gen cần phân cắt. Mặt khác, trong công nghệ CRISPR-Cas9, sự phân cắt ADN được gây ra bằng cách sử dụng ADN nucleaza Cas9 và ARN dẫn đường bắt chước cấu trúc kẹp tóc của tracrARN-crARN. Trong công nghệ CRISPR-Cpf1, sự phân cắt ADN được gây ra bằng cách sử

dụng ADN nucleaza Cpf1 và ARN dẫn đường bất chức crARN. Theo khía cạnh này, tính đặc hiệu phân cắt ADN của Cas9 hoặc Cpf1 được xác định bởi ARN dẫn đường và cho phép phân cắt tại vị trí đích. Những người có hiểu biết trung bình trong lĩnh vực này có thể thu được trình tự ADN của vị trí ADN hệ gen được phân cắt và vùng lân cận của nó và thiết kế ADN nucleaza nhân tạo hoặc ARN dẫn đường có vùng nhận biết đích thích hợp của nó.

Trong trường hợp chuyển ADN nucleaza có thể lập trình đến vật chủ, polyribonucleotit nucleaza hoặc polypeptit nucleaza của nó có thể được chuyển trực tiếp đến tế bào chủ. Ưu tiên chuyển polynucleotit mã hóa nucleaza sang tế bào chủ và biểu hiện nucleaza trong các tế bào.

Trong trường hợp chuyển polynucleotit mã hóa ADN nucleaza có thể lập trình sang tế bào chủ, tốt hơn là từ vector biểu hiện chứa polynucleotit mã hóa ADN nucleaza được chuyển đến tế bào chủ. Vector biểu hiện không bị giới hạn cụ thể miễn là vector biểu hiện được giữ lại ổn định trong tế bào chủ và có khả năng tăng sinh trong đó. Vector biểu hiện có thể được lựa chọn một cách thích hợp bởi những người có hiểu biết trung bình trong lĩnh vực này. Ví dụ về vector biểu hiện phù hợp đối với nấm thuộc chi *Rhizopus* bao gồm pUC18/19, pUC118/119, pBR322, pMW218/219, pPTR1/2 (Takara Bio Inc.), pRI909/910 (Takara Bio Inc.), pDJB2 (D.J. Ballance et al., Gene, 36, 321-331, 1985), pAB4-1 (van Hartingsveldt W et al., Mol Gen Genet, 206, 71-75, 1987), pLeu4 (M.I.G. Roncero et al., Gene, 84, 335-343, 1989), pPyr225 (C.D. Skory et al., Mol Genet Genomics, 268, 397-406, 2002), và pFG1 (Gruber, F. et al., Curr Genet, 18, 447-451, 1990).

Để biểu hiện hiệu quả ADN nucleaza có thể lập trình trong tế bào chủ, được ưu tiên là polynucleotit mã hóa ADN nucleaza có thể lập trình nên được liên kết chức năng với trình tự khởi đầu có chức năng trong tế bào chủ. Loại trình tự khởi đầu được sử dụng có thể được lựa chọn một cách thích hợp bởi những người có hiểu biết trung bình trong lĩnh vực này theo tế bào chủ. Ví dụ về trình tự khởi đầu thực hiện chức năng trong nấm thuộc chi *Rhizopus* bao gồm trình tự khởi đầu ldhA (Patent Hoa Kỳ số 6268189), trình tự khởi đầu pgk1 (WO 2001/73083), trình tự khởi đầu pgk2 (WO 2001/72967), trình tự khởi đầu pdcA và trình tự khởi đầu amyA (Archives of Microbiology, 2006, 186: 41-50), các trình tự khởi đầu tef và 18S rARN (Công bố đơn xin cấp Patent Hoa Kỳ số 2010/112651), và trình tự khởi đầu adh1 (JP-A-2015-155759).

Trong trường hợp sử dụng ADN nucleaza có thể lập trình, phá vỡ, v.v., ADN đích bởi ADN nucleaza có thể lập trình được xúc tiến bằng cách biểu hiện exonucleaza ngoại lai trong vật chủ (Scientific Report, 2013, 3: 1253, DOI: 10.1038/srep01253; and Nat Methods, 2012, 9: 973-975). Do đó, theo một phương án được ưu tiên, exonucleaza hoặc polynucleotit mã hóa exonucleaza được chuyển đến tế bào chủ cùng với ADN nucleaza có thể lập trình hoặc polynucleotit mã hóa ADN nucleaza có thể lập trình, tiếp theo là đồng biểu hiện của chúng trong tế bào chủ. Exonucleaza không bị giới hạn cụ thể miễn là exonucleaza có nguồn gốc từ nấm sợi. Các ví dụ về chúng bao gồm tốt hơn là exonucleaza có nguồn gốc từ nấm thuộc chi *Rhizopus*, tốt hơn nữa là nấm thuộc về exonucleaza 1 hoặc exonucleaza 2, tốt hơn nữa là exonucleaza có nguồn gốc từ *Rhizopus oryzae* hoặc *Rhizopus delemar*.

Tốt hơn là, vectơ biểu hiện chứa polynucleotit mã hóa exonucleaza được chuyển vào tế bào chủ, và exonucleaza được đồng biểu hiện với ADN nucleaza có thể lập trình. Tốt hơn là, polynucleotit mã hóa exonucleaza được liên kết chức năng với trình tự khởi đầu trên vectơ biểu hiện. Các loại vectơ biểu hiện và trình tự khởi đầu có thể được sử dụng giống như trong trường hợp của ADN nucleaza.

ADN sợi đơn được sử dụng trong sáng chế được sử dụng như ADN người cho đối với ADN hệ gen của nấm sợi chủ. Theo sáng chế, "ADN người cho" đề cập đến đoạn ADN được sử dụng để thiết kế ADN hệ gen của vật chủ thông qua tái tổ hợp tương đồng. ADN người cho là ADN sợi đơn chứa hai trình tự ADN (còn được gọi tương ứng là trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều) tương đồng với ADN hệ gen của vật chủ. Trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều lần lượt chứa các vùng tương đồng với vùng ngược chiều và vùng xuôi chiều của vị trí được phân cắt bởi ADN nucleaza có thể lập trình trong ADN hệ gen của vật chủ. Vị trí phân cắt cho ADN nucleaza có thể lập trình có thể được đặt giữa trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều. Ngoài ra, một phần hoặc toàn bộ vị trí phân cắt có thể được chứa trong trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều bất kỳ. ADN người cho có thể có thêm trình tự chèn để chèn vào ADN hệ gen của nấm sợi chủ, giữa trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều. Trình tự chèn không bị giới hạn cụ thể và có thể chứa trình tự của vùng gen tùy ý, ví dụ, vùng mã hóa của gen, vùng kiểm soát như trình tự khởi đầu, intron, và các vùng không mã hóa khác và có thể

chứa thêm gen đánh dấu chọn lọc chẳng hạn như gen kháng thuốc. Tốt hơn là, trong trình tự chèn, vùng mã hóa của gen được liên kết chức năng với vùng kiểm soát.

ADN sợi đơn làm ADN người cho được chuyển vào vật chủ, gây ra sự tái tổ hợp tương đồng giữa vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt trong ADN hệ gen của vật chủ do trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều của nó để trình tự này từ vùng ngược chiều đến vùng xuôi chiều được thay thế bằng trình tự đó. Do đó, tùy thuộc vào sự có mặt hoặc không có mặt trình tự chèn trong ADN sợi đơn, sự thay thế, chèn hoặc xóa bỏ điểm trực tiếp của trình tự xảy ra trong ADN hệ gen của vật chủ sau khi tái tổ hợp tương đồng. Cụ thể hơn, khi ADN sợi đơn chứa trình tự chèn, thì trình tự chèn này được thay thế hoặc chèn vào ADN hệ gen của vật chủ. Ngoài ra, khi ADN sợi đơn về cơ bản chỉ có trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều và không chứa trình tự chèn, sự xóa bỏ được thực hiện trong ADN hệ gen của vật chủ.

ADN người cho có thể được xây dựng bằng cách chuẩn bị các ADN mã hóa trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều, và nếu cần, trình tự chèn sẽ được chuyển đến vật chủ, và liên kết với các ADN này. Những người có hiểu biết trung bình trong lĩnh vực này có thể thu được trình tự ADN của vùng lân cận vị trí phân cắt của hệ gen vật chủ, hoặc vùng đích được thay thế bằng ADN sợi đơn, và thiết kế các ADN của trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều trên cơ sở của chúng. ADN mã hóa trình tự chèn có thể là ADN của một loại vùng gen bất kỳ hoặc có thể là ADN chứa

hai hoặc nhiều loại vùng gen. Ví dụ, trình tự chèn có thể bao gồm đoạn ADN mã hóa gen đã chuyển sang vật chủ, và đoạn ADN mã hóa gen đánh dấu như gen kháng thuốc. Trong ví dụ được ưu tiên, mỗi đoạn ADN của trình tự tương đồng ngược chiều, đoạn ADN mã hóa trình tự chèn, và đoạn ADN của trình tự tương đồng xuôi chiều được xây dựng theo phương pháp tiêu chuẩn và sau đó được liên kết theo thứ tự, để thu được ADN người cho.

Đầu tiên, ADN người cho có thể được xây dựng như ADN sợi đôi và sau đó bị biến tính thành ADN sợi đơn. Ví dụ, plasmid ADN sợi đôi chứa trình tự tương đồng ngược chiều, trình tự tương đồng xuôi chiều, và trình tự chèn được xây dựng bằng cách tách dòng hoặc tương tự và được sử dụng làm khuôn trong PCR để thu được đoạn ADN sợi đôi. Plasmid hoặc ADN sợi đôi có thể bị siêu âm, làm nguội sau khi biến tính nhiệt, xử lý bằng chất kiềm mạnh như NaOH, v.v., và do đó được biến tính thành ADN sợi đơn.

Ngoài ra, ADN người cho có thể được xây dựng làm ADN sợi đơn ngay từ đầu. ADN sợi đơn có thể được chuẩn bị bằng phương pháp tổng hợp hóa học (có thể mua từ Medical & Biological Laboratories Co., Ltd., v.v.), PCR bất đối xứng, hoặc tương tự. ADN sợi đơn cũng có thể được chuẩn bị bằng cách sử dụng thực khuẩn thể Ff như M13 hoặc plasmid như pBluescript có Fori, và *E. coli* như JM109 có yếu tố f. Các ví dụ khác về phương pháp chuẩn bị ADN sợi đơn có thể bao gồm phương pháp biotin hóa một trong các đoạn mồi, thực hiện PCR, thu sản phẩm PCR bằng cách sử dụng các hạt streptavidin, sau đó ly tâm sản phẩm PCR thành các ADN sợi đơn, và rửa giải ADN sợi đơn chưa được

biotin hóa, và phương pháp phosphoryl hóa một trong các đoạn môi, và thực hiện PCR, tiếp theo là xử lý bằng Lambda Exonucleaza.

Các ví dụ được ưu tiên về phương pháp chuẩn bị ADN sợi đơn bao gồm phương pháp sử dụng thực khuẩn thể Ff và *E. coli* có yếu tố f, phương pháp sử dụng đoạn môi được biotin hóa, và phương pháp sử dụng đoạn môi được phosphoryl hóa. Trong số đó, phương pháp sử dụng đoạn môi được phosphoryl hóa được ưu tiên hơn.

Trong ADN sợi đơn, mỗi chiều dài của trình tự tương đồng ngược chiều và trình tự tương đồng xuôi chiều tốt hơn là từ 5 bp trở lên, 10 bp trở lên, 15 bp trở lên, 20 bp trở lên, 25 bp trở lên, 30 bp trở lên, 35 bp trở lên, 40 bp trở lên, 45 bp trở lên, 50 bp trở lên, 100 bp trở lên, 250 bp trở lên, 500 bp trở lên, 750 bp trở lên, hoặc 1000 bp trở lên. Khi ADN sợi đơn có trình tự chèn giữa các trình tự tương đồng ngược chiều và xuôi chiều, chiều dài chuỗi của trình tự chèn có thể là từ 1 bp trở lên và tốt hơn là từ 0,3 kbp trở lên, tốt hơn nữa là từ 1 kbp trở lên, tốt hơn nữa là từ 3 kbp trở lên và tốt hơn là từ 50 kbp trở xuống, tốt hơn nữa là từ 20 kbp trở xuống, tốt hơn nữa là từ 10 kbp trở xuống, còn tốt hơn nữa là từ 9 kbp trở xuống, còn tốt hơn nữa là từ 5 kbp trở xuống. Ngoài ra, chiều dài chuỗi của trình tự chèn tốt hơn là từ 1 bp đến 50 kbp, từ 0,3 kbp đến 50 kbp, từ 1 kbp đến 50 kbp, hoặc từ 3 kbp đến 50 kbp, tốt hơn nữa là từ 1 bp đến 20 kbp, từ 0,3 kbp đến 20 kbp, từ 1 kbp đến 20 kbp, hoặc từ 3 kbp đến 20 kbp, tốt hơn nữa là từ 1 bp đến 10 kbp, từ 0,3 kbp đến 10 kbp, từ 1 kbp đến 10 kbp, hoặc từ 3 kbp đến 10 kbp, còn tốt hơn nữa là từ 1 bp đến 9 kbp, từ 0,3 kbp đến 9 kbp, từ 1 kbp

đến 9 kbp, hoặc từ 3 kbp đến 9 kbp, còn tốt hơn nữa là từ 1 bp đến 5 kbp, từ 0,3 kbp đến 5 kbp, từ 1 kbp đến 5 kbp, hoặc từ 3 kbp đến 5 kbp.

Do đó, chiều dài chuỗi của ADN sợi đơn được sử dụng trong sáng chế tốt hơn là từ 10 bp trở lên, tốt hơn nữa là từ 20 bp trở lên, tốt hơn nữa là từ 30 bp trở lên, còn tốt hơn nữa là từ 40 bp trở lên, còn tốt hơn nữa là từ 50 bp trở lên, còn tốt hơn nữa là từ 60 bp trở lên, còn tốt hơn nữa là từ 70 bp trở lên, còn tốt hơn nữa là từ 80 bp trở lên, còn tốt hơn nữa là từ 90 bp trở lên, còn tốt hơn nữa là từ 100 bp trở lên, còn tốt hơn nữa là từ 200 bp trở lên, còn tốt hơn nữa là từ 500 bp trở lên, còn tốt hơn nữa là từ 1000 bp trở lên, còn tốt hơn nữa là từ 1500 bp trở lên, còn tốt hơn nữa là từ 2000 bp trở lên, và tốt hơn là từ 50 kbp trở xuống, tốt hơn nữa là từ 20 kbp trở xuống, tốt hơn nữa là từ 12 kbp trở xuống, còn tốt hơn nữa là từ 11 kbp trở xuống.

Ngoài ra, chiều dài chuỗi của ADN sợi đơn được sử dụng trong sáng chế tốt hơn là từ 10 bp đến 50 kbp, từ 10 bp đến 20 kbp, từ 10 bp đến 12 kbp, từ 10 bp đến 11 kbp, từ 20 bp đến 50 kbp, từ 20 bp đến 20 kbp, từ 20 bp đến 12 kbp, từ 20 bp đến 11 kbp, từ 30 bp đến 50 kbp, từ 30 bp đến 20 kbp, từ 30 bp đến 12 kbp, từ 30 bp đến 11 kbp, từ 40 bp đến 50 kbp, từ 40 bp đến 20 kbp, từ 40 bp đến 12 kbp, từ 40 bp đến 11 kbp, từ 50 bp đến 50 kbp, từ 50 bp đến 20 kbp, từ 50 bp đến 12 kbp, từ 50 bp đến 11 kbp, từ 60 bp đến 50 kbp, từ 60 bp đến 20 kbp, từ 60 bp đến 12 kbp, từ 60 bp đến 11 kbp, từ 70 bp đến 50 kbp, từ 70 bp đến 20 kbp, từ 70 bp đến 12 kbp, từ 70 bp đến 11 kbp, từ 80 bp đến 50 kbp, từ 80 bp đến 20 kbp, từ 80 bp đến 12 kbp, từ 80 bp đến 11 kbp, từ 90 bp đến 50 kbp, từ 90 bp đến 20 kbp, từ 90 bp đến 12 kbp, từ 90 bp đến 11 kbp, từ 100 bp

đến 50 kbp, từ 100 bp đến 20 kbp, từ 100 bp đến 12 kbp, từ 100 bp đến 11 kbp, từ 200 bp đến 50 kbp, từ 200 bp đến 20 kbp, từ 200 bp đến 12 kbp, từ 200 bp đến 11 kbp, từ 500 bp đến 50 kbp, từ 500 bp đến 20 kbp, từ 500 bp đến 12 kbp, từ 500 bp đến 11 kbp, từ 1000 bp đến 50 kbp, từ 1000 bp đến 20 kbp, từ 1000 bp đến 12 kbp, từ 1000 bp đến 11 kbp, từ 1500 bp đến 50 kbp, từ 1500 bp đến 20 kbp, từ 1500 bp đến 12 kbp, từ 1500 bp đến 11 kbp, từ 2000 bp đến 50 kbp, từ 2000 bp đến 20 kbp, từ 2000 bp đến 12 kbp, hoặc từ 2000 bp đến 11 kbp.

Polynucleotit mã hóa ADN nucleaza có thể lập trình hoặc vector chứa polynucleotit, và ADN sợi đơn có thể được chuyển sang tế bào chủ bằng cách sử dụng phương pháp biến nạp chung, ví dụ, phương pháp điện chuyển, phương pháp biến nạp, phương pháp chuyển nạp, phương pháp tiếp hợp, phương pháp dung hợp tế bào nguyên sinh, phương pháp bắn gen, hoặc phương pháp sử dụng *Agrobacterium*.

Các tế bào tái tổ hợp đã trải qua quá trình tái tổ hợp tương đồng quan tâm bởi ADN sợi đơn được chuyển có thể được chọn trên cơ sở xóa bỏ gen được mang ban đầu, biểu hiện của gen được chuyển, biểu hiện của gen đánh dấu, v.v..

Sáng chế còn đề cập đến các chất, phương pháp sản xuất, sử dụng, phương pháp, v.v., sau đây như là các phương án được lấy làm ví dụ của sáng chế. Tuy nhiên, sáng chế không bị giới hạn bởi các phương án này.

(1) Phương pháp sản xuất nấm sợi đột biến, trong đó phương pháp này bao gồm bước chuyển ADN nucleaza có thể lập trình và ADN sợi đơn vào nấm sợi chủ, và thay thế vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt cho ADN

nucleaza có thể lập trình trong ADN hệ gen của vật chủ bởi ADN sợi đơn thông qua tái tổ hợp tương đồng.

(2) Phương pháp theo điểm (1), trong đó tốt hơn là chiều dài của ADN sợi đơn là:

từ 10 bp đến 50 kbp, từ 10 bp đến 20 kbp, từ 10 bp đến 12 kbp, từ 10 bp đến 11 kbp, từ 20 bp đến 50 kbp, từ 20 bp đến 20 kbp, từ 20 bp đến 12 kbp, từ 20 bp đến 11 kbp, từ 30 bp đến 50 kbp, từ 30 bp đến 20 kbp, từ 30 bp đến 12 kbp, từ 30 bp đến 11 kbp, từ 40 bp đến 50 kbp, từ 40 bp đến 20 kbp, từ 40 bp đến 12 kbp, từ 40 bp đến 11 kbp, từ 50 bp đến 50 kbp, từ 50 bp đến 20 kbp, từ 50 bp đến 12 kbp, từ 50 bp đến 11 kbp, từ 60 bp đến 50 kbp, từ 60 bp đến 20 kbp, từ 60 bp đến 12 kbp, từ 60 bp đến 11 kbp, từ 70 bp đến 50 kbp, từ 70 bp đến 20 kbp, từ 70 bp đến 12 kbp, từ 70 bp đến 11 kbp, từ 80 bp đến 50 kbp, từ 80 bp đến 20 kbp, từ 80 bp đến 12 kbp, từ 80 bp đến 11 kbp, từ 90 bp đến 50 kbp, từ 90 bp đến 20 kbp, từ 90 bp đến 12 kbp, từ 90 bp đến 11 kbp, từ 100 bp đến 50 kbp, từ 100 bp đến 20 kbp, từ 100 bp đến 12 kbp, từ 100 bp đến 11 kbp, từ 200 bp đến 50 kbp, từ 200 bp đến 20 kbp, từ 200 bp đến 12 kbp, từ 200 bp đến 11 kbp, từ 500 bp đến 50 kbp, từ 500 bp đến 20 kbp, từ 500 bp đến 12 kbp, từ 500 bp đến 11 kbp, từ 1000 bp đến 50 kbp, từ 1000 bp đến 20 kbp, từ 1000 bp đến 12 kbp, từ 1000 bp đến 11 kbp, từ 1500 bp đến 50 kbp, từ 1500 bp đến 20 kbp, từ 1500 bp đến 12 kbp, từ 1500 bp đến 11 kbp, từ 2000 bp đến 50 kbp, từ 2000 bp đến 20 kbp, từ 2000 bp đến 12 kbp, hoặc từ 2000 bp đến 11 kbp.

(3) Phương pháp theo điểm (1) hoặc (2), trong đó tốt hơn là ADN sợi đơn bao gồm hai trình tự ADN tương ứng bao gồm các vùng tương đồng với vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt trong ADN hệ gen của vật chủ.

(4) Phương pháp theo điểm (3), trong đó tốt hơn là mỗi chiều dài của hai trình tự ADN là từ 5 bp trở lên, 10 bp trở lên, 15 bp trở lên, 20 bp trở lên, 25 bp trở lên, 30 bp trở lên, 35 bp trở lên, 40 bp trở lên, 45 bp trở lên, 50 bp trở lên, 100 bp trở lên, 250 bp trở lên, 500 bp trở lên, 750 bp trở lên, hoặc 1000 bp trở lên.

(5) Phương pháp theo điểm bất kỳ trong số các điểm từ (1) đến (4), trong đó tốt hơn là ADN sợi đơn bao gồm trình tự chèn được chèn vào ADN hệ gen của vật chủ.

(6) Phương pháp theo điểm (5), trong đó chiều dài của trình tự chèn là:

tốt hơn là từ 1 bp đến 50 kbp, từ 0,3 kbp đến 50 kbp, từ 1 kbp đến 50 kbp, hoặc từ 3 kbp đến 50 kbp,

tốt hơn nữa là từ 1 bp đến 20 kbp, từ 0,3 kbp đến 20 kbp, từ 1 kbp đến 20 kbp, hoặc từ 3 kbp đến 20 kbp,

tốt hơn nữa là từ 1 bp đến 10 kbp, từ 0,3 kbp đến 10 kbp, từ 1 kbp đến 10 kbp, hoặc từ 3 kbp đến 10 kbp,

còn tốt hơn nữa là từ 1 bp đến 9 kbp, từ 0,3 kbp đến 9 kbp, từ 1 kbp đến 9 kbp, hoặc từ 3 kbp đến 9 kbp,

còn tốt hơn nữa là từ 1 bp đến 5 kbp, từ 0,3 kbp đến 5 kbp, từ 1 kbp đến 5 kbp, hoặc từ 3 kbp đến 5 kbp.

(7) Phương pháp theo điểm bất kỳ trong số các điểm từ (1) đến (4), trong đó tốt hơn là ADN sợi đơn không chứa trình tự chèn được chèn vào ADN hệ gen của vật chủ.

(8) Phương pháp theo điểm bất kỳ trong số các điểm từ (1) đến (7), trong đó ADN nucleaza có thể lập trình là:

tốt hơn là ADN nucleaza nhân tạo dựa trên công nghệ TALEN, CRISPR-Cas9, CRISPR-Cpf1 hoặc ZFN,

tốt hơn nữa là TALEN.

(9) Phương pháp theo điểm bất kỳ trong số các điểm từ (1) đến (8), trong đó nấm sợi là:

tốt hơn là nấm thuộc ngành *Zygomycota*, ngành *Ascomycota*, hoặc ngành *Basidiomycota*,

tốt hơn nữa là nấm thuộc ngành *Zygomycota* hoặc ngành *Ascomycota*,

còn tốt hơn nữa là nấm thuộc ngành *Zygomycota*.

(10) Phương pháp theo (9), trong đó tốt hơn là nấm thuộc ngành *Zygomycota* là nấm thuộc chi *Rhizopus*.

(11) Phương pháp theo điểm (10), trong đó tốt hơn là nấm thuộc chi *Rhizopus* là *Rhizopus delemar* hoặc *Rhizopus oryzae*.

(12) Phương pháp theo điểm bất kỳ trong số các điểm từ (1) đến (11), trong đó phương pháp này tốt hơn là bao gồm thêm biểu hiện exonucleaza ngoại lai trong nấm sợi chủ kết hợp với chuyển ADN nucleaza có thể lập trình và ADN sợi đơn.

(13) Phương pháp theo điểm (12), trong đó exonucleaza ngoại lai là:

tốt hơn là exonucleaza có nguồn gốc từ nấm sợi,

tốt hơn nữa là exonucleaza có nguồn gốc từ nấm thuộc chi *Rhizopus*,
còn tốt hơn nữa là exonucleaza có nguồn gốc từ *Rhizopus oryzae* hoặc
Rhizopus delemar.

Ví dụ thực hiện sáng chế

Sau đây, sáng chế sẽ được mô tả chi tiết hơn có tham khảo đến các ví dụ.
Tuy nhiên, sáng chế không bị giới hạn bởi những ví dụ này.

Các đoạn môi PCR được sử dụng trong các ví dụ hiện tại được biểu thị
trong các Bảng từ 1 đến 6.

Bảng 1

Đoạn mồi	Trình tự (5'→3')	SEQ ID No.
oJK162	cgagctcgaattatttaaataaacagcaagttaataatctagaggg	10
oJK163	tatgaccatgattacgatgagaggcaaaatgaagcgtac	11
oJK164	atttaaataattcgagctcgggtacccgggg	12
oJK165	cgtaatcatgggtcatagctg	13
oJK202	tagagggaagagagaattgaaatagg	14
oJK204	ttttgtatttaattgtattaattgataatg	15
oJK205	aattaaataacaaaatcatttaattacgcatttc	16
oJK216	catgattacgcggccgcgccattataatgcactagt	17
oJK210	ctcttttccctctaataagaggcaaaatgaagcgtac	18
oJK211	atttaaataatcatgggtcatagctgttc	19
trpC-lost-F	tttaaattagagggaagagagaattgaaatag	20
trpC-lost-R	tccctctaatttaaataatcgagctcgggtacc	21
adhpro-R	ttttgtatttaattgtattaattgataatg	22
adhter-F	tcatttaattacgcattttcatttac	23
adhpro-TALEN-F	aattaaataacaaaatggactacaaagaccatgacgggtg	24
TALEN-adhter-R	gcgtaattaaaatgattaaaagtattctcgcggttatta	25
adhpro-LifeTALEN-F	aattaaataacaaaatgggaaaacctattcctaatacctctgctg	26
LifeTALEN-adhter-R	gcgtaattaaaatgatcagaagttgatctcgcggtgttgaaacttc	27
pPTR1-sal1-F	gggtaccgagctcgaattc	28
pPTR1-sal1-R	ggggatcctctagagtcgac	29
sal1-lhdhpro-F3	ctctagaggatcccctaggtgtggctgtggtgaccatattg	30
lhdhpro-R	gagaattatattgtaaagaaaaataaag	31
lhdhpro-exo1-F2	tacaatataattctcatgaaaatccaagttgcttctcctattgaccaatc	32

Bảng 2

Đoạn mồi	Trình tự (5'→3')	SEQ ID No.
exo1-pdcter-R2	atgaattctaagattttatcttctttcatgagaaacactaaacttgataac	33
pdcter-F	aatcttagaattcatcttttttg	34
pdcter-sal1-R	tcgagctcggtagccactctaccgtctgctctttgtct	35
adhpro-exo1-F	aattaaataacaaaaatgaaaatccaagttgcttctcctattgac	36
exo1-adhtr-R	gcgtaattaaaatgattatcttctttcatgagaaacactaaacttg	37
pUC18-Pae1-F3	ctgcaggctgactctagaggatccccgggtaccg	38
pUC18-Hind3-R3	gcttggcactggccgtcggtttacaacgtcgtgac	39
PDC1-upstr-F	cggccagtgccaaagcgcagacttcaacagttggctttttaagta	40
PDC1-upstr-R	cattttgcctctcatgttttaaatgtttgttagagtattgaata	41
trpCpro-R	acagcaagtaataatctagagggcgc	42
trpCter-F	atgagaggcaaaatgaagcgtacaaagag	43
PDC1-downstr-F	attaacttgctgttcaatcttagaattcattttttttgtatcattcg	44
PDC1-downstr-R	agagtcgacctgcaggcgtcaataagagcttgaagggttgccggatc	45
PDC2-upstr-F	cggccagtgccaaagcttatgggaaaaatgtgaaatcatcccact	46
PDC2-upstr-R	cattttgcctctcatgtttagttcaaataatattttttgtga	47
PDC2-downstr-F	attaacttgctgttcagttctctttctgtaataatccttgatttc	48
PDC2-downstr-R	agagtcgacctgcagccttttacagaaaacaagcacatctttac	49
trpC-inactive-F	taacgcatcagttccatcctgaaagtatcc	50
trpC-inactive-R	ggaactgatgcgttaaaaagaggggaaa	51
pdclter-F	aatcttagaattcattttttttgtatc	52
trpC-adh1pro-F	attaacttgctgttctagagggaagagagaattgaaatagg	53
FT1-pdclTer-R	atgaattctaagattttaatagaaacctgcttaaatgcaagacc	54
PYC-pdclTer-R	atgaattctaagatttaggcttctctttgacaaccttggccac	55

Bảng 3

Đoạn mồi	Trình tự (5'→3')	SEQ ID No.
PDC1-upstr-F2	gcagacttcaacagttggctttttaagta	56
PDC1-downstr-R2	gcgtcaataagagcttgaaggttggtgccgcatc	57
pdcl-up2	cattccacacaggatttgtgc	58
trpC(d)-1	gtgagatgtgatcatttgtacatg	59
pdcl-down1	agaagacaacctagaccctc	60
trpC(d)-7	atagctcttggtgggattcg	61
pdcl-down3	gctgctcgagatcctccaaggatc	62
trpC(d)-5	ttcgaccaagggttccatac	63
PDC1-downstr-R-P	gcgtcaataagagcttgaaggttggtgccgcatc	64
pdcl_750-F	tagattcaatttgattggataaagttcatc	65
pdcl_750-R-P	ataacaaacaatggctaaaagtggaaccccc	66
pdcl_500-F	tttactagttaaagcaaaaaacatgagca	67
pdcl_500-R-P	ttcaatttacatttctttatgaatatgcct	68
pdcl_250-F	gtgtttatctgttcacaagtactggtaagc	69
pdcl_250-R-P	ttcaatataaaccaagtcctaaaagaaat	70
pdcl_100-F	gctgatatgacattgcgacgaaaatagtat	71
pdcl_100-R-P	aaaaaaaaagcttatttcaaaaatatgat	72
pdcl_50-F	tcgttcaaaaaaatcattattcaatactc	73
pdcl_50-R-P	gtaatgaagtatagaacgaatgatacaaaa	74
pdcl_45-F	caaaaaaatcattattcaatactctacaa	75
pdcl_45-R-P	gaagtatagaacgaatgatacaaaaaaaaaaat	76

Bảng 4

Đoạn mồi	Trình tự (5'→3')	SEQ ID No.
pdcl_40-F	aaaatcattattcaataactctacaaaacaa	77
pdcl_40-R-P	atagaacgaatgatacaaaaaaaaaaatgaa	78
pdcl_35-F	cattattcaataactctacaaaacaaattta	79
pdcl_35-R-P	acgaatgatacaaaaaaaaaaatgaattcta	80
pdcl_30-F	ttcaataactctacaaaacaaatttaaaaac	81
pdcl_30-R-P	tgatacaaaaaaaaaaatgaattctaagatt	82
pdcl_25-F	tactctacaaaacaaatttaaaaacatgag	83
pdcl_25-R-P	caaaaaaaaaaatgaattctaagattgaaca	84
pdcl_20-F	tacaaaacaaatttaaaaacatgagaggca	85
pdcl_20-R-P	aaaaaatgaattctaagattgaacagcaag	86
pdcl_15-F	aacaaatttaaaaacatgagaggcaaaatg	87
pdcl_15-R-P	atgaattctaagattgaacagcaagttaat	88
pdcl_10-F	atttaaaaacatgagaggcaaaatgaagcg	89
pdcl_10-R-P	ttctaagattgaacagcaagttaataatct	90
pdcl_5-F	aaaacatgagaggcaaaatgaagcgtacaa	91
pdcl_5-R-P	agattgaacagcaagttaataatctagagg	92
PDC2-upstr-F2	ttatgggaaaaatgtgaaaatcatcccact	93
PDC2-downstr-R2-P	cctttttacagaaaacaagcacatctttac	94
trpC-ki-F2-P	cttttcatgaccaacaaatcagacac	95

Bảng 5

Đoạn môi	Trình tự (5'→3')	SEQ ID No.
trpCter-R	atgagaggcaaaatgaagcgtacaaagag	104
trpCter-cicCpro-F	cattttgcctctcatcttacgcaggtgatagtagccgcc	105
cipCpro-adhtr-R	gcgtaattaaaatgaggttagagtatgaagaaaaaaaaa	106
trpC-lost-F2	tttaaatcttacgcaggtgatagtagccgc	107
trpC-lost-R2	tgcgtaagatttaaatgaattcgagctcggtac	108
cipCpro-R	ggtagagtatgaagaaaaaaaaaaaaacg	109
cipCpro-LifeTALEN-F	cttcatactctaaccatgggaaaacctattcctaatacctctgctg	110
cipCpro-exo1-F	cttcatactctaaccatgaaaatccaagttgcttctccta	111
pdc3-upstr-F	cggccagtgccaaagcccgtcaggggtgaatgagatatttt	112
pdc3-upstr-R2	aaaagatgtgagttataaaaggatgatgcaagc	113
pdc3-downstr-F2	taactcacatcttttattcttttctatccctc	114
pdc3-downstr-R	agagtcgacctgcagacctgtagaaaggtagcatgcattc	115
pdc3-upstr-R	cattttgcctctcatgtgagttataaaaggatgatgcaag	116
pdc3-downstr-F	attaacttgctgttcacatcttttattcttttctatccctc	117
pdc3-upstr-F2	ccgtcaggggtgaatgagatatt	118
pdc3-downstr-R2-P	acctgtagaaaggtagcatgcattc	119
pdc3-up	gacctcaatcactatccttgg	120

Bảng 6

Đoạn mồi	Trình tự (5'→3')	SEQ ID No.
pUC18-adh1pro-F	cggccagtgccaaagcattattattagagggaagaaagaaaga	121
adh1pro-pUC18-R	agagtcgacctgcagttttgtattttgtattaattgataa	122
adh1pro-adh1ter-F	aacaaaatcattttaattacgcattttcatttttactaa	123
adh1ter-pUC18-R	agagtcgacctgcagagcagaacatgagctctggaagcgagacac	124
adh1pro-adh1ter-R	taaaatgattttgtattttgtattaattgataatga	125
adh1pro(o)-R	ttttgtattttgtattaattgataatg	126
adh1ter(o)-F	tcattttaattacgcattttcatttttac	127
adh1pro(o)-LifeTALEN-F	aaataaataacaaaatgggaaaacctattcctaatacctctgct	128
LifeTALEN-adh1ter(o)-R	gcgtaattaaaatgatacagaagtgatctcgccgttggtgaact	129
ldhA-upstr-F	cggccagtgccaaagcaaaagaataagaaaagatgtgtcag	130
ldhA-upstr-R2	taaattaaggacttgagctgaacgatgtcag	131
ldhA-downstr-F2	caagtccttaatttacaaataataaatcatgtt	132
ldhA-downstr-R	agagtcgacctgcagaacgacaaacatggctatcaaggga	133
ldhA-upstr-R	tgtgcaagtgaacaaaggacttgagctgaacgatgtcag	134
ldhA-downstr-F	gctcttgacaatgcataatttacaaataataaatcatgtt	135
ade1pro-R	tgcatgtcaagagcgtgtcgcaac	136
ade1ter-F	ttgttcacttgacagcgtgatatgcaag	137
ldhA-upstr-F2	aaaagaataagaaaagatgtgtcaggac	138
ldhA-downstr-R2-P	aacgacaaacatggctatcaagggaac	139
ldhA-up	gaaactacagtattccctcgtg	140
ade1-15	tgcttccatattgcaataggc	141
exo1-adh1ter-R2	gcgtaattaaaatgattatcttcttcatgagaaacacta	142

Ví dụ 1: Chuẩn bị thể đột biến của chi *Rhizopus*

(1) Chuẩn bị chủng khuyết dưỡng tryptophan

Chủng *Rhizopus delemar* JCM (Japan Collection of Microorganisms/Riken, Japan) 5557 (sau đây, được gọi là chủng 5557) hoặc chủng khuyết dưỡng tryptophan có nguồn gốc từ chủng 5557 được sử dụng làm chủng bố mẹ của thể gen tái tổ hợp tương đồng. Chủng khuyết dưỡng tryptophan thu được bằng cách sàng lọc giữa các chủng bị đột biến bởi chiếu xạ chùm ion của chủng 5557. Chiếu xạ chùm ion được thực hiện tại cơ sở của máy gia tốc ion Takasaki cho ứng dụng bức xạ tiên tiến (TIARA: Takasaki ion accelerators for advanced radiation application), Viện nghiên cứu bức xạ tiên tiến Takasaki, thuộc Viện Khoa học và Công nghệ Lượng tử và Phóng xạ Quốc gia. Chủng này được chiếu xạ từ 100 đến 1250 Gray với năng lượng 220 MeV với $^{12}\text{C}^{5+}$ được tăng tốc bằng cách sử dụng AVF xyclotron. Các bào tử được thu thập từ các tế bào nấm được chiếu xạ. Trong số đó, thu được chủng *Rhizopus delemar* 02T6 có đột biến xóa bỏ một bazơ trong vùng gen trpC và biểu hiện khuyết dưỡng tryptophan. Chủng 5557 hoặc chủng 02T6 được sử dụng làm chủng bố mẹ để chuẩn bị thể gen tái tổ hợp tương đồng của nấm thuộc chi *Rhizopus* trong các ví dụ tiếp theo.

(2) Chuẩn bị vectơ plasmit

Một đoạn ADN của vùng gen trpC được tổng hợp bằng PCR sử dụng ADN bộ gen của chủng 5557 làm khuôn với các đoạn mồi oJK162 (SEQ ID NO: 10) và oJK163 (SEQ ID NO: 11). Tiếp theo, đoạn ADN được khuếch đại bằng PCR sử dụng plasmit pUC18 làm khuôn với các đoạn mồi oJK164 (SEQ ID NO: 12) và oJK165 (SEQ ID NO: 13). Hai đoạn này được gắn bằng cách sử

dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-trpC.

Sau đó, đoạn trình tự khởi đầu và đoạn trình tự kết thúc của *adh1* được khuếch đại bằng PCR sử dụng ADN bộ gen của chủng 5557 làm khuôn với các đoạn mồi oJK202 (SEQ ID NO: 14) và oJK204 (SEQ ID NO: 15) và các đoạn mồi oJK205 (SEQ ID NO: 16) và oJK216 (SEQ ID NO: 17), tương ứng. Tiếp theo, đoạn ADN được khuếch đại bằng PCR sử dụng plasmit pUC18-trpC được xây dựng như được mô tả ở trên làm khuôn với các đoạn mồi oJK210 (SEQ ID NO: 18) và oJK211 (SEQ ID NO: 19). Ba đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-trpC-Padh-Tadh. Trong plasmit thu được, trình tự khởi đầu và trình tự kết thúc *adh1* được sắp xếp theo thứ tự xuôi chiều của vùng gen *trpC*.

Hơn nữa, vector plasmit trong đó vùng gen *trpC* được loại bỏ khỏi pUC18-trpC-Padh-Tadh đã được chuẩn bị. Cụ thể, đoạn ADN được khuếch đại bằng PCR sử dụng pUC18-trpC-Padh-Tadh được xây dựng như mô tả ở trên làm khuôn với các đoạn mồi là *trpC-lost-F* (SEQ ID NO: 20) và *trpC-lost-R* (SEQ ID NO: 21). Đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-Padh-Tadh.

(3) Chuẩn bị TALEN để phá vỡ gen *pdcl*, *pdcl2* và *trpC*

Các TALEN nhắm vào locus gen *pdcl*, *pdcl2* hoặc *trpC* đã được chuẩn bị. Việc chuẩn bị TALEN được gửi đến Transposagen Biopharmaceuticals, Inc. đối với các TALEN nhắm vào locus gen *pdcl*, và Life Technologies Corp. đối với các TALEN nhắm vào locus gen *pdcl2* hoặc *trpC* để thu được Custom XTN TALEN (tên thương mại của TALEN do Transposagen Biopharmaceuticals, Inc. cung cấp) hoặc GeneArt Precision TALs (tên thương mại của TALEN do Life

Technologies Corp. cung cấp). Mỗi kit này chứa hai polynucleotit mã hóa Left-TALEN và Right-TALEN cho gen đích. Kit nhắm vào gen *pdcl* (SEQ ID NO: 1) chứa LeftTALEN-*pdcl* (SEQ ID NO: 2) và RightTALEN-*pdcl* (SEQ ID NO: 3) mã hóa các TALEN nhắm vào trình tự 5'-TGCCTGCTATTAAAATCG-3' (SEQ ID NO: 96) trong mạch mang mã của gen *pdcl* và trình tự 5'-TTGATTTCCTTAAGACGG-3' (SEQ ID NO: 97) tương ứng trong mạch đối mã của chúng. Kit nhắm vào gen *pdcl* (SEQ ID NO: 4) chứa LeftTALEN-*pdcl* (SEQ ID NO: 5) và RightTALEN-*pdcl* (SEQ ID NO: 6) mã hóa các TALEN nhắm vào trình tự 5'-TGGTGTTGCTGGTTCTTAT-3' (SEQ ID NO: 98) trong mạch mang mã của gen *pdcl* và trình tự 5'-TGCCGACAATGTGAATCAC-3' (SEQ ID NO: 99) tương ứng trong mạch đối mã của chúng. Kit nhắm vào gen *trpC* (SEQ ID NO: 7) chứa LeftTALEN-*trpC* (SEQ ID NO: 8) và RightTALEN-*trpC* (SEQ ID NO: 9) mã hóa các TALEN nhắm vào trình tự 5'-TGCCAAGGCGCCAATGTAG-3' (SEQ ID NO: 100) trong mạch mang mã của gen *trpC* và trình tự 5'-TCGGAAATGGTGATTTTGT-3' (SEQ ID NO: 101) tương ứng trong mạch đối mã của chúng.

Polynucleotit mã hóa LeftTALEN cho gen *pdcl* đã được chèn vào vector biểu hiện pUC18-*trpC*-*Padh*-*Tadh* cho *R. delemar* đã được chuẩn bị trong đoạn (2) để chuẩn bị vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1*. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-*trpC*-*Padh*-*Tadh* làm khuôn với các đoạn mồi *adhpro*-R (SEQ ID NO: 22) và *adhtr*-F (SEQ ID NO: 23). Sau đó, đoạn LeftTALEN-*pdcl* được khuếch đại bằng PCR sử dụng LeftTALEN-*pdcl* làm khuôn với các đoạn mồi *adhpro*-TALEN-F (SEQ ID NO: 24) và TALEN-*adhtr*-R (SEQ ID NO: 25).

Hai đoạn này chứa các vùng chồng lấn nhau bởi 15 bazơ. Hai đoạn này được nối bằng cách sử dụng In-Fusion HD cloning kit (Clontech Laboratories, Inc.) để thu được plasmid padh-LeftTALEN-pdc1 chứa LeftTALEN-pdc1.

Tương tự, vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1* được chuẩn bị từ polynucleotit mã hóa LeftTALEN cho gen *pdc1* bằng cách sử dụng vector biểu hiện pUC18-Padh-Tadh cho *R. delemar* đã được chuẩn bị trong đoạn (2). Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-Padh-Tadh làm khuôn với các đoạn môi *adhpro*-R (SEQ ID NO: 22) và *adhter*-F (SEQ ID NO: 23). Sau đó, đoạn LeftTALEN-pdc1 được khuếch đại bằng PCR sử dụng LeftTALEN-pdc1 làm khuôn với các đoạn môi *adhpro*-TALEN-F (SEQ ID NO: 24) và TALEN-*adhter*-R (SEQ ID NO: 25), và được gắn với đoạn vector để thu được plasmid padh-LeftTALEN-pdc1-2 chứa LeftTALEN-pdc1.

Vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1* được chuẩn bị từ polynucleotit mã hóa RightTALEN cho gen *pdc1* bằng cách sử dụng vector biểu hiện pUC18-Padh-Tadh cho *R. delemar* đã được chuẩn bị trong đoạn (2). Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-Padh-Tadh làm khuôn với các đoạn môi *adhpro*-R (SEQ ID NO: 22) và *adhter*-F (SEQ ID NO: 23). Sau đó, đoạn RightTALEN-pdc1 được khuếch đại bằng PCR sử dụng RightTALEN-pdc1 làm khuôn với các đoạn môi *adhpro*-TALEN-F (SEQ ID NO: 24) và TALEN-*adhter*-R (SEQ ID NO: 25), và được gắn với đoạn vector để thu được plasmid padh-RightTALEN-pdc1 chứa RightTALEN-pdc1.

Các vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1* được chuẩn bị từ các polynucleotit mã hóa Left-TALEN và Right-TALEN cho gen *pdh2* bằng cách sử dụng vector biểu hiện pUC18-Padh-Tadh cho *R. delemar* được chuẩn bị trong đoạn (2). Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-Padh-Tadh làm khuôn với các đoạn mồi *adhpro*-R (SEQ ID NO: 22) và *adhtr*-F (SEQ ID NO: 23). Sau đó, đoạn LeftTALEN-*pdh2* hoặc RightTALEN-*pdh2* được khuếch đại bằng PCR sử dụng LeftTALEN-*pdh2* hoặc RightTALEN-*pdh2* làm khuôn với các đoạn mồi *adhpro*-LifeTALEN-F (SEQ ID NO: 26) và LifeTALEN-*adhtr*-R (SEQ ID NO: 27), và được gắn với đoạn vector để thu được plasmit *padh*-LeftTALEN-*pdh2* hoặc *padh*-RightTALEN-*pdh2* chứa LeftTALEN-*pdh2* hoặc RightTALEN-*pdh2*.

Tương tự, các vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1* được chuẩn bị từ các polynucleotit mã hóa Left-TALEN và Right-TALEN cho gen *trpC* bằng cách sử dụng vector biểu hiện pUC18-Padh-Tadh cho *R. delemar* được chuẩn bị trong đoạn (2). Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-Padh-Tadh làm khuôn với các đoạn mồi *adhpro*-R (SEQ ID NO: 22) và *adhtr*-F (SEQ ID NO: 23). Sau đó, đoạn LeftTALEN-*trpC* hoặc RightTALEN-*trpC* được khuếch đại bằng PCR sử dụng LeftTALEN-*trpC* hoặc RightTALEN-*trpC* làm khuôn với các đoạn mồi *adhpro*-LifeTALEN-F (SEQ ID NO: 26) và LifeTALEN-*adhtr*-R (SEQ ID NO: 27), và được gắn với đoạn vector để thu được plasmit *padh*-LeftTALEN-*trpC* hoặc *padh*-RightTALEN-*trpC* chứa LeftTALEN-*trpC* hoặc RightTALEN-*trpC*.

(4) Chuẩn bị vector biểu hiện exonucleaza

Đoạn vector pUC18 được khuếch đại bằng PCR sử dụng plasmit pUC18 (Takara Bio Inc.) làm khuôn với các đoạn mồi pPTR1-sal1-F (SEQ ID NO: 28) và pPTR1-sal1-R (SEQ ID NO: 29). Ngoài ra, đoạn trình tự khởi đầu ldh được khuếch đại bằng cách sử dụng dung dịch bộ gen tinh khiết của chủng *Rhizopus oryzae* NRBC5384 (sau đây, được gọi là chủng 5384) làm khuôn với các đoạn mồi sal1-ldhpro-F3 (SEQ ID NO: 30) và ldhpro-R (SEQ ID NO: 31). Đoạn gen exonucleaza được khuếch đại bằng cách sử dụng cùng một khuôn như trên với các đoạn mồi ldhpro-exo1-F2 (SEQ ID NO: 32) và exo1-pdcter-R2 (SEQ ID NO: 33). Đoạn trình tự kết thúc pdc được khuếch đại bằng cách sử dụng cùng một khuôn như trên với các đoạn mồi pdcTer-F (SEQ ID NO: 34) và pdcTer-sal1-R (SEQ ID NO: 35). Bốn đoạn được khuếch đại này được gán bằng cách sử dụng In-Fusion HD cloning kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pldh-exo1.

Ngoài ra, đoạn gen exonucleaza được khuếch đại bằng cách sử dụng dung dịch bộ gen tinh khiết của chủng 5384 làm khuôn với các đoạn mồi adhpro-exo1-F (SEQ ID NO: 36) và exo1-adhter-R (SEQ ID NO: 37). Đoạn vector được khuếch đại bằng PCR sử dụng pUC18-trpC-Padh-Tadh làm khuôn với các đoạn mồi adhpro-R (SEQ ID NO: 22) và adhter-F (SEQ ID NO: 23). Hai đoạn được khuếch đại này được gán bằng cách sử dụng In-Fusion HD cloning kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit padh-exo1.

(5) Chuẩn bị plasmit cho trpC knock-in nhắm vào từng locus gen

Một plasmit ptrpC-knock-in (pdc1) để loại bỏ ORF gen pdc1 và kích hoạt vùng gen trpC tại locus gen pdc1 đã được chuẩn bị. Cụ thể, đoạn vector pUC18 được khuếch đại bằng cách sử dụng pUC18 làm khuôn với các đoạn mồi pUC18-Pae1-F3 (SEQ ID NO: 38) và pUC18-Hind3-R3 (SEQ ID NO: 39), đoạn

vị trí trình tự khởi đầu của gen *pdcl* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi PDC1-upstr-F (SEQ ID NO: 40) và PDC1-upstr-R (SEQ ID NO: 41), đoạn vùng gen *trpC* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi *trpC*pro-R (SEQ ID NO: 42) và *trpC*ter-F (SEQ ID NO: 43), đoạn vị trí trình tự kết thúc của gen *pdcl* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi PDC1-downstr-F (SEQ ID NO: 44) và PDC1-downstr-R (SEQ ID NO: 45) được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để xây dựng *ptrpC*-knock-in (*pdcl*).

Tương tự, plasmit *ptrpC*-knock-in (*pdcl*) để loại bỏ ORF gen *pdcl* và kích hoạt vùng gen *trpC* tại locus gen *pdcl* đã được chuẩn bị. Cụ thể, đoạn vector pUC18 được khuếch đại bằng cách sử dụng pUC18 làm khuôn với các đoạn mồi pUC18-Pae1-F3 (SEQ ID NO: 38) và pUC18-Hind3-R3 (SEQ ID NO: 39), đoạn vị trí trình tự khởi đầu của gen *pdcl* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi PDC2-upstr-F (SEQ ID NO: 46) và PDC2-upstr-R (SEQ ID NO: 47), đoạn vùng gen *trpC* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi *trpC*pro-R (SEQ ID NO: 42) và *trpC*ter-F (SEQ ID NO: 43), đoạn vị trí trình tự kết thúc của gen *pdcl* được khuếch đại bằng cách sử dụng bộ gen của chủng 5557 làm khuôn với các đoạn mồi PDC2-downstr-F (SEQ ID NO: 48) và PDC2-downstr-R (SEQ ID NO: 49) được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để xây dựng *ptrpC*-knock-in (*pdcl*).

Sau đó, plasmit *ptrpC*-knock-in (*trpC*) để phá vỡ gen *trpC* tại locus gen *trpC* bằng cách kích hoạt vùng gen *trpC* (SEQ ID NO: 7) có sự xóa bỏ 500 bp từ codon mở đầu trong vùng gen *trpC* đã được chuẩn bị. Cụ thể, đoạn được khuếch

đại bằng cách sử dụng pUC18-trpC-Padh-Tadh làm khuôn với các đoạn mồi trpC-inactive-F (SEQ ID NO: 50) và trpC-inactive-R (SEQ ID NO: 51) và được gắn ở cả hai đầu sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để xây dựng ptrpC-knock-in (trpC) chứa trình tự gen trpC có sự xóa bỏ 500 bp so với codon mở đầu.

(6) Chuẩn bị plasmit cho (trpC + FT1) knock-in nhắm vào locus gen *pdcl*

Một plasmit p(trpC + FT1)-knock-in (*pdcl*) được chuẩn bị bằng cách chuyển trình tự khởi đầu *adh1* và gen FT1 vào đầu 5' của trình tự vùng gen trpC trong plasmit ptrpC-knock-in (*pdcl*). Cụ thể, đoạn vector được khuếch đại bằng cách sử dụng ptrpC-knock-in (*pdcl*) làm khuôn với các đoạn mồi trpCpro-R (SEQ ID NO: 42) và *pdcl*ter-F (SEQ ID NO: 52), đoạn chứa trình tự khởi đầu *adh1* và gen FT1 được khuếch đại bằng cách sử dụng plasmit FT1 (plasmit mang gen FT1 (SEQ ID NO: 102) xuôi chiều của trình tự khởi đầu *adh1* của PUC18-trpC-Padh-Tadh) làm khuôn với các đoạn mồi trpC-*adh1*pro-F (SEQ ID NO: 53) và FT1-*pdcl*Ter-R (SEQ ID NO: 54) được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để xây dựng p(trpC + FT1)-knock-in (*pdcl*).

(7) Chuẩn bị plasmit cho (trpC + PYC) knock-in nhắm vào locus gen *pdcl*

p(trpC + PYC)-knock-in (*pdcl*) được chuẩn bị theo cách tương tự như trong đoạn (6). Cụ thể, đoạn vector được khuếch đại bằng cách sử dụng ptrpC-knock-in (*pdcl*) làm khuôn với các đoạn mồi trpCpro-R (SEQ ID NO: 42) và *pdcl*ter-F (SEQ ID NO: 52), đoạn chứa trình tự khởi đầu *adh1* và gen PYC được khuếch đại bằng cách sử dụng plasmit PYC (plasmit mang gen PYC (SEQ ID NO: 103) xuôi chiều của trình tự khởi đầu *adh1* của PUC18-trpC-Padh-Tadh) làm khuôn với các đoạn mồi trpC-*adh1*pro-F (SEQ ID NO: 53) và PYC-

pdc1Ter-R (SEQ ID NO: 55) được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để xây dựng p(trpC + PYC)-knock-in (pdc1).

(8) Chuẩn bị ADN sợi đôi

ADN sợi đôi được chuẩn bị theo số 1 của Bảng 7. Cụ thể, đoạn ADN được khuếch đại bằng cách sử dụng plasmit ptrpC-knock-in (pdc1) làm khuôn với các đoạn mồi PDC1-upstr-F2 (SEQ ID NO: 56) và PDC1-downstr-R2 (SEQ ID NO: 57). Khuôn bị phân hủy bằng cách xử lý Dpn1 (Toyobo Co., Ltd.). Sau đó, sản phẩm được tinh chế bằng phương pháp xử lý phenol/cloroform/rượu isoamyl và xử lý kết tủa bằng etanol. Sản phẩm tinh khiết được hòa tan trong một lượng nước không có DNaza thích hợp để thu được dung dịch ADN sợi đôi cho sự phá vỡ gen pdc1.

(9) Chuẩn bị ADN sợi đơn

Các ADN sợi đơn được chuẩn bị theo số 2 đến 20 của Bảng 7. Cụ thể, đoạn ADN có chiều dài bằng "chiều dài chèn" của Bảng 7 được khuếch đại bằng cách sử dụng "khuôn" của từng điều kiện trong Bảng 7 làm khuôn với "Đoạn mồi 1" và "Đoạn mồi 2". Khuôn đã bị phân hủy do xử lý Dpn1 (Toyobo Co., Ltd.). Sau đó, sản phẩm được tinh chế bằng phương pháp xử lý phenol/cloroform/rượu isoamyl và xử lý kết tủa bằng etanol. Sản phẩm tinh khiết được xử lý thêm bằng cách sử dụng Lambda Exonucleaza (NEW ENGLAND BioLabs Inc.) và sau đó được tinh chế theo cách tương tự như trên để thu được ADN sợi đơn. Quá trình xử lý Lambda Exonucleaza được thực hiện qua đêm ở 37°C.

Bảng 7

Số	ADN	Đoạn môi 1	Đoạn môi 2	Khuôn	Locus gen	Chiều dài chèn (bp)	Chiều dài trình tự tương đồng (bp)
1	Sợi đôi	PDC1-upstr-F2	PDC1-downstr-R2	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	1000
2	Sợi đơn	PDC1-upstr-F2	PDC1-downstr-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	1000
3	Sợi đơn	PDC1-upstr-F2	PDC1-downstr-R-P	p(trpC+FT1)-knock-in (p _{dc1})	p _{dc1}	5811	1000
4	Sợi đơn	PDC1-upstr-F2	PDC1-downstr-R-P	p(trpC+PYC)-knock-in (p _{dc1})	p _{dc1}	8838	1000
5	Sợi đơn	p _{dc1} _750-F	p _{dc1} _750-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	750
6	Sợi đơn	p _{dc1} _500-F	p _{dc1} _500-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	500
7	Sợi đơn	p _{dc1} _250-F	p _{dc1} _250-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	250
8	Sợi đơn	p _{dc1} _100-F	p _{dc1} _100-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	100
9	Sợi đơn	p _{dc1} _50-F	p _{dc1} _50-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	50
10	Sợi đơn	p _{dc1} _45-F	p _{dc1} _45-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	45
11	Sợi đơn	p _{dc1} _40-F	p _{dc1} _40-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	40
12	Sợi đơn	p _{dc1} _35-F	p _{dc1} _35-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	35
13	Sợi đơn	p _{dc1} _30-F	p _{dc1} _30-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	30
14	Sợi đơn	p _{dc1} _25-F	p _{dc1} _25-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	25
15	Sợi đơn	p _{dc1} _20-F	p _{dc1} _20-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	20
16	Sợi đơn	p _{dc1} _15-F	p _{dc1} _15-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	15
17	Sợi đơn	p _{dc1} _10-F	p _{dc1} _10-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	10
18	Sợi đơn	p _{dc1} _5-F	p _{dc1} _5-R-P	p _{trpC} -knock-in (p _{dc1})	p _{dc1}	4298	5
19	Sợi đơn	PDC2-upstr-F2	PDC2-downstr-r2-P	p _{trpC} -knock-in (p _{dc1})	p _{dc2}	4298	1000
20	Sợi đơn	trpC-ki-F2-P	trpCpro-R	p _{trpC} -knock-in (trpC)	trpC	3798	1000

* Đoạn mỗi 2 của số 2 đến 19 đã được phosphoryl hóa tận cùng 5'. Đoạn mỗi 1 của số 20 đã được phosphoryl hóa đầu 5'.

"Chiều dài trình tự tương đồng" thể hiện rằng các trình tự tương đồng có chiều dài được hiển thị trong bảng có mặt ở cả đầu ngược chiều và đầu xuôi chiều của ADN sợi đơn hoặc ADN sợi đôi.

(10) Chuyển gen bằng bắn gen

Mỗi ADN sợi đơn hoặc ADN sợi đôi được chuẩn bị trong đoạn (8) hoặc (9) (số 1 đến 20 trong Bảng 7), mỗi vector biểu hiện Left-TALEN và mỗi vector biểu hiện Right-TALEN đã được chuẩn bị trong đoạn (3), và mỗi vector biểu hiện exonucleaza đã được chuẩn bị trong đoạn (4) được trộn để chuẩn bị dung dịch ADN.

Các vector biểu hiện TALEN được sử dụng là padh-LeftTALEN-pdc1 và padh-RightTALEN-pdc1 cho số 1 và 2, padh-LeftTALEN-pdc1-2 và padh-RightTALEN-pdc1 cho số 3 đến 18, padh-LeftTALEN-pdc2 và padh-RightTALEN-pdc2 cho số 19, và padh-LeftTALEN-trpC và padh-RightTALEN-trpC cho số 20. Vector biểu hiện exonucleaza được sử dụng là pldh-exo1 cho số 1 và 2, và padh-exo1 cho những loại khác. Các plasmit chuyển gen chỉ được xử lý bằng enzym giới hạn đối với các điều kiện của số 1 và 2. Cụ thể, padh-LeftTALEN-pdc1, padh-LeftTALEN-pdc1-2 và padh-RightTALEN-pdc1 được xử lý bằng enzym giới hạn ScaI, và plasmit pldh-exo1 được xử lý bằng enzym giới hạn PstI. Tỷ lệ nồng độ giữa vector biểu hiện Left-TALEN, vector biểu hiện Right-TALEN, vector biểu hiện exonucleaza, và ADN sợi đơn hoặc ADN sợi đôi trong dung dịch ADN được đặt thành khoảng 2:4:2:1 cho số 1 và 2 và khoảng 1:1:1:2 cho số 3 đến 20.

Bổ sung 10 μL mỗi dung dịch ADN đã chuẩn bị (khoảng 1 đến 3 $\mu\text{g}/\mu\text{L}$) và trộn với 100 μL dung dịch chứa hạt vàng (60 mg/mL, INBIO GOLD, kích thước hạt: 1 μm). Tiếp theo, bổ sung 40 μL spermidin 0,1 M vào đó, và hỗn hợp được khuấy kỹ bằng cách trộn đều. Bổ sung 100 μL CaCl_2 2,5 M vào đó, và hỗn hợp được khuấy trong 1 phút bằng cách trộn đều và sau đó ly tâm ở tốc độ 6000 vòng/phút trong 30 giây để loại bỏ dịch nổi phía trên. Bổ sung 200 μL EtOH 70% vào chất kết tủa thu được, và hỗn hợp được khuấy trong 30 giây bằng cách trộn đều và sau đó ly tâm ở tốc độ 6000 vòng/phút trong 30 giây để loại bỏ dịch nổi phía trên. Các chất kết tủa thu được này được ngâm lại trong 100 μL EtOH 100%.

Các bào tử của chủng 02T6 hoặc chủng 5557 thu được trong đoạn (1) được chuyển gen bằng cách sử dụng từng dung dịch chứa hạt vàng ADN được mô tả ở trên và GDS-80 (súng bắn gen từ Nepa Gene Co., Ltd.). Chủng 02T6 được sử dụng cho các điều kiện từ số 1 đến 19, và các bào tử sau khi chuyển gen được nuôi cấy tĩnh ở 30°C trong khoảng 1 tuần trên môi trường thạch vô cơ (20 g/L glucoza, 1 g/L amoni sulfat, 0,6 g/L kali dihydro phosphat, 0,25 g/L magie sulfat heptahydrat, 0,09 g/L kẽm sulfat heptahydrat, và 15 g/L thạch). Các bào tử của chủng 5557 được sử dụng cho điều kiện số 20, và các bào tử sau khi chuyển gen được nuôi cấy tĩnh ở 30°C trong khoảng 1 tuần trên môi trường PDB (39 g/L PDB).

(11) Lựa chọn chủng khuyết gen *pdc1* (số 1 đến số 18)

Các bào tử được thu thập từ các tế bào nấm được nuôi cấy trong đoạn (10). Các chủng nấm được phân lập bằng cách sử dụng môi trường thạch vô cơ (20

g/L glucoza, 1 g/L amoni sulfat, 0,6 g/L kali dihydro phosphat, 0,25 g/L magie sulfat heptahydrat, 0,09 g/L kẽm sulfat heptahydrat, và 15 g/L thạch) được điều chỉnh đến độ pH 3. Một phần sợi nấm của các chủng nấm phát triển được thu thập bằng tăm, sau đó được tạo huyền phù trong 10 mM Tris-HCl (độ pH 8,5), và ủ ở 95°C trong 10 phút. Sau đó, huyền phù được pha loãng thích hợp với 10 mM Tris-HCl (độ pH 8,5) để chuẩn bị dung dịch khuôn hệ gen cho phương pháp PCR khuôn lạc. Phương pháp PCR khuôn lạc được thực hiện bằng cách sử dụng dung dịch khuôn hệ gen, các đoạn mồi pdc1-up2 (SEQ ID NO: 58) và trpC(d)-1 (SEQ ID NO: 59), và KOD FX Neo (Toyobo Co., Ltd.). Phương pháp PCR khuôn lạc sử dụng các đoạn mồi này khuếch đại đoạn ADN có chiều dài thích hợp nếu đoạn gen trpC hoặc đoạn gen khác với gen trpC được kích hoạt tại locus gen pdc1. Bằng phương pháp PCR khuôn lạc, thu được chủng nấm có đoạn khuếch đại ADN thu được là chủng knock-in (chủng khuyết gen pdc1).

(12) Lựa chọn chủng khuyết gen pdc2 (số 19)

Các bào tử được thu thập từ các tế bào nấm được nuôi cấy trong đoạn (10). Các chủng nấm được phân lập bằng cách sử dụng môi trường thạch vô cơ (20 g/L glucoza, 1 g/L amoni sulfat, 0,6 g/L kali dihydro phosphat, 0,25 g/L magie sulfat heptahydrat, 0,09 g/L kẽm sulfat heptahydrat, và 15 g/L thạch) được điều chỉnh đến độ pH 3. Một phần sợi nấm của các chủng nấm phát triển được thu thập bằng tăm, sau đó được tạo huyền phù trong 10 mM Tris-HCl (độ pH 8,5), và ủ ở 95°C trong 10 phút. Sau đó, huyền phù được pha loãng thích hợp với 10 mM Tris-HCl (độ pH 8,5) để chuẩn bị dung dịch khuôn hệ gen cho phương pháp PCR khuôn lạc. Phương pháp PCR khuôn lạc được thực hiện sử dụng dung dịch khuôn hệ gen, các đoạn mồi pdc2-down1 (SEQ ID NO: 60) và trpC(d)-7 (SEQ ID NO: 61), và KOD FX Neo (Toyobo Co., Ltd.). Phương pháp PCR khuôn lạc

sử dụng các đoạn mồi này khuếch đại đoạn ADN có chiều dài thích hợp nếu đoạn gen *trpC* được kích hoạt tại locus gen *pdC2*. Bằng phương pháp PCR khuôn lạc, thu được chủng nấm có đoạn khuếch đại ADN thu được là chủng knock-in (chủng khuyết gen *pdC2*).

(13) Lựa chọn chủng khuyết gen *trpC* (số 20)

Các bào tử được thu thập từ các tế bào nấm được nuôi cấy trong đoạn (10), và nuôi cấy trong môi trường 5-FAA (10 g/L glucoza, 3,36 g/L bazơ nitơ nấm men không có axit amin, 0,03 g/L tryptophan, 5 g/L axit 5-fluoroanthranilic, và 15 g/L thạch) được điều chỉnh đến độ pH 6. Các chủng nấm phát triển được nuôi cấy cấp hai trong môi trường 5-FAA. Một phần sợi nấm của các chủng nấm phát triển được thu thập bằng tăm, sau đó được tạo huyền phù trong 10 mM Tris-HCl (độ pH 8,5), và được ủ ở 95°C trong 10 phút. Sau đó, huyền phù được pha loãng thích hợp với 10 mM Tris-HCl (độ pH 8,5) để chuẩn bị dung dịch khuôn hệ gen cho phương pháp PCR khuôn lạc. Phương pháp PCR khuôn lạc được thực hiện sử dụng dung dịch khuôn hệ gen, các đoạn mồi *pdC1-down3* (SEQ ID NO: 62) và *trpC(d)-5* (SEQ ID NO: 63), và KOD FX Neo (Toyobo Co., Ltd.). Vị trí dài của đoạn khuếch đại ADN được thay đổi bởi phương pháp PCR khuôn lạc sử dụng các đoạn mồi này nếu đoạn gen *trpC* có sự xóa bỏ 500 bp được kích hoạt tại locus gen *trpC*. Bằng phương pháp PCR khuôn lạc, thu được chủng nấm có dài của đoạn khuếch đại ADN bị thay đổi là chủng knock-in (chủng khuyết gen *trpC*).

(14) So sánh hiệu quả tái tổ hợp tương đồng

(a) ADN sợi đơn và ADN sợi đôi

Hiệu quả tái tổ hợp tương đồng được so sánh giữa ADN sợi đôi và ADN sợi đơn (số 1 và 2 của Bảng 7) được sử dụng làm ADN người cho. Các kết quả

được thể hiện trong Bảng 8. Không thể thu được chủng tái tổ hợp tương đồng có trình tự chuỗi dài được chuyển bằng cách sử dụng ADN sợi đôi, trong khi chủng tái tổ hợp tương đồng có trình tự chuỗi dài được chuyển đã thu được thành công bằng cách sử dụng ADN sợi đơn. Hơn nữa, chủng tái tổ hợp tương đồng có trình tự chuỗi dài được chuyển dài nhất là khoảng 8,8 kbp đã thu được thành công bằng cách sử dụng ADN sợi đơn làm ADN người cho (số 2 đến 4 của Bảng 7).

Bảng 8

Điều kiện thử nghiệm số	Số lượng chủng phân lập	Số lượng các chủng tái tổ hợp tương đồng	Hiệu quả tái tổ hợp tương đồng (%)
1	108	0	0
2	8	1	12,5

(b) Chiều dài của vùng tương đồng với ADN hệ gen vật chủ

Ảnh hưởng của chiều dài chuỗi của các vùng tương đồng với ADN hệ gen vật chủ trong ADN sợi đơn đến hiệu quả tái tổ hợp tương đồng đã được nghiên cứu (số 2, 8, và 9 của Bảng 7). Các kết quả được thể hiện trong Bảng 9.

Bảng 9

Điều kiện thử nghiệm số	Chiều dài trình tự tương đồng (bp)	Số lượng các chủng phân lập	Số lượng các chủng tái tổ hợp tương đồng	Hiệu quả tái tổ hợp tương đồng (%)
2	1000	8	1	12,5
8	100	24	22	92
9	50	24	19	79

(15) Thu được chủng khuyết gen *pdC3*

(a) Chuẩn bị vector plasmit

Vector plasmit được chuẩn bị bằng cách thay đổi trình tự khởi đầu adh1 của pUC18-trpC-Padh-Tadh được xây dựng trong đoạn (2) thành trình tự khởi đầu cipC. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-trpC-Padh-Tadh làm khuôn với các đoạn mồi adhter-F (SEQ ID NO: 23) và trpCter-R (SEQ ID NO: 104). Ngoài ra, đoạn vùng trình tự khởi đầu cipC được khuếch đại bằng PCR sử dụng ADN hệ gen của chủng 5557 làm khuôn với các đoạn mồi trpCter-cipCpro-F (SEQ ID NO: 105) và cipCpro-adhter-R (SEQ ID NO: 106). Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-trpC-PcipC-Tadh.

Sau đó, vector plasmit trong đó vùng gen trpC đã được loại bỏ khỏi pUC18-trpC-PcipC-Tadh được xây dựng như mô tả ở trên đã được chuẩn bị. Cụ thể, đoạn ADN được khuếch đại bằng PCR sử dụng pUC18-trpC-PcipC-Tadh làm khuôn với các đoạn mồi trpC-lost-F2 (SEQ ID NO: 107) và trpC-lost-R2 (SEQ ID NO: 108). Đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-PcipC-Tadh.

(b) Chuẩn bị TALEN để phá vỡ gen pdc3

Các TALEN nhắm vào locus gen pdc3 đã được chuẩn bị. Việc chuẩn bị TALEN đã được gửi đến Life Technologies Corp. để thu được GeneArt PerfectMatch TALs (tên thương mại của TALEN do Life Technologies Corp. cung cấp). Kit này chứa hai polynucleotit mã hóa Left-TALEN và Right-TALEN cho gen đích. Kit nhắm vào gen pdc3 (SEQ ID NO: 143) chứa LeftTALEN-pdc3 (SEQ ID NO: 144) và RightTALEN-pdc3 (SEQ ID NO: 145)

mã hóa các TALEN nhắm vào trình tự 5'-CCGGAATCGACACGATTTT-3' (SEQ ID NO: 146) tương ứng trong mạch mang mã của gen *pdC3* và trình tự 5'-CGTAACTTACCATATTGTA-3' (SEQ ID NO: 147) trong mạch đối mã của chúng.

Polynucleotit mã hóa Left-TALEN cho gen *pdC3* đã được chèn vào vector biểu hiện pUC18-PcipC-Tadh cho *R. delemar* đã được chuẩn bị trong đoạn (15)(a) để chuẩn bị vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *cipC* và trình tự kết thúc *adh1*. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-PcipC-Tadh làm khuôn với các đoạn mồi *cipC*pro-R (SEQ ID NO: 109) và *adh*ter-F (SEQ ID NO: 23). Sau đó, đoạn LeftTALEN-*pdC3* được khuếch đại bằng PCR sử dụng LeftTALEN-*pdC3* làm khuôn với các đoạn mồi *cipC*pro-LifeTALEN-F (SEQ ID NO: 110) và LifeTALEN-*adh*ter-R (SEQ ID NO: 27). Hai đoạn này chứa các vùng chồng lấn nhau bởi 15 bazơ. Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để thu được plasmit *pcipC*-LeftTALEN-*pdC3* chứa LeftTALEN-*pdC3*.

Tương tự như vậy, polynucleotit mã hóa RightTALEN cho gen *pdC3* được chèn vào vector biểu hiện pUC18-PcipC-Tadh cho *R. delemar* đã chuẩn bị trong đoạn (15)(a) để chuẩn bị vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *cipC* và trình tự kết thúc *adh1*. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-PcipC-Tadh làm khuôn với các đoạn mồi *cipC*pro-R (SEQ ID NO: 109) và *adh*ter-F (SEQ ID NO: 23). Sau đó, đoạn RightTALEN-*pdC3* được khuếch đại bằng PCR sử dụng RightTALEN-*pdC3* làm khuôn với các

đoạn môi cipCpro-LifeTALEN-F (SEQ ID NO: 110) và LifeTALEN-adhter-R (SEQ ID NO: 27). Hai đoạn này chứa các vùng chồng lẫn nhau bởi 15 bazơ. Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để thu được pcipC-RightTALEN-pdc3 plasmit chứa RightTALEN-pdc3.

(c) Chuẩn bị vectơ biểu hiện exonucleaza

Đoạn gen exonucleaza được khuếch đại bằng PCR sử dụng plasmit pldh-exo1 được chuẩn bị trong đoạn (4) làm khuôn với các đoạn môi cipCpro-exo1-F (SEQ ID NO: 111) và exo1-adhter-R2 (SEQ ID NO: 142). Đoạn vector được khuếch đại bằng PCR sử dụng plasmit pUC18-PcipC-Tadh được chuẩn bị trong đoạn (15)(a) làm khuôn với các đoạn môi cipCpro-R (SEQ ID NO: 109) và adhter-F (SEQ ID NO: 23). Hai đoạn này chứa các vùng chồng lẫn nhau bởi 15 bazơ. Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để thu được plasmit pcipC-exo1.

(d) Chuẩn bị plasmit cho trpC knock-in nhắm vào locus gen pdc3

Plasmit ptrpC-knock-in (pdc3) để loại bỏ ORF gen pdc3 và kích hoạt vùng gen trpC tại locus gen pdc3 đã được chuẩn bị. Cụ thể, đoạn vector pUC18 được khuếch đại bằng cách sử dụng pUC18 làm khuôn với các đoạn môi pUC18-Pae1-F3 (SEQ ID NO: 38) và pUC18-Hind3-R3 (SEQ ID NO: 39). Đoạn trình tự khởi đầu gen pdc3 được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5557 làm khuôn với các đoạn môi pdc3-upstr-F (SEQ ID NO: 112) và pdc3-upstr-R2 (SEQ ID NO: 113). Đoạn trình tự kết thúc gen pdc3 được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5557 làm khuôn với các đoạn

mỗi pdc3-downstr-F2 (SEQ ID NO: 114) và pdc3-downstr-R (SEQ ID NO: 115). Ba đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị pknock-in (pdc3).

Sau đó, đoạn ADN được khuếch đại bằng cách sử dụng pknock-in (pdc3) làm khuôn với các đoạn mỗi pdc3-upstr-R (SEQ ID NO: 116) và pdc3-downstr-F (SEQ ID NO: 117). Đoạn vùng gen trpC được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5557 làm khuôn với các đoạn mỗi trpCpro-R (SEQ ID NO: 42) và trpCter-F (SEQ ID NO: 43). Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị ptrpC-knock-in (pdc3).

(e) Chuẩn bị ADN sợi đơn

ADN sợi đơn thu được bằng các quy trình tương tự như trong đoạn (9) ngoại trừ: plasmit ptrpC-knock-in (pdc3) được sử dụng làm khuôn PCR; và pdc3-upstr-F2 (SEQ ID NO: 118) và pdc3-downstr-R2-P (SEQ ID NO: 119; phosphoryl hóa tận cùng 5') được sử dụng làm các đoạn mỗi PCR.

(f) Chuyển gen bằng bắn gen

Bằng các quy trình tương tự như trong đoạn (10), dung dịch chứa hạt vàng ADN chứa ADN sợi đơn đã được chuẩn bị, và sự chuyển gen đến các bào tử của chủng 02T6 được thực hiện sử dụng dung dịch này, sau đó là nuôi cấy các bào tử thu được. ADN sợi đơn được sử dụng là ADN sợi đơn được chuẩn bị trong đoạn (15)(e). Các vector biểu hiện TALEN được sử dụng là pcipC-LeftTALEN-pdc3 và pcipC-RightTALEN-pdc3 được chuẩn bị trong đoạn (15)(b). Vector biểu hiện exonucleaza được sử dụng là pcipC-exo1 được chuẩn bị trong đoạn (15)(c).

Tỷ lệ nồng độ giữa pcipC-LeftTALEN-pdc3, pcipC-RightTALEN-pdc3, pcipC-exo1, và ADN sợi đơn trong dung dịch ADN được đặt thành khoảng 1:1:1:2.

(g) Lựa chọn chủng khuyết gen pdc3

Bằng các quy trình tương tự như trong đoạn (11), các bào tử được thu thập từ các tế bào nấm được nuôi cấy trong đoạn (15)(f), và việc phân lập các chủng nấm và chuẩn bị dung dịch khuôn hệ gen được thực hiện. Sau đó, chủng khuyết gen pdc3 với đoạn vùng gen trpC được kích hoạt tại locus gen pdc3 đã được chọn bằng phương pháp PCR khuôn lạc sử dụng dung dịch khuôn hệ gen làm khuôn. Phương pháp PCR khuôn lạc được thực hiện sử dụng dung dịch khuôn hệ gen, các đoạn mồi pdc3-up (SEQ ID NO: 120) và trpC(d)-1 (SEQ ID NO: 59), và KOD FX Neo (Toyobo Co., Ltd.). Phương pháp PCR khuôn lạc sử dụng các đoạn mồi này khuếch đại đoạn ADN có chiều dài thích hợp nếu đoạn vùng gen trpC được kích hoạt tại locus gen pdc3. Bằng phương pháp PCR khuôn lạc, thu được chủng nấm có mảnh khuếch đại ADN thu được là chủng knock-in (chủng khuyết gen pdc3).

(h) Hiệu quả tái tổ hợp tương đồng tại locus gen pdc3

Chủng khuyết gen pdc3 thu được này được kiểm tra hiệu quả tái tổ hợp tương đồng (Bảng 10).

Bảng 10

Chiều dài trình tự tương đồng ngược chiều (bp)	Chiều dài trình tự tương đồng xuôi chiều (bp)	Số lượng các chủng phân lập	Số lượng các chủng tái tổ hợp tương đồng	Hiệu quả tái tổ hợp tương đồng (%)
1000	950	20	8	40

Ví dụ 2: Chuẩn bị thể đột biến của chi *Rhizopus* (*Rhizopus oryzae*)

(1) Chuẩn bị chủng khuyết dưỡng adenin

Chủng *Rhizopus oryzae* NRBC5384 (sau đây, được gọi là chủng 5384) được sử dụng làm chủng bố mẹ của thể gen tái tổ hợp tương đồng. Chủng khuyết dưỡng adenin thu được bằng cách sàng lọc trong số các chủng bị đột biến bởi sự chiếu xạ tia tử ngoại các bào tử của chủng 5384. Quá trình chiếu xạ tia tử ngoại được thực hiện bằng cách chiếu xạ trong 1 phút bằng đèn diệt khuẩn GL15 (Hitachi Appliances, Inc.). Các bào tử được chiếu xạ đã được thu thập. Trong số đó, đã thu được chủng *Rhizopus oryzae* AM002 có đột biến trong vùng gen *ade1* và biểu hiện khuyết dưỡng adenin. Chủng *Rhizopus oryzae* AM002 được sử dụng làm chủng bố mẹ để chuẩn bị thể gen tái tổ hợp tương đồng của *Rhizopus oryzae* trong các ví dụ tiếp theo.

(2) Chuẩn bị vector plasmit cho *Rhizopus oryzae*

Đoạn trình tự khởi đầu của *adh1* được khuếch đại bằng PCR sử dụng ADN hệ gen của chủng 5384 làm khuôn với các đoạn mồi pUC18-*adh1*pro-F (SEQ ID NO: 121) và *adh1*pro-pUC18-R (SEQ ID NO: 122). Tiếp theo, đoạn ADN được khuếch đại bằng PCR sử dụng pUC18 làm khuôn với các đoạn mồi pUC18-Pae1-F3 (SEQ ID NO: 38) và pUC18-Hind3-R3 (SEQ ID NO: 39). Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-Padh (RO).

Sau đó, đoạn trình tự kết thúc của *adh1* được khuếch đại bằng PCR sử dụng ADN hệ gen của chủng 5384 làm khuôn với các đoạn mồi *adh1*pro-

adh1ter-F (SEQ ID NO: 123) và adh1ter-pUC18-R (SEQ ID NO: 124). Hơn nữa, đoạn ADN được khuếch đại bằng PCR sử dụng pUC18-*Padh* (RO) làm khuôn với các đoạn mồi pUC18-*Pae1*-F3 (SEQ ID NO: 38) và adh1pro-adh1ter-R (SEQ ID NO: 125). Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị plasmit pUC18-*Padh*-*Tadh* (RO).

(3) Chuẩn bị TALEN để phá vỡ gen *ldhA*.

Các TALEN nhắm vào locus gen *ldhA* đã được chuẩn bị. Việc chuẩn bị TALEN đã được gửi đến Life Technologies Corp. để thu được GeneArt Precision TALs (tên thương mại của TALEN do Life Technologies Corp. cung cấp). Kit này chứa hai polynucleotit mã hóa Left-TALEN và Right-TALEN cho gen đích. Kit nhắm vào gen *ldhA* (SEQ ID NO: 148) chứa LeftTALEN-*ldhA* (SEQ ID NO: 149) và RightTALEN-*ldhA* (SEQ ID NO: 150) mã hóa các TALEN nhắm vào trình tự 5'-TGCAGATGCTGCCAGTATA-3' (SEQ ID NO: 151) trong mạch mang mã của gen *ldhA* và trình tự 5'-TCCTCTGCGCTACCTGCTC-3' (SEQ ID NO: 152) trong mạch đối mã của nó, tương ứng.

Polynucleotit mã hóa Left-TALEN cho gen *ldhA* được chèn vào vector biểu hiện pUC18-*Padh*-*Tadh* (RO) cho *R. oryzae* đã được chuẩn bị trong đoạn (2) để chuẩn bị vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu *adh1* và trình tự kết thúc *adh1*. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-*Padh*-*Tadh* (RO) làm khuôn với các đoạn mồi adh1pro(o)-R (SEQ ID NO: 126) và adh1ter(o)-F (SEQ ID NO: 127). Sau đó, đoạn

LeftTALEN-ldhA được khuếch đại bằng PCR sử dụng LeftTALEN-ldhA làm khuôn với các đoạn mồi adh1pro(o)-LifeTALEN-F (SEQ ID NO: 128) và LifeTALEN-adh1ter(o)-R (SEQ ID NO: 129). Hai đoạn này chứa các vùng chồng lấn nhau bởi 15 bazơ. Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để thu được plasmid padh-LeftTALEN-ldhA (RO) chứa LeftTALEN-ldhA.

Tương tự, polynucleotit mã hóa Right-TALEN cho gen ldhA đã được chèn vào vector biểu hiện pUC18-Padh-Tadh (RO) để chuẩn bị vector biểu hiện TALEN dưới sự kiểm soát của trình tự khởi đầu adh1 và trình tự kết thúc adh1. Cụ thể, đoạn vector được khuếch đại bằng PCR sử dụng pUC18-Padh-Tadh (RO) làm khuôn với các đoạn mồi adh1pro(o)-R (SEQ ID NO: 126) và adh1ter(o)-F (SEQ ID NO: 127). Sau đó, đoạn RightTALEN-ldhA được khuếch đại bằng PCR sử dụng RightTALEN-ldhA làm khuôn với các đoạn mồi adh1pro(o)-LifeTALEN-F (SEQ ID NO: 128) và LifeTALEN-adh1ter(o)-R (SEQ ID NO: 129). Hai đoạn này chứa các vùng chồng lấn nhau bởi 15 bazơ. Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để thu được plasmid padh-RightTALEN-ldhA (RO) chứa RightTALEN-ldhA.

(4) Chuẩn bị plasmid cho ade1 knock-in nhắm vào locus gen ldhA

Plasmid pade1-knock-in (ldhA) để loại bỏ một phần gen ldhA ORF và kích hoạt vùng gen ade1 tại locus gen ldhA đã được chuẩn bị. Cụ thể, đoạn vector pUC18 được khuếch đại bằng cách sử dụng pUC18 làm khuôn với các đoạn mồi pUC18-Pae1-F3 (SEQ ID NO: 38) và pUC18-Hind3-R3 (SEQ ID NO: 39).

Đoạn vị trí trình tự khởi đầu gen *ldhA* chứa một phần gen *ldhA* được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5384 làm khuôn với các đoạn mồi *ldhA*-upstr-F (SEQ ID NO: 130) và *ldhA*-upstr-R2 (SEQ ID NO: 131). Đoạn vị trí trình tự kết thúc gen *ldhA* được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5384 làm khuôn với các đoạn mồi *ldhA*-downstr-F2 (SEQ ID NO: 132) và *ldhA*-downstr-R (SEQ ID NO: 133). Ba đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị pknock-in (*ldhA*).

Sau đó, đoạn ADN được khuếch đại bằng cách sử dụng pknock-in (*ldhA*) làm khuôn với các đoạn mồi *ldhA*-upstr-R (SEQ ID NO: 134) và *ldhA*-downstr-F (SEQ ID NO: 135). Đoạn vùng gen *ade1* được khuếch đại bằng cách sử dụng ADN hệ gen của chủng 5384 làm khuôn với các đoạn mồi *ade1*pro-R (SEQ ID NO: 136) và *ade1*ter-F (SEQ ID NO: 137). Hai đoạn này được gắn bằng cách sử dụng In-Fusion HD Cloning Kit (Clontech Laboratories, Inc.) để chuẩn bị *pade1* knock-in (*ldhA*).

(5) Chuẩn bị ADN sợi đôi

Đoạn ADN được khuếch đại bằng cách sử dụng plasmit *pade1* knock-in (*ldhA*) làm khuôn với các đoạn mồi *ldhA*-upstr-F2 (SEQ ID NO: 138) và *ldhA*-downstr-R2-P (SEQ ID NO: 139; phosphoryl hóa tận cùng 5'). Khuôn bị phân hủy do xử lý Dpn1 (Toyobo Co., Ltd.). Sau đó, sản phẩm được tinh chế bằng phương pháp xử lý phenol/cloroform/rượu isoamyl và xử lý kết tủa bằng ethanol. Sản phẩm tinh khiết được hòa tan trong một lượng nước không có DNaza thích hợp để thu được dung dịch ADN sợi đôi để phá vỡ gen *ldhA*.

(6) Chuẩn bị ADN sợi đơn

Sau đó, ADN sợi đôi thu được trong đoạn (5) được xử lý bằng Lambda Exonucleaza (NEW ENGLAND BioLabs Inc.) và sau đó được tinh chế theo cách tương tự như trên để thu được ADN sợi đơn. Quá trình xử lý Lambda Exonucleaza được thực hiện qua đêm ở 37°C.

(7) Chuyển gen bằng bắn gen

Bằng các quy trình tương tự như trong Ví dụ 1(10), dung dịch chứa hạt vàng ADN chứa ADN sợi đơn hoặc sợi đôi đã được chuẩn bị, và việc chuyển gen vào các bào tử của chủng AM002 được thực hiện sử dụng dung dịch này, sau đó là nuôi cấy các bào tử thu được. ADN sợi đơn hoặc ADN sợi đôi được sử dụng là ADN sợi đơn hoặc ADN sợi đôi được chuẩn bị trong đoạn (6) hoặc (5). Các vector biểu hiện TALEN được sử dụng là padh-LeftTALEN-ldhA (RO) và padh-RightTALEN-ldhA (RO) được chuẩn bị trong đoạn (3). Vector biểu hiện exonucleaza được sử dụng là pldh-exo1 được chuẩn bị trong Ví dụ 1(4). Tỷ lệ nồng độ giữa padh-LeftTALEN-ldhA (RO), padh-RightTALEN-ldhA (RO), pldh-exo1, và ADN sợi đơn hoặc ADN sợi đôi trong dung dịch ADN được đặt thành khoảng 1:1:1:2.

(8) Lựa chọn chủng khuyết gen ldhA

Các bào tử được thu thập từ tế bào nấm được nuôi cấy trong đoạn (7). Việc phân lập các chủng nấm và chuẩn bị dung dịch khuôn hệ gen được thực hiện theo các quy trình tương tự như trong Ví dụ 1(11). Sau đó, chủng khuyết gen ldhA với đoạn vùng gen ade1 được kích hoạt tại locus gen ldhA đã được chọn lọc bằng phương pháp PCR khuôn lạc sử dụng dung dịch khuôn hệ gen

làm khuôn. Phương pháp PCR khuôn lạc được thực hiện sử dụng dung dịch khuôn hệ gen, đoạn môi *ldhA*-up (SEQ ID NO: 140) và *ade1*-15 (SEQ ID NO: 141), và KOD FX Neo (Toyobo Co., Ltd.). Phương pháp PCR khuôn lạc sử dụng các đoạn môi này sẽ khuếch đại đoạn ADN có chiều dài thích hợp nếu đoạn vùng gen *ade1* được kích hoạt tại locus gen *ldhA*. Bằng phương pháp PCR khuôn lạc, thu được chủng nấm có đoạn khuếch đại ADN thu được là chủng knock-in (chủng khuyết gen *ldhA*).

(9) So sánh hiệu quả tái tổ hợp tương đồng ở *Rhizopus oryzae*

(a) ADN sợi đơn và ADN sợi đôi

Hiệu quả tái tổ hợp tương đồng được so sánh giữa ADN sợi đôi và ADN sợi đơn được sử dụng làm ADN người cho. Các kết quả được thể hiện trong Bảng 11. Không thể thu được chủng tái tổ hợp tương đồng có trình tự chuỗi dài được chuyển bằng cách sử dụng ADN sợi đôi, trong khi chủng tái tổ hợp tương đồng có trình tự chuỗi dài được chuyển đã thu được thành công bằng cách sử dụng ADN sợi đơn.

Bảng 11

Điều kiện thử nghiệm	Chiều dài trình tự tương đồng (bp)	Số lượng các chủng phân lập	Số lượng các chủng tái tổ hợp tương đồng	Hiệu quả tái tổ hợp tương đồng (%)
ADN sợi đôi	1000	24	0	0
ADN sợi đơn	1000	24	18	75

* "Chiều dài trình tự tương đồng" thể hiện rằng các trình tự tương đồng có chiều dài được hiển thị trong bảng có mặt ở hai đầu ngược chiều và xuôi chiều của ADN sợi đơn hoặc ADN sợi đôi.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất nấm đột biến thuộc chi *Rhizopus*, trong đó phương pháp này bao gồm:

chuyển ADN nucleaza có thể lập trình và ADN sợi đơn vào nấm chủ thuộc chi *Rhizopus*, và

thay thế vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt cho ADN nucleaza có thể lập trình trong ADN hệ gen của vật chủ bằng ADN sợi đơn thông qua tái tổ hợp tương đồng;

trong đó chiều dài của ADN sợi đơn là từ 2000 bp đến 11 kbp,

ADN sợi đơn bao gồm hai trình tự ADN tương ứng bao gồm các vùng tương đồng với vùng ngược chiều và vùng xuôi chiều của vị trí phân cắt trong ADN hệ gen của nấm chủ thuộc chi *Rhizopus*, và

mỗi chiều dài của hai trình tự ADN là từ 25 bp trở lên.

2. Phương pháp theo điểm 1, trong đó ADN sợi đơn bao gồm trình tự chèn được chèn vào ADN hệ gen của nấm chủ thuộc chi *Rhizopus*.

3. Phương pháp theo điểm 2, trong đó chiều dài của trình tự được chèn là từ 1 kbp đến 10 kbp.

4. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 3, trong đó ADN nucleaza có thể lập trình là ADN nucleaza nhân tạo dựa trên công nghệ TALEN, CRISPR-Cas9, CRISPR-Cpf1 hoặc ZFN.

5. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 4, trong đó nấm thuộc chi *Rhizopus* là *Rhizopus delemar* hoặc *Rhizopus oryzae*.
6. Phương pháp theo điểm bất kỳ trong số các điểm từ 1 đến 5, trong đó phương pháp này bao gồm thêm bước biểu hiện exonucleaza ngoại lai trong nấm chủ thuộc chi *Rhizopus*.
7. Phương pháp theo điểm 6, trong đó exonucleaza ngoại lai là exonucleaza có nguồn gốc từ nấm thuộc chi *Rhizopus*.

DANH MỤC TRÌNH TỰ

<110> Kao Corporaiton

<120> Phương pháp sản xuất nấm sợi đột biến

<130> KS1478

<150> JP2016-019676

<151> 2016-02-04

<150> JP2016-152972

<151> 2016-08-03

<160> 152

<170> Phiên bản sáng chế 3.5

<210> 1

<211> 1735

<212> ADN

<213> Rhizopus delemar

<220>

<223> pdc1

<400> 1

atgcctgcta ttaaaatcgg tcaacatctc cttaaccgtc ttaaggaaat caacattgat	60
gttgtctttg gtgttcctgg tgatttcaac atggtaagca gacaattgaa ttgaacgaga	120
gcctataaac ttattatttc tatagccott gttggatatc attgaagatg acccagaact	180
tacctggggg aacaatgcc aagaattgaa tgcattctat gcagctgatg gttatgctcg	240

tattcgtggg	gcagggtgctg	ttgtcactac	ctttgggtgta	ggtgagctgt	ctgctgtcaa	300
cggtattgct	ggttcatact	ctgagatgct	tcccgtgatt	cacatcgtcg	gtactccttc	360
tactaaatcc	caagctgccg	gtgccatgct	tcaccactct	ttgggtgacg	gtaactttga	420
tgtgtttcttc	aacatgtcct	ccatgattgc	ctgtgcctct	gctcacctca	agaacaaaac	480
ggccattgca	gaaattgacc	gtgtgatctc	ccaagctggt	ctctccaagc	gtacagggtta	540
cattgggtatc	cctatcgatc	tgatcaagac	tgagggttga	gtacctgagc	ccattcctgc	600
cctcaagacc	gaattaccca	aaaaccacgc	tgatgtccaa	gcgattgcct	tgagagtggg	660
cacggatgcg	atcgccaaag	cccaattccc	tgtgattggt	gtcgatggct	gtgtgcttcg	720
ccagagatgc	caaaaggcag	tacaagcctt	tatcgaacgt	actggtttcc	ctacttatgt	780
tgtccttatg	ggtaaggggtg	ccgttgacga	atcctctgtg	agttaccgtg	gctgctactc	840
gggtaatgtc	acattggaag	cagtgaatga	agagatcaag	caagccgatt	tgatcatcga	900
agtgggctcc	atcaagtctg	atttcaacac	gggtaacttt	tcatactctc	tcgaccgttc	960
caagacgata	acottgcact	cctttgccac	catcgtgttt	tgtgctgaat	acaaaaaggt	1020
ctccatgctc	gaattcattc	ctctcttgac	ccaagccctt	cccgaacaac	cccgtcaatt	1080
caacctgggt	ccccgcccaa	gaccgcgtacc	tatccaaccc	ggtaccgaaa	tcaccacaaa	1140
ctacttttgg	cacaaggtag	ccgaattcat	ggatgagaac	gccattgtct	gtgccgagac	1200
cggtacagct	gaatttgctt	cactcaacat	ggacggaccc	aagggaacga	cttatatcac	1260
ccaattcctc	tggggctcta	tcggtttctc	agtaggtgcc	gctgtgggtg	ctgcgatcgc	1320
cgctcgtgat	cgctcgtgtg	atctctttgt	cggtgatggg	tccttccaat	tgacctgtca	1380
agaaatctct	ggcttccttc	gccatgggtt	gacacctgtg	atcttcttgc	tgaacaatga	1440
cggttacttg	atcgaaaaac	tcattcacgg	tcccgaacgt	gcctataata	actttcaa	1500
gtgggaatac	agcaagacgc	ttgattat	cggtgctcat	cttgaacaca	acaagtccat	1560
gggtgttcct	cccgttggct	tcgaaggcaa	ggtagccaca	cgcgatgaat	ttgaatccgc	1620
catgagacag	gttcaagcca	atcctgacaa	gattcatttc	cttgaagtca	ttatgcctca	1680
atttgactct	cctcgtgaac	ttgaactcct	ggttgccaac	tctgaaaacc	gttaa	1735

<210> 2

<211> 3006

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> LeftTALEN-pdc1

<400> 2

atggactaca aagaccatga cgggtgattat aaagatcatg acatcgatta caaggatgac	60
gatgacaaga tggcccccaa gaagaagagg aagggtgggca ttcaccgcgg ggtacctatg	120
gtggacttga ggacactcgg ttattcgcaa cagcaacagg agaaaatcaa gcctaaggtc	180
aggagcaccg tcgcgcaaca ccacgaggcg cttgtggggc atggcttcac tcatgcgcat	240
attgtcgcgc ttccacagca ccctgcggcg cttgggacgg tggctgtcaa ataccaagat	300
atgattgcgg ccctgcccga agccacgcac gaggcaattg taggggtcgg taaacagtgg	360
tcgggagcgc gagcacttga ggcgctgctg actgtggcgg gtgagcttag ggggcctccg	420
ctccagctcg acaccgggca gctgctgaag atcgcgaa gaagggggagt aacagcggta	480
gaggcagtgc acgcctggcg caatgcgctc accggggccc cttgaactt gaccccagac	540
caggtagtcg caatcgcgaa caataatggg ggaaagcaag ccctggaaac cgtgcaaagg	600
ttgttgccgg tcctttgtca agaccacggc ctgactcccg atcaagttgt agcgattgcg	660
tcgcatgacg gagggaaaca agcattggag actgtccaac ggctccttcc cgtgttgtgt	720
caagcccacg gtttgacgcc tgcacaagtg gtcgccatcg ccagccatga tggcggtaa	780
caggcgctgg aaacagtaca gcgcctgctg cctgtactgt gccaggatca tggactgacc	840
ccagaccagg tagtcgcaat cgcgtcaaac ggagggggaa agcaagccct ggaaaccgtg	900
caaaggttgt tgccggtcct ttgtcaagac cacggcctga cccagacca ggtagtcgca	960
atcgcgaaaca ataatggggg aaagcaagcc ctggaaaccg tgcaaagggtt gttgccggtc	1020
ctttgtcaag accatggcct gactcccgat caagttgtag cgattgcgtc gcatgacgga	1080
gggaaacaag cattggagac tgtccaacgg ctccctcccg tgttgtgtca agcccacggt	1140
ttgacgcctg cacaagtggg cgccatcgcc agcaatggcg gcggtaaagca ggcgctggaa	1200
acagtacagc gcctgctgcc tgtactgtgc caggatcatg gactgacccc agaccaggta	1260
gtcgcaatcg cgtcgaaatc tgggggaaag caagccctgg aaaccgtgca aaggttgttg	1320
ccggtccttt gtcaagacca cggcctgacc ccagaccagg tagtcgcaat cgcgtcaaac	1380
ggagggggaa agcaagccct ggaaaccgtg caaaggttgt tgccggtcct ttgtcaagac	1440
cacggcctga ctcccgatca agttgtagcg attgcgtcca acggtggagg gaaacaagca	1500

ttggagactg tccaacggct ccttcccgtg ttgtgtcaag cccatggatt gaccccagac	1560
caggtagtcg caatcgcgtc aaacattggg ggaaagcaag ccctggaaac cgtgcaaagg	1620
ttgttgccgg tcctttgtca agaccacggc ctgactcccg atcaagttgt agcgattgcg	1680
tcgaacattg gagggaaaca agcattggag actgtccaac ggctccttcc cgtgttgtgt	1740
caagcccacg gtttgacgcc tgcacaagtg gtcgccatcg ccagcaatat tggcggtaag	1800
caggcgctgg aaacagtaca gcgcctgctg cctgtactgt gccaggatca tggactgacc	1860
ccagaccagg tagtcgcaat cgcgtcaaac attgggggaa agcaagccct ggaaaccgtg	1920
caaaggttgt tgccggtcct ttgtcaagac cacggcctga cccagacca ggtagtcgca	1980
atcgcgtcaa acggaggggg aaagcaagcc ctggaaaccg tgcaaagggtt gttgccggtc	2040
ctttgtcaag accacggcct gactcccgat caagttgtag cgattgcgtc gcatgacgga	2100
gggaaacaag cattggagac tgtccaacgg ctccctcccg tgttgtgtca agcccacggt	2160
ctgacacccg aacagggtgg cgccattgct aataataacg gaggacggcc agccttgag	2220
tccatcgtag cccaattgtc caggcccgat ccgcgcttgg ctgcgttaac gaatgaccat	2280
ctggtggcgt tggcatgtct tggtgacga ccgcgctcg atgcagtcaa aaagggtctg	2340
cctcatgctc ccgcattgat caaaagaacc aaccggcgga ttcccagag aacttcccat	2400
cgagtcgcgg gatcccaact agtcaaaagt gaactggagg agaagaaatc tgaacttcgt	2460
cataaattga aatatgtgcc tcatgaatat attgaattaa ttgaaattgc cagaaattcc	2520
actcaggata gaattcttga aatgaaggta atggaatttt ttatgaaagt ttatggatat	2580
agaggtaaac atttggttgg atcaaggaaa ccggacggag caatttatac tgcggatct	2640
cctattgatt acggtgtgat cgtggatact aaagcttata gcggagggtta taatctgcca	2700
attggccaag cagatgaaat gcaacgatat gtcgaagaaa atcaaacacg aaacaaacat	2760
atcaacccta atgaatggtg gaaagtctat ccatcttctg taacggaatt taagttttta	2820
tttgtgagtg gtcactttta aggaaactac aaagctcagc ttacacgatt aaatcatatc	2880
actaattgta atggagctgt tcttagtgta gaagagcttt taattggtgg agaaatgatt	2940
aaagccggca cattaacctt agaggaagtc agacggaaat ttaataacgg cgagataaac	3000
ttttaa	3006

<210> 3

<211> 3006

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> RightTALEN-pdc1

<400> 3

atggactaca aagaccatga cgggtgattat aaagatcatg acatcgatta caaggatgac	60
gatgacaaga tggcccccaa gaagaagagg aagggtgggca ttcaccgcgg ggtacctatg	120
gtggacttga ggacactcgg ttattcgcaa cagcaacagg agaaaatcaa gcctaaggtc	180
aggagcaccg tcgcgcaaca ccacgaggcg cttgtggggc atggcttcac tcatgcgcat	240
attgtcgcgc ttccacagca ccctgcggcg cttgggacgg tggctgtcaa ataccaagat	300
atgattgcgg ccctgcccga agccacgcac gaggcaattg taggggtcgg taaacagtgg	360
tcgggagcgc gagcacttga ggcgctgctg actgtggcgg gtgagcttag ggggcctccg	420
ctccagctcg acaccgggca gctgctgaag atcgcgaaag gagggggagt aacagcggta	480
gaggcagtgc acgcctggcg caatgcgctc accggggccc cttgaactt gaccccagac	540
caggtagtcg caatcgcgtc aaacggaggg ggaaagcaag ccctggaaac cgtgcaaagg	600
ttgttgccgg tcctttgtca agaccacggc ctgactcccg atcaagttgt agcgattgcg	660
aataacaatg gagggaaaca agcattggag actgtccaac ggctccttcc cgtgttgtgt	720
caagcccacg gtttgacgcc tgcacaagtg gtcgccatcg ccagcaatat tggcggtaa	780
caggcgctgg aaacagtaca gcgcctgctg cctgtactgt gccaggatca tggactgacc	840
ccagaccagg tagtcgcaat cgcgtcaaac ggagggggaa agcaagccct ggaaaccgtg	900
caaaggttgt tgccggtcct ttgtcaagac cacggcctga cccagacca ggtagtcgca	960
atcgcgtaaa acggaggggg aaagcaagcc ctggaaaccg tgcaaagggtt gttgccggtc	1020
ctttgtcaag accacggcct gactcccgat caagttgtag cgattgcgtc gaacggtgga	1080
gggaaacaag cattggagac tgtccaacgg ctcttccccg tgttgtgtca agcccacggt	1140
ttgacgcctg cacaagtggc cgccatcgcc agccatgatg gcggtaaagca ggcgctggaa	1200
acagtacagc gcctgctgcc tgtactgtgc caggatcatg gactgacccc agaccaggta	1260
gtcgcaatcg cgtcacatga cgggggaaag caagccctgg aaaccgtgca aagggtgttg	1320
ccggtccttt gtcaagacca cggcctgacc ccagaccagg tagtcgcaat cgcgtcaaac	1380
ggagggggaa agcaagccct ggaaaccgtg caaaggttgt tgccggtcct ttgtcaagac	1440

```

cacggcctga ctcccgatca agttgtagcg attgcgtcca acggtggagg gaaacaagca 1500
ttggagactg tccaacggct ccttcccgtg ttgtgtcaag cccatggatt gaccccagac 1560
caggtagtcg caatcgcgtc aaacattggg ggaaagcaag ccctggaaac cgtgcaaagg 1620
ttgttgccgg tcctttgtca agaccacggc ctgactcccg atcaagttgt agcgattgcg 1680
tcgaacattg gagggaaaca agcattggag actgtccaac ggctccttcc cgtgttgtgt 1740
caagcccacg gtttgacgcc tgcacaagtg gtcgccatcg ccaacaacaa cggcggtaag 1800
caggcgctgg aaacagtaca gcgcctgctg cctgtactgt gccaggatca tggactgacc 1860
ccagaccagg tagtcgcaat cgcgtcaaac attgggggaa agcaagccct ggaaaccgtg 1920
caaaggttgt tgccggtcct ttgtcaagac cacggcctga ccccagacca ggtagtcgca 1980
atcgcgtcac atgacggggg aaagcaagcc ctggaaaccg tgcaaagggtt gttgccggtc 2040
ctttgtcaag accatggcct gactcccgat caagttgtag cgattgcgaa taacaatgga 2100
gggaaacaag cattggagac tgtccaacgg ctccctcccg tgttgtgtca agcccacggt 2160
ctgacacccg aacagggtgg cgccattgct aataataacg gaggacggcc agccttgag 2220
tccatcgtag cccaattgtc caggcccgat ccgcggttgg ctgcgttaac gaatgaccat 2280
ctggtggcgt tggcatgtct tggtgagcga ccgcgctcg atgcagtcaa aaagggctctg 2340
cctcatgctc ccgcattgat caaaagaacc aaccggcgga ttcccagag aacttcccat 2400
cgagtcgcgg gatcccaact agtcaaaagt gaactggagg agaagaaatc tgaacttcgt 2460
cataaattga aatatgtgcc tcatgaatat attgaattaa ttgaaattgc cagaaattcc 2520
actcaggata gaattcttga aatgaaggta atggaatttt ttatgaaagt ttatggatat 2580
agaggtaaac atttggttgg atcaaggaaa ccggacggag caatttatac tgtcggatct 2640
cctattgatt acggtgtgat cgtggatact aaagcttata gcggagggtta taatctgcca 2700
attggccaag cagatgaaat gcaacgatat gtcgaagaaa atcaaacacg aaacaaacat 2760
atcaacccta atgaatggtg gaaagtctat ccatcttctg taacggaatt taagttttta 2820
tttgtgagtg gtcactttaa aggaaactac aaagctcagc ttacacgatt aaatcatatc 2880
actaattgta atggagctgt tcttagtgta gaagagcttt taattggtgg agaaatgatt 2940
aaagccggca cattaacctt agaggaagtc agacggaaat ttaataacgg cgagataaac 3000
ttttaa 3006

```

<210> 4

<211> 1733

<212> ADN

<213> Rhizopus delemar

<220>

<223> pdc2

<400> 4

atgccttcta tccaaatcgg tcaacatctt ctttaaccgtc tcaaggaaat taacattgat	60
gtagtctttg gtgttcctgg tgacttcaac atggtaaata aatataaaga gtgatgaaca	120
agttgtatth atatccttct tagcctttac tggatatcat tgaagacgac ccaaagctta	180
cgtggggtaa taatgctaata gaattgaacg cttcttatgc agctgatggg tatgctcgta	240
ttcgtgggtgc ggggtgctgc gtcactacgt ttgggtgttg tgaactctct gctgtgaatg	300
gtgttgctgg ttcttatgct gagatgcttc ctgtgattca cattgtcggc acaccttcta	360
ctaaatctca agctgctggg gctatgcttc accactctct ggggtgatgg aattttgatg	420
tgttcttcaa catgtcctcc atgattgcct gtgcacttac tcatctcaag aagcagactg	480
ctatcgcaga aattgaccgt gtcactcttc aagctgtcct ctcaaacgt acgggttaca	540
tgggtatccc tattgacttg atcaagaccg aagtcgaaat tccagaggaa ttaagtcctc	600
ttcagaccac acttcccaag aacaaccag aagtccaagc gatcgccttg aaggctgca	660
ccgaagccat tcagtctgcc aagcatcctg tgatcattgt cgatggctgt gtccttcgtc	720
accgctgtca gaaaccgctc caggaattca tcaccgcttc cggtttccca acctacgtgg	780
ctcccatggg caagggcgcc gtggatgagt ccattgagaa tttccgtggg tgttattcgg	840
gtaacgtgac actggaggct gtcaatgaag aaatcaaact ggccgatttg attattgaga	900
ttggttcgat caagtctgat ttcaacacag gtaacttttc atactccctc gatcgttcca	960
agacgatcac cttgcattct tttgctacga tcgtgttttg tgcagagtac caaaaggctc	1020
ccatgatgga attcattcct ctcttgacct aggcccttcc ccaacaacct cgtatgttca	1080
acttgggtcc acgtgcaaag cctgtgccga tccaaccgga aactgaaatc acccacaatt	1140
acttttggca caaggtacct gactacatgg aagaaaacgc gattgtctgt gcagagaccg	1200
gtacagccga atttgcttca ctcaacatgg atggccccaa gggtagcaca tatatcacc	1260
aaatcctttg gggttctatc ggtttcacgg taggtgcctc tgtaggtgcc gctatcgccg	1320
ctcgtgatcg tcgtgtgtat ctctttgtcg gtgatgggtc cttccaactg acctgccaag	1380

aaatcgctgt tttccttctg catggcctca cacctgtcat cttcttgctg aacaatgacg 1440
 gttacttgat tgaaaaactc attcacgggc ctgatcgctg ttacaataat taccaaatgt 1500
 ggaattacca caagaccctt gactatttct gtgctcatct tgaacacaac aagtccatgg 1560
 gtgttcctcc cgttggtttc gaaggcaagg tggccacacg tgatgaattc gaatcggcca 1620
 tgaagcaagt tcaagccaac cctaacaaga ttcatcttct tgaagtcatt atgcctcaat 1680
 ttgatgcccc tcgtgaactt gaattattgg ttgccaactc tgaaaaccga taa 1733

<210> 5

<211> 3075

<212> ADN

<213> Nhân tạo

<220>

<223> LeftTALEN-pdc2

<400> 5

atgggaaaac ctattcctaa tcctctgctg ggccctggatt ctaccggagg catggcccct 60
 aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag 120
 cagcaggaga agatcaagcc caaggtgaga tctacagtgg cccagcacca cgaagccctg 180
 gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcacc ctcgcccctg 240
 ggaacagtgg ccgtgaaata tcaggatatg attgccgccc tgccctgaggc cacacacgaa 300
 gccattgtgg gagtgggaaa acagtggctt ggagccagag ccctggaagc cctgctgaca 360
 gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt 420
 gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca 480
 ggagcccctc tgaacctgac ccccgaaacag gtggtggcca ttgccagcaa caacggcggc 540
 aagcaggccc tggaaccgt gcagagactg ctgcccgctg tgtgccaggc ccatggcctg 600
 acacctgaac aggtggtggc tatcgctctt aacaacggag gaaaacaggc tctggaaaca 660
 gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg 720
 gctattgctt ccaacggcgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca 780
 gtgctgtgcc aggtcacgg actgaccccc gaacaggtgg tggccattgc cagcaacaac 840

ggcggcaagc	aggccctgga	aaccgtgcag	agactgctgc	ccgtgctgtg	ccaggcccat	900
ggcctgacac	ctgaacaggt	ggtggctatc	gcctctaacg	gcggaggaaa	acaggctctg	960
gaaacagtgc	agcggctgct	gcctgtgctg	tgtcaggctc	acggcttgac	tccagaacag	1020
gtggtggcta	ttgcttccaa	cggcgggggg	aaacaggccc	tggaaaactgt	gcagcgcctg	1080
ctgccagtgc	tgtgccaggc	tcacgggctg	acccccgaac	aggtgggtggc	cattgccagc	1140
aacaacggcg	gcaagcaggc	cctggaaacc	gtgcagagac	tgctgcccgt	gctgtgccag	1200
gcccattggcc	tgacacctga	acaggtggtg	gctatcgctt	ctcacgacgg	aggaaaacag	1260
gctctggaaa	cagtgcagcg	gctgctgcct	gtgctgtgtc	aggtcacagg	cttgactcca	1320
gaacagggtg	tggctattgc	ttccaacggc	gggggggaaac	aggccctgga	aactgtgcag	1380
cgctgctgc	cagtgtgtg	ccaggctcac	ggcctcactc	ccgaacaggt	ggtggccatt	1440
gccagcaaca	acggcggcaa	gcaggccctg	gaaaccgtgc	agagactgct	gcccgtgctg	1500
tgccaggccc	atggcctgac	acctgaacag	gtggtggcta	tcgcctctaa	caacggagga	1560
aaacaggctc	tggaaacagt	gcagcggctg	ctgcctgtgc	tgtgtcaggc	tcacggcttg	1620
actccagaac	aggtgggtgg	tattgcttcc	aacggcgggg	ggaaacaggc	cctggaaaact	1680
gtgcagcgcc	tgctgccagt	gctgtgccag	gctcacggac	tgacccccga	acaggtggtg	1740
gccattgcca	gcaacggcgg	cggcaagcag	gccctggaaa	ccgtgcagag	actgctgccc	1800
gtgctgtgcc	aggcccatgg	cctgacacct	gaacagggtg	tggctatcgc	ctctcacgac	1860
ggaggaaaac	aagcactcga	gacagtgcag	cggctgctgc	ctgtgctgtg	tcaggctcac	1920
ggcttgactc	cagaacaggt	ggtggctatt	gcttccaacg	gcgggggggaa	acaggccctg	1980
gaaactgtgc	agcgctgct	gccagtgtg	tgccaggctc	acgggctgac	ccccgaacag	2040
gtggtggcca	ttgccagcaa	cggcggcggc	aagcaggccc	tggaaaccgt	gcagagactg	2100
ctgcccgtgc	tgtgccaggc	ccatggcctg	acacctgaac	aggtgggtgg	tatcgctctt	2160
aatatcggag	gaaaacaggc	tctggaaaca	gtgcagcggc	tgctgcctgt	gctgtgtcag	2220
gctcacggct	tgactccaca	gcaggtcgtg	gcaattgcta	gcaacggcgg	cggacggccc	2280
gccctggaga	gcattgtggc	ccagctgtct	agacctgata	ctgccctggc	cgccctgaca	2340
aatgatcacc	tgggtggcct	ggcctgtctg	ggaggcagac	ctgccctgga	tgccgtgaaa	2400
aaaggactgc	ctcacgcccc	tgccctgata	aagagaacaa	atagaagaat	ccccgagcgg	2460
acctctcaca	gagtggccgg	atcccagctg	gtgaaatctg	agctggagga	gaagaagtct	2520
gagctgagac	acaagctgaa	gtacgtgcct	cacgagtaca	tcgagctgat	cgagatcgcc	2580
agaaatagca	cccaggatag	aatcctggag	atgaagggtg	tggagttctt	catgaagggtg	2640

tacggctaca gaggaaagca cctgggagga agcagaaaac ctgacggagc catttataca	2700
gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac	2760
aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg	2820
aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc	2880
aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggccagct gaccaggctg	2940
aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga	3000
gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagtt caacaacggc	3060
gagatcaact tctga	3075

<210> 6

<211> 3075

<212> ADN

<213> Nhân tạo

<220>

<223> RightTALEN-pdc2

<400> 6

atgggaaaac ctattcctaa tcctctgctg ggccctggatt ctaccggagg catggcccct	60
aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag	120
cagcaggaga agatcaagcc caaggtgaga tctacagtgg ccacagacca cgaagccctg	180
gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcacc tgccgcctg	240
ggaacagtgg ccgtgaaata tcaggatatg attgccgcc tgccctgaggc cacacacgaa	300
gccattgtgg gagtgggaaa acagtgggtc ggagccagag ccctggaagc cctgctgaca	360
gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt	420
gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca	480
ggagcccctc tgaacctgac cccgaacag gtggtggcca ttgccagcaa caacggcggc	540
aagcaggccc tggaaaccgt gcagagactg ctgcccgctg tgtgccaggc ccatggcctg	600
acacctgaac aggtggtggc tatcgctct cagcagggag gaaaacaggc tctggaaaca	660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg	720

gctattgctt	cccacgacgg	ggggaaacag	gccctggaaa	ctgtgcagcg	cctgctgcca	780
gtgctgtgcc	aggtcacgg	actgaccccc	gaacagggtg	tggccattgc	cagcaacaac	840
ggcggcaagc	aggccctgga	aaccgtgcag	agactgctgc	ccgtgctgtg	ccaggcccat	900
ggcctgacac	ctgaacagg	ggtggctatc	gcctctaata	tcggaggaaa	acaggctctg	960
gaaacagtgc	agcggctgct	gcctgtgctg	tgtcaggctc	acggcttgac	tccagaacag	1020
gtggtggcta	ttgcttccca	cgacgggggg	aaacaggccc	tggaaactgt	gcagcgctg	1080
ctgccagtgc	tgtgccaggc	tcacgggctg	acccccgaac	aggtggtggc	cattgccagc	1140
aacatcggcg	gcaagcaggc	cctggaaacc	gtgcagagac	tgctgcccgt	gctgtgccag	1200
gcccattggcc	tgacacctga	acagggtggtg	gctatcgctt	ctaataatcg	aggaaaacag	1260
gctctggaaa	cagtgcagcg	gctgctgcct	gtgctgtgtc	aggtcacgg	cttgactcca	1320
gaacagggtg	tggctattgc	ttccaacggc	ggggggaaaac	aggccctgga	aactgtgcag	1380
cgctgctgc	cagtgtgtg	ccaggctcac	ggcctcactc	ccgaacagg	ggtggccatt	1440
gccagcaaca	acggcggcaa	gcaggccctg	gaaaccgtgc	agagactgct	gcccgtgctg	1500
tgccaggccc	atggcctgac	acctgaacag	gtggtggcta	tcgcctctaa	cggcggagga	1560
aaacaggctc	tggaaacagt	gcagcggctg	ctgcctgtgc	tgtgtcaggc	tcacggcttg	1620
actccagaac	aggtggtggc	tattgcttcc	aacaacgggg	ggaaacaggc	cctggaaaact	1680
gtgcagcgcc	tgctgccagt	gctgtgccag	gctcacggac	tgacccccga	acagggtggtg	1740
gccattgcca	gcaacatcgg	cggcaagcag	gccctggaaa	ccgtgcagag	actgctgccc	1800
gtgctgtgcc	aggcccatgg	cctgacacct	gaacagggtg	tggctatcgc	ctctaataatc	1860
ggaggaaaac	aagcactcga	gacagtgcag	cggctgctgc	ctgtgctgtg	tcaggctcac	1920
ggcttgactc	cagaacagg	ggtggctatt	gcttccaacg	gcggggggaa	acaggccctg	1980
gaaactgtgc	agcgctgct	gccagtgtg	tgccaggctc	acgggctgac	ccccgaacag	2040
gtggtggcca	ttgccagcca	cgacggcggc	aagcaggccc	tggaaaccgt	gcagagactg	2100
ctgcccgtgc	tgtgccaggc	ccatggcctg	acacctgaac	aggtggtggc	tatcgctctt	2160
aatatcggag	gaaaacaggc	tctggaaaca	gtgcagcggc	tgctgcctgt	gctgtgtcag	2220
gctcacggct	tgactccaca	gcaggctgtg	gcaattgcta	gccacgacgg	cggacggccc	2280
gcctggaga	gcattgtggc	ccagctgtct	agacctgatc	ctgccctggc	cgccctgaca	2340
aatgatcacc	tgggtggcct	ggcctgtctg	ggaggcagac	ctgccctgga	tgccgtgaaa	2400
aaaggactgc	ctcacgcccc	tgccctgatc	aagagaacaa	atagaagaat	ccccgagcgg	2460
acctctcaca	gagtggccgg	atcccagctg	gtgaaatctg	agctggagga	gaagaagtct	2520

gagctgagac acaagctgaa gtacgtgcct cacgagtaca tcgagctgat cgagatcgcc 2580
 agaaatagca cccaggatag aatcctggag atgaaggtga tggagttctt catgaaggtg 2640
 tacggctaca gaggaaagca cctgggagga agcagaaaac ctgacggagc catttataca 2700
 gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac 2760
 aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg 2820
 aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc 2880
 aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggccagct gaccaggctg 2940
 aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga 3000
 gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagtt caacaacggc 3060
 gagatcaact tctga 3075

<210> 7

<211> 2298

<212> ADN

<213> Rhizopus delemar

<220>

<223> trpC

<400> 7

atgaccactt tacttattga caactacgac agttttactt ataatgtcta tcaataacttg 60
 agctgccaaag ggcccaatgt agttgtctac agaaacgaca aaatcaccat ttccgaaatt 120
 gagcaattgg ctctctgcaa tattgtcatc tcacctggcc ctggccaccc ttccaccgat 180
 gccggtgtct ctcgagaggc cattcgagct tttgcaggaa agattcccat cttgggtatt 240
 tgtatgggtc agcaatgtat gtatgaagtg tacggtggta aagtgtcata tgcaggtgat 300
 attgtgcatg gcaaggcatc cagcatcaag catgacagtc gaggtatctt caagggcggt 360
 cctcaaaaaca acatggtcac tcgttaccat tcccttgctg gcatgccttc tactttacct 420
 gaaacattag aagtcactgc gactaccgac gatggtatca tcatgggcat tcgacacaag 480
 gaatacactg tcgaaggtgt tcagttccat cctgaaagta tcctttgtga acacggacat 540
 acgatgatca acaacttctt aagcttgctg ggtggcacct gggaagagaa tcctgcagcc 600

ggtgttgtct ttaagaaagc tcgttccgaa acacccaaaa tcagtgtctag tgaatcccaa	660
ctcgatctct ctcagcaaca acctgccgca gcaccttcca tcttgacccg catttactct	720
caacgactca aggatgttca ggcagccaag gagattcccc gccagacatt tgaagattta	780
gaaaaacttt taaagttgca cgtcgcccca cctcttcaag acgtcgtcgc tcgctgtcgt	840
caaagcaagc ccgccttgat ggccgaagtc aagcgtgcct ctccctcgaa aggaaacatt	900
gatgtttcgg ccaacgcggc tgagcaggca cttcaatatg ctttagcagg tgcaagcgtc	960
gtctctgttc tgactgaacc caaatggttc cgcggtacga ttcatgatat gcatcaggtc	1020
cgagaggcct tgagccatct gccaaccgt ccttgtgtgt tgagaaagga ttttattgtc	1080
gatcgctatc aaatcttggg aggtcgtctg tacggtgctg atactatctt gttgatcgtg	1140
gcatgctga atgatgaaca actgcacgaa ttgtatcact atgcgaaatc attaggtatg	1200
gaacccttgg tcgaagtcaa taatacggaa gagatggccc gtgccaatgc tttgggcgca	1260
cgtctgggtg gtgttaataa tcgcaacttg cacagctttg atgttgatat ggaaaccacg	1320
agtcgattgg tagagatggg gcctgaagga acgatcttgt gtgcactttc tggattact	1380
ggacgagctg atgttgaaat gtacgtcaaa cagggtgtgc acgctgtctt ggtgggtgaa	1440
gccctgatgc gtgcttgga tttgaaggag tttgtgtctg atttgttggg tcatgaaaag	1500
aaggatcctg tgcctgtgtc caaggaatca aaatcttcac tagtcaagggt atgtggtatc	1560
tctagtgtgg atgcagcagt tgaagcagcc aagtcagggg ctgacttgat tggctttatc	1620
tttgtgaaa agtccaaacg aaaagtgtct ttggaaagag ctcaagaaat cgtgtcctca	1680
gtgcgtgcgt tggatattca agtcaaacga acgttatcaa atgatgattc tcaactggat	1740
tggttccaga tgcacaagcg tctcttgga aagcgagcaa gaaaaccttt ggtagtggc	1800
gtgtttgtga atcaatcgat tgaatacatg actgagggtg caacgacagt cggactggac	1860
cttattcagc tgcattggaac cgaatcaacg gagcttgac gctatttacc cgtgcctgtc	1920
atcaaagctt tccatatacg cagtgggtgag ttcaatgaag ctgagatacc aaacctaaat	1980
caaccaggct cttatcatta tgtcttactg gacgctaaag tgcccagctt accatcggat	2040
caacaagggtg gacgtgggtg caagtttgat tgggtcaattg ctaccaaaat cgtgaaacat	2100
aggcactttg agtttttggg taatcaagat ttccctgtca tcttggctgg tgggttgat	2160
cctaccaatg tggcatctgc cattcaacag gtgaaaccct ggattgtgga tgtgtcgagt	2220
ggtgttgaaa cagatggagt gaaggattta gaaaagattc gtgcctttgt taaaactgtc	2280
cagtcaacac aattttta	2298

<210> 8

<211> 3075

<212> ADN

<213> Nhân tạo

<220>

<223> LeftTALEN-trpC

<400> 8

atgggaaaac ctattcctaa tcctctgctg ggccctggatt ctaccggagg catggcccct	60
aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag	120
cagcaggaga agatcaagcc caaggtgaga tctacagtgg cccagcacca cgaagccctg	180
gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcaccc tgccgccctg	240
ggaacagtgg ccgtgaaata tcaggatatg attgccgccc tgccctgaggc cacacacgaa	300
gccattgtgg gagtgggaaa acagtggctc ggagccagag ccctggaagc cctgctgaca	360
gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt	420
gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca	480
ggagcccctc tgaacctgac ccccgaaacag gtggtggcca ttgccagcaa caacggcggc	540
aagcaggccc tggaaaccgt gcagagactg ctgcccgctg tgtgccaggc ccatggcctg	600
acacctgaac aggtggtggc tatcgccctc cagcagggag gaaaacaggc tctggaaaca	660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg	720
gctattgctt cccacgacgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca	780
gtgctgtgcc aggtcacgg actgaccccc gaacagggtg tggtccattgc cagcaacatc	840
ggcggcaagc aggccttggg aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat	900
ggcctgacac ctgaacaggg ggtggctatc gcctctaata tcggaggaaa acaggctctg	960
gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag	1020
gtggtggcta ttgcttccaa caacgggggg aaacaggccc tggaaactgt gcagcgctg	1080
ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtggtggc cattgccagc	1140
aacaacggcg gcaagcaggc cctggaaacc gtgcagagac tgctgcccgt gctgtgccag	1200
gcccattggc tgacacctga acaggtggtg gctatcgct ctcacgacgg aggaaaacag	1260

gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggctcacgg cttgactcca	1320
gaacagggtgg tggctattgc ttccaacaac ggggggaaac aggccctgga aactgtgcag	1380
cgctgtctgc cagtgtctgt ccaggctcac ggctcactc ccgaacagggt ggtggccatt	1440
gccagccacg acggcggcaa gcaggccctg gaaaccgtgc agagactgct gcccgtgctg	1500
tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctca cgacggagga	1560
aaacaggctc tggaaacagt gcagcggctg ctgcctgtgc tgtgtcaggc tcacggcttg	1620
actccagaac aggtggtggc tattgcttcc aatattgggg ggaaacaggc cctggaaact	1680
gtgcagcgcc tgtgtccagt gctgtgccag gctcacggac tgacccccga acagggtggtg	1740
gccattgcca gcaacatcgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc	1800
gtgctgtgcc aggccatgg cctgacacct gaacagggtg tggctatcgc ctctaacggc	1860
ggaggaaaac aagcactcga gacagtgcag cggtgctgc ctgtgctgtg tcaggctcac	1920
ggcttgactc cagaacagggt ggtggctatt gcttccaaca acggggggaa acaggccctg	1980
gaaactgtgc agcgctgct gccagtgtg tgcaggctc acgggctgac ccccgaaacag	2040
gtggtggcca ttgccagcaa cggcggcggc aagcaggccc tggaaaccgt gcagagactg	2100
ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtggtggc tatcgctct	2160
aatatcggag gaaaacaggc tctggaaaca gtgcagcggc tgtgcctgt gctgtgtcag	2220
gctcacggct tgactccaca gcaggctgtg gcaattgcta gcaacaacgg cggacggccc	2280
gccctggaga gcattgtggc ccagctgtct agacctgatc ctgccctggc cgccctgaca	2340
aatgatcacc tgggtggcct ggctgtctg ggaggcagac ctgccctgga tgccgtgaaa	2400
aaaggactgc ctcacgcccc tgccctgatc aagagaacaa atagaagaat ccccgagcgg	2460
acctctcaca gagtggccgg atcccagctg gtgaaatctg agctggagga gaagaagtct	2520
gagctgagac acaagctgaa gtacgtgcct cacgagtaca tcgagctgat cgagatcgcc	2580
agaaatagca ccaggatag aatcctggag atgaagggtga tggagttctt catgaagggtg	2640
tacggctaca gaggaaagca cctgggagga agcagaaaac ctgacggagc catttataca	2700
gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac	2760
aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg	2820
aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc	2880
aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggccagct gaccaggctg	2940
aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga	3000
gagatgatta aggccggaac actgacactg gaggagggtga gaagaaagtt caacaacggc	3060

gagatcaact tctga

3075

<210> 9

<211> 3075

<212> ADN

<213> Nhân tạo

<220>

<223> RightTALEN-trpC

<400> 9

atgggaaaac ctattcctaa tcctctgctg ggcctggatt ctaccggagg catggcccct	60
aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag	120
cagcaggaga agatcaagcc caaggtgaga tctacagtgg cccagcacca cgaagccctg	180
gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcacc cgtccgcccctg	240
ggaacagtgg ccgtgaaata tcaggatatg attgccgccc tgccctgaggc cacacacgaa	300
gccattgtgg gagtgggaaa acagtggctc ggagccagag ccctggaagc cctgctgaca	360
gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt	420
gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca	480
ggagcccctc tgaacctgac cccgaacag gtggtggcca ttgccagcca cgacggcggc	540
aagcaggccc tggaaaccgt gcagagactg ctgcccgctg tgtgccaggc ccatggcctg	600
acacctgaac aggtggtggc tatcgccctc aacaacggag gaaaacaggc tctggaaaca	660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg	720
gctattgctt ccaacaacgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca	780
gtgctgtgcc aggtcacgg actgaccccc gaacagggtg tggccattgc cagcaacatc	840
ggcggcaagc aggccttgga aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat	900
ggcctgacac ctgaacagg ggtggctatc gcctctaata tcggaggaaa acaggctctg	960
gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag	1020
gtggtggcta ttgcttccaa tattgggggg aaacaggccc tggaaactgt gcagcgccctg	1080
ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtggtggc cattgccagc	1140

aacggcgggcg gcaagcaggc cctggaaacc gtgcagagac tgctgcccgt gctgtgccag	1200
gcccattggcc tgacacctga acagggtggcg gctatcgccct ctaacaacgg agggaaaacag	1260
gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggtcacagg cttgactcca	1320
gaacagggtgg tggctattgc ttccaacaac ggggggaaaac aggccctgga aactgtgcag	1380
cgctgtctgc cagtgtgtg ccaggctcac ggctcactc ccgaacagggt ggtggccatt	1440
gccagcaacg gcggcgggcaa gcaggccctg gaaaccgtgc agagactgct gcccgtgctg	1500
tgccaggccc atggcctgac acctgaacag gtgggtggcta tcgcctctaa caacggagga	1560
aaacaggctc tggaacacagt gcagcggtcg ctgcctgtgc tgtgtcaggc tcacggcttg	1620
actccagaac aggtgggtggc tattgcttcc aatattgggg ggaaacaggc cctggaaaact	1680
gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acagggtggcg	1740
gccattgcca gcaacggcgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc	1800
gtgctgtgcc agggccatgg cctgacacct gaacagggtg tggctatcgc ctctaacggc	1860
ggaggaaaac aagcactcga gacagtgcag cggctgctgc ctgtgctgtg tcaggctcac	1920
ggcttgactc cagaacagggt ggtggctatt gcttccaacg gcggggggaa acaggccctg	1980
gaaactgtgc agcgccctgct gccagtgtg tgccaggctc acgggctgac ccccgaaacag	2040
gtgggtggcca ttgccagcaa cggcgggcggc aagcaggccc tggaaaccgt gcagagactg	2100
ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtgggtggc tatcgccctc	2160
aacaacggag gaaaacaggc tctggaaaac gtgcagcggc tgctgcctgt gctgtgtcag	2220
gctcacggct tgactccaca gcaggctgtg gcaattgcta gcaacggcgg cggacggccc	2280
gccctggaga gcattgtggc ccagctgtct agacctgac ctgccctggc cgccctgaca	2340
aatgatcacc tgggtggccct ggccctgtctg ggaggcagac ctgccctgga tgccgtgaaa	2400
aaaggactgc ctcacgcccc tgccctgac aagagaacaa atagaagaat ccccgagcgg	2460
acctctcaca gagtggccgg atcccagctg gtgaaatctg agctggagga gaagaagtct	2520
gagctgagac acaagctgaa gtacgtgcct cacgagtaca tcgagctgat cgagatcgcc	2580
agaaatagca cccaggatag aatcctggag atgaagggtg tggagttctt catgaagggtg	2640
tacggctaca gaggaaagca cctgggagga agcagaaaac ctgacggagc catttataca	2700
gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac	2760
aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg	2820
aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc	2880
aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggccagct gaccaggctg	2940

aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga 3000
 gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagtt caacaacggc 3060
 gagatcaact tctga 3075

<210> 10

<211> 46

<212> ADN

<213> Nhân tạo

<220>

<223> oJK162

<400> 10

cgagctcgaa ttattttaaataaacagcaag ttaataatct agaggg 46

<210> 11

<211> 39

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> oJK163

<400> 11

tatgaccatg attacgatga gaggcaaaat gaagcgtac 39

<210> 12

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> oJK164

<400> 12

atttaaataa ttcgagctcg gtacccgggg

30

<210> 13

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> oJK165

<400> 13

cgtaatcatg gtcatactg

20

<210> 14

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> oJK202

<400> 14

tagagggaaa aagagagaat tgaaatagg

29

<210> 15

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> oJK204

<400> 15

ttttgttatt taattgtatt aattgataat g 31

<210> 16

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> oJK205

<400> 16

aattaaataa caaaatcatt ttaattacgc attttc 36

<210> 17

<211> 37

<212> ADN

<213> Nhân tạo

<220>

<223> oJK216

<400> 17

catgattacg cggccgcgcc attataaatgc actagtg

37

<210> 18

<211> 38

<212> ADN

<213> Nhân tạo

<220>

<223> oJK210

<400> 18

ctcttttttcc ctctaattgag aggcaaaatg aagcgtac

38

<210> 19

<211> 31

<212> ADN

<213> Nhân tạo

<220>

<223> oJK211

<400> 19

atttaaatgt aatcatgggc atagctgttt c

31

<210> 20

<211> 35

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-lost-F

<400> 20

tttaaattag agggaaaaag agagaattga aatag

35

<210> 21

<211> 35

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-lost-R

<400> 21

tccctctaatt ttaaattgaat tcgagctcgg taccc

35

<210> 22

<211> 31

<212> ADN

<213> Nhân tạo

<220>

<223> adhpro-R

<400> 22

ttttgttatt taattgtatt aattgataat g

31

<210> 23

<211> 27

<212> ADN

<213> Nhân tạo

<220>

<223> adhter-F

<400> 23

tcattttaat tacgcatttt catttac

27

<210> 24

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> adhpro-TALEN-F

<400> 24

aattaaataa caaaaatgga ctacaaagac catgacggtg

40

<210> 25

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> TALEN-adhter-R

<400> 25

gcgtaattaa aatgattaaa agtttatctc gccggttatta

40

<210> 26

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> adhpro-LifeTALEN-F

<400> 26

aattaaataa caaaaatggg aaaacctatt cctaatcctc tgctg

45

<210> 27

<211> 47

<212> ADN

<213> Nhân tạo

<220>

<223> LifeTALEN-adhter-R

<400> 27

gcgtaattaa aatgatcaga agttgatctc gccgttggtg aactttc

47

<210> 28

<211> 19

<212> ADN

<213> Nhân tạo

<220>

<223> pPTR1-sal1-F

<400> 28

gggtaccgag ctccaattc

19

<210> 29

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> pPTR1-sal1-R

<400> 29

ggggatcctc tagagtcgac

20

<210> 30

<211> 41

<212> ADN

<213> Nhân tạo

<220>

<223> sal1-ldhpro-F3

<400> 30

ctctagagga tcccctaggt gtggctgtgg tgaccatatt g

41

<210> 31

<211> 28

<212> ADN

<213> Nhân tạo

<220>

<223> ldhpro-R

<400> 31

gagaattata ttgtaaagaa aaataaag

28

<210> 32

<211> 50

<212> ADN

<213> Nhân tạo

<220>

<223> ldhpro-exo1-F2

<400> 32

tacaatataa ttctcatgaa aatccaagtt gcttctccta ttgaccaatc

50

<210> 33

<211> 51

<212> ADN

<213> Nhân tạo

<220>

<223> exo1-pdcter-R2

<400> 33

atgaattcta agatTTTTatc ttctttcatg agaaacacta aacttgataa c

51

<210> 34

<211> 24

<212> ADN

<213> Nhân tạo

<220>

<223> pdcTer-F

<400> 34

aatcttagaa ttcatctttt ttg

24

<210> 35

<211> 39

<212> ADN

<213> Nhân tạo

<220>

<223> pdcTer-sal1-R

<400> 35

tcgagctcgg tacccactct accgtctgct cttttgtct

39

<210> 36

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> adhpro-exo1-F

<400> 36

aattaaataa caaaaatgaa aatccaagtt gcttctccta ttgac

45

<210> 37

<211> 46

<212> ADN

<213> Nhân tạo

<220>

<223> ex01-adhter-R

<400> 37

gcgtaattaa aatgattatc ttctttcatg agaaacacta aacttg

46

<210> 38

<211> 34

<212> ADN

<213> Nhân tạo

<220>

<223> pUC18-Pae1-F3

<400> 38

ctgcaggtcg actctagagg atccccgggt accg

34

<210> 39

<211> 35

<212> ADN

<213> Nhân tạo

<220>

<223> pUC18-Hind3-R3

<400> 39

gcttggcact ggccgtcggt ttacaacgtc gtagac

35

<210> 40

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-upstr-F

<400> 40

cggccagtgc caagcgcaga cttcaacagt tggctttttt aagta

45

<210> 41

<211> 47

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-upstr-R

<400> 41

cattttgcct ctcatgtttt taaatttggt ttgtagagta ttgaata

47

<210> 42

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> trpCpro-R

<400> 42

gaacagcaag ttaataatct agagggcgc

29

<210> 43

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> trpCter-F

<400> 43

atgagaggca aaatgaagcg tacaaagag

29

<210> 44

<211> 49

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-downstr-F

<400> 44

attaacttgc tgttcaatct tagaattcat tttttttttg tatcattcg

49

<210> 45

<211> 49

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-downstr-R

<400> 45

agagtcgacc tgcaggcgtc aataagagct tgaagggttg tgccggatc 49

<210> 46

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-upstr-F

<400> 46

cggccagtgc caagcttatg ggaaaaatgt gaaaatcatc ccact 45

<210> 47

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-upstr-R

<400> 47

cattttgcct ctcatgttta gttcaaataa tatttttttt tgtga

45

<210> 48

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-downstr-F

<400> 48

attaacttgc tgttcagttc tctttctgta ataatccttg atttc

45

<210> 49

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-downstr-R

<400> 49

agagtcgacc tgcagccttt ttacagaaaa caagcacatc ttta

45

<210> 50

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-inactive-F

<400> 50

taacgcatca gttccatcct gaaagtatcc

30

<210> 51

<211> 28

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-inactive-R

<400> 51

ggaactgatg cgtaaanaag aggggaaa

28

<210> 52

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> pdclter-F

<400> 52

aatcttagaa ttcatTTTTT ttttgtatc

29

<210> 53

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-adhlpro-F

<400> 53

attaacttgc tgttctagag ggaaaaagag agaattgaaa tagg

44

<210> 54

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> FT1-pdc1Ter-R

<400> 54

atgaattcta agattttaat agaaaccctg cttaaattgca agacc

45

<210> 55

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> PYC-pdc1Ter-R

<400> 55

atgaattcta agatttttagg cttcctcttt gacaaccttg gccac

45

<210> 56

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-upstr-F2

<400> 56

gcagacttca acagttggct tttttaagta

30

<210> 57

<211> 34

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-downstr-R2

<400> 57

gcgtaataa gagcttgaag gttggtgccg gatc

34

<210> 58

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1-up2

<400> 58

cattcccaca ggatttgtgc

20

<210> 59

<211> 25

<212> ADN

<213> Nhân tạo

<220>

<223> trpC(d)-1

<400> 59

gtgagatggt gatcatttgc acatg

25

<210> 60

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> pdc2-down1

<400> 60

agaagacaac ctagaccctc

20

<210> 61

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> trpC(d)-7

<400> 61

atagctcttg gtgggattcg

20

<210> 62

<211> 25

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1-down3

<400> 62

gctgctcgag atcctccaag gtatc

25

<210> 63

<211> 20

<212> ADN

<213> Nhân tạo

<220>

<223> trpC(d)-5

<400> 63

ttcgaccaag ggttcatac

20

<210> 64

<211> 34

<212> ADN

<213> Nhân tạo

<220>

<223> PDC1-downstr-R-P

<400> 64

gcgtcaataa gagcttgaag gttggtgccg gatac

34

<210> 65

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_750-F

<400> 65

tagattcaat ttgattggat aaagttcatc

30

<210> 66

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_750-R-P

<400> 66

ataacaaaca atggctaaaa gtggaccccc

30

<210> 67

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_500-F

<400> 67

tttactagtt taaagcaaaa aacatgagca

30

<210> 68

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_500-R-P

<400> 68

ttcaattttac atttctttat gaatatgcct

30

<210> 69

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_250-F

<400> 69

gtgttttatct gttcacaagt actggtaagc

30

<210> 70

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_250-R-P

<400> 70

ttcaatataa accaagtccc taaaagaaat

30

<210> 71

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_100-F

<400> 71

gctgatatga cattgcgacg aaaatagtat

30

<210> 72

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_100-R-P

<400> 72

aaaaaaaaag cttatatttca aaaatatgat

30

<210> 73

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_50-F

<400> 73

tcggttcaaaa aaaatcatta ttcaatactc

30

<210> 74

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_50-R-P

<400> 74

gtaatgaagt atagaacgaa tgatacaaaa

30

<210> 75

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_45-F

<400> 75

caaaaaaaaaat cattattcaa tactctacaa

30

<210> 76

<211> 32

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_45-R-P

<400> 76

gaagtataga acgaatgata caaaaaaaaaa at

32

<210> 77

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_40-F

<400> 77

aaaatcatta ttcaatactc tacaaaacaa

30

<210> 78

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_40-R-P

<400> 78

atagaacgaa tgatacaaaa aaaaaatgaa

30

<210> 79

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_35-F

<400> 79

cattattcaa tactctacaa aacaaattta

30

<210> 80

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_35-R-P

<400> 80

acgaatgata caaaaaaaaa atgaattcta

30

<210> 81

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_30-F

<400> 81

ttcaatactc tacaaaacaa atttaaaaac

30

<210> 82

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_30-R-P

<400> 82

tgatacaaaa aaaaaatgaa ttctaagatt

30

<210> 83

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_25-F

<400> 83

tactctacaa aacaaattta aaaacatgag

30

<210> 84

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_25-R-P

<400> 84

caaaaaaaaa atgaattcta agattgaaca

30

<210> 85

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_20-F

<400> 85

tacaaaacaa atttaaaaac atgagaggca

30

<210> 86

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_20-R-P

<400> 86

aaaaaatgaa ttctaagatt gaacagcaag

30

<210> 87

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_15-F

<400> 87

aacaaatttta aaaacatgag aggcaaaatg

30

<210> 88

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_15-R-P

<400> 88

atgaattcta agattgaaca gcaagttaat

30

<210> 89

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_10-F

<400> 89

atttaaaaac atgagaggca aaatgaagcg

30

<210> 90

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_10-R-P

<400> 90

ttctaagatt gaacagcaag ttaataatct

30

<210> 91

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_5-F

<400> 91

aaaacatgag aggcaaaatg aagcgtacaa

30

<210> 92

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> pdc1_5-R-P

<400> 92

agattgaaca gcaagttaat aatctagagg

30

<210> 93

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-upstr-F2

<400> 93

ttatgggaaa aatgtgaaaa tcatcccact

30

<210> 94

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> PDC2-downstr-R2-P

<400> 94

ccttttttaca gaaaacaagc acatcttttac

30

<210> 95

<211> 27

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-ki-F2-P

<400> 95

cttttcatga cccaacaaat cagacac

27

<210> 96

<211> 18

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của LeftTALEN-pdc1

<400> 96

tgcttgctat taaaatcg

18

<210> 97

<211> 18

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của RightTALEN-pdc1

<400> 97

ttgatttcct taagacgg

18

<210> 98

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của LeftTALEN-pdc2

<400> 98

tggtggttgct gggtcttat

19

<210> 99

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của RightTALEN-pdc2

<400> 99

tgccgacaat gtgaatcac

19

<210> 100

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của LeftTALEN-trpC

<400> 100

tgccaaggcg ccaatgtag

19

<210> 101

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của RightTALEN-trpC

<400> 101

tcggaaatgg tgattttgt

19

<210> 102

<211> 513

<212> ADN

<213> Rhizopus delemar

<220>

<223> FT1

<400> 102

atgtcttcta tcgaaacctc caaaatctca agttttatga tcaacaatat tgatgaacttt	60
gttgacagatt ggcatgttcg cgttacttgg atcgcatcca tgacactctg ggtcttttgg	120
ggattgggtt ggggttttccg caacttcttt gtttcaaact ctcttgcatc gaactctgct	180
cctgaagcaa atgctgctga tgatacggaa gcatcaaaga aaaaactctt tagcggtacc	240
agtgacagtt ttgctcttcg tctagatcgt gctcatcaag ttgtaaaaga tgccttggtc	300
tctcttctct gtcttctttc catgaactcc ttgctcgtg cttcaactcg tgctgtcatg	360
atcctcgctt gggtcttcac tgcctttgct gtctgttgg ttgctgtcgt attccttggt	420
gataaccgct ttgttcgttt gacgtattca cttgtctttt acgctcttgg tcttgctatt	480
gccggtcttg catttaagca gggtttctat taa	513

<210> 103

<211> 3540

<212> ADN

<213> Rhizopus delemar

<220>

<223> PYC

<400> 103

atgcctgctg caccagtacg tgaacactct gtggatacca ttcgtagaaa tagcgaagtg	60
atgggtaacc tgagaaaatt gatgggtggt aatcggtgtg aaattgctat ccgtgtcttt	120
cgtacagctc atgaactctc tatgaagaca gtagctatct tctctcatga agatagatta	180
tccatgcaca gatataaggc tgatgaatcc tatcaactcg gtcgtattgg tcaatacaca	240
cctgtaggtg cttatctggc acaagatgaa gtcgttcgaa tcgcaaagga acgtggtgtg	300
agcatgattc atcctgggta tggtttcttg tctgaaaatg ctgaattcgc tcgcaagggtg	360

gaagctgcag gaatcacttt cattgggtccc tctcctgatg tcattgaaag tttaggcgat	420
aagacaaaag ccagaacgat tgccatgaag tgtgaagtcc ctgttggtccc tggcacacct	480
ggacccgtca gtgaatacaa agatgccctg aactttatca aagaatatgg ttttcctatc	540
atcatcaagg ctgccatggg tgggtgggtgt cgtgggtatgc gtgtgggttcg tgacgaagcc	600
agtctagagg acgcgtttac ccgtgcgaaa tctgaagctt tggctgcctt tggatgatggc	660
actgtcttta tcgaacgttt ccttgataag cctcgtcata tcgaggttca attgttggca	720
gatcgtgcag gtaacgtggg ccatctcttt gaacgtgatt gctctgttca gcgtcgtcac	780
caaaaggctg ttgaaatcgc acctgccaaa aacttggata acaaggtagc tgaggccatc	840
ttgaacgatg cgatcaagat tgccaaggct gtaaagtaca agaatgcggg tactgcagaa	900
ttcttggtgg ataacaaaa ccgtcactac tttatcgaaa tcaatcctcg tatccaagtc	960
gaacatacca tcacagaaga aatcacgggt atcgatatcg ttgccgctca aattcagatt	1020
gctgccggtg ccctcttgcc tcaattgggt cttacccaac aacgtatccg tcaacgtggg	1080
ttcgcgatcc agtgctgtgt gacaaccgag gaccccgaaa agaatttcca gcctgacacg	1140
ggtaagatcg aagtttaccg ttccctctgt ggtaacggtg ttcgtctgga tgggtggtgct	1200
ggttacgcag gtgctatcat taccctcac tatgattcac ttttggtaa agtctcttgt	1260
tctggatcca cctacgaagt cgctcgtcga aagatcgtcc gtgccttggg cgaattcaga	1320
atccgtgggtg tcaagaccaa tatccccttc ttacaacgtc tcttgacca tgataccttc	1380
atcaacggta actgctggac aactttcatt gatgatactc ccgatctttt ccgtcttggt	1440
caattccaaa accgtgctca aagactcttg ggttacctgg gtgatgtcgt cgtcaatggt	1500
tctcaaatca aggggtcaaat gggatgatccc attctgaagc aagagatcga aatccccgtg	1560
ttgcgtgaaa gtggtagtga caagacggtc gatgtctctg ctctgctac ggaaggctgg	1620
agaaagatca ttgtggaaca aggacctgaa gctttcgcaa aagctgtccg tgcttaccct	1680
gggtgtcttga tcaccgatac cacctggaga gacgctcatc agagtatttatt ggccactcgt	1740
gtgagaactg tcgatctctt gcgtatcgcc cctgctacct ctcacgcttt ggccaacgcc	1800
ttttcattgg aatgttgggg aggtgctacg tttgatgttg ccatgcgttt cttcatgaa	1860
gatccttggg accgtcttgc tgctttgcga aagttggtac ccaatgtacc cttccaaatg	1920
cttttgctg gtgccaatgc ggtaggttac acctcttacc ctgataatgt tatctatgaa	1980
ttctgtgaca aggcagtcaa gtgtggtatg gatgtcttcc gtatctttga ctctctcaat	2040
tatgttgaga acatgagatt gggatttgac gctgtcaaga aggccggtgg tgttgttgaa	2100
gccaccatct gttacacggg tgatgtctcc aaccctaacc gcaagaagta cgacttgaag	2160

tactaccttg accttacaca atccttggtg aacgaaggta ttcacatctt gggatatcaag 2220
 gacatggctg gtcttctcaa acccgaggca gccaaagtac tggctctccag tatccgtgcc 2280
 aagttccccg acttgcccat ccacgttcac acacacgata ccgcaggtaac ggggtgttgct 2340
 agcatgatgg ccgctgccgc tgctgggtgct gacattgttg atgttgccgt ggacgccatg 2400
 tccggcatga cctctcaacc agcgatgggt gccattgtcg ctggactgga acagaccaat 2460
 ttgggtaccg gtatccgcat ggaagacatt catgccatca attcttactg ggagcaatgc 2520
 cgtttgcttt actcttgctt cgaagccaac gtgcgttcgg ccgattcggg tgtctatgaa 2580
 catgaaatgc ctggtggaca atataccaac ttgatgttcc aagcccaaca actcggtttg 2640
 ggaactcagt ggaagcaaact caagaaggct tacaaggagg ccaacgaact ctgtggtgac 2700
 ctggtcaagg tcacgccttc gtccaaggtc gtgggtgac ttgctcaatt catggtttcc 2760
 aaccaactct ctgcaaaga atttgaagaa cgcgcctcga gtctctctct gccacctct 2820
 gtcacgcagt tcttccaagg ttatctcggc caaccctatg gcggtttccc cgagcccttg 2880
 cgctccaaca tccttcgtga tctccctcgc ctcgacggtc gccctggtgc tagcctgcct 2940
 ccgttggaca tggctaaaact caaggaagag ttggttgaaa agtacggttc gagcatccgt 3000
 gattacgacg tgatctcggc tgctctttac cccaaggctt ttgccgacta ccgtgatacc 3060
 gtcagtcaat acggtgatct ctccgttttg cctacacgct actttttgtc caagcccag 3120
 atcaatgaag aattccatgt ggagattgaa gaaggaaaga cgttgatcat caagttattg 3180
 gccgtcggtc ctctgaacaa tgacggtaaa cgtgatgttt actttgaatt gaacggtgaa 3240
 gctcgtgtgg tgggcattgt ggatcgcaat tctgctattg aaatcgtcac acgtgaaaag 3300
 gccaacccct ctaaccccgg tgacattggt gtcctatgt cgggtgtggt tgtcgagatc 3360
 cgtgccaaagg aaggtagcca tgtcaaggcc ggtgatcctc ttgctgttct ctctgctatg 3420
 aagatggaaa cagtggtcac tgtccccgtg gctggtagag ttgagcgtgt tgctatccaa 3480
 gaaggtgatt cattatccgc tggtgatttg gtggccaagg ttgtcaaaga ggaagcctaa 3540

<210> 104

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> trpCter-R

<400> 104

atgagaggca aaatgaagcg tacaaagag

29

<210> 105

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> trpCter-cicCpro-F

<400> 105

cattttgcct ctcatcttac gcaggttgat agtagccgcc

40

<210> 106

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> cipCpro-adhter-R

<400> 106

gcgtaattaa aatgaggtta gagtatgaag aaaaaaaaaa

40

<210> 107

<211> 31

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-lost-F2

<400> 107

tttaaattctt acgcagggttg atagtagccg c 31

<210> 108

<211> 33

<212> ADN

<213> Nhân tạo

<220>

<223> trpC-lost-R2

<400> 108

tgcgtaagat ttaaatgaat tcgagctcgg tac 33

<210> 109

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> cipCpro-R

<400> 109

ggttagagta tgaagaaaaa aaaaaaacg 29

<210> 110

<211> 45

<212> ADN

<213> Nhân tạo

<220>

<223> cipCpro-LifeTALEN-F

<400> 110

cttcatactc taaccatggg aaaacctatt cctaatcctc tgctg

45

<210> 111

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> cipCpro-exo1-F

<400> 111

cttcatactc taaccatgaa aatccaagtt gcttctccta

40

<210> 112

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-upstr-F

<400> 112

cggccagtgc caagcccgtc aggggtgaat gagatatatt

40

<210> 113

<211> 33

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-upstr-R2

<400> 113

aaaagatgtg agttataaaa ggatgatgca agc

33

<210> 114

<211> 33

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-downstr-F2

<400> 114

taactcacat cttttattct ttttctatcc ctc

33

<210> 115

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-downstr-R

<400> 115

agagtcgacc tgcagacctg ttagaaaggt acatgcattc

40

<210> 116

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-upstr-R

<400> 116

cattttgcct ctcatgtgag ttataaaaagg atgatgcaag

40

<210> 117

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-downstr-F

<400> 117

attaacttgc tgttcattctt ttattctttt tctatccctc

40

<210> 118

<211> 23

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-upstr-F2

<400> 118

ccgtcagggg tgaatgagat att

23

<210> 119

<211> 25

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-downstr-R2-P

<400> 119

acctgttaga aaggtacatg cattc

25

<210> 120

<211> 21

<212> ADN

<213> Nhân tạo

<220>

<223> pdc3-up

<400> 120

gacctcaatc actatccttg g

21

<210> 121

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> pUC18-adhlpro-F

<400> 121

cggccagtg c caagcattat tattagaggg aaaaaaaga aaga

44

<210> 122

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> adhlpro-pUC18-R

<400> 122

agagtcgacc tgcagttttg ttattttattt gtattaattg ataa

44

<210> 123

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> adhlpro-adhlter-F

<400> 123

aacaaaatca ttttaattac gcattttcat tttttactaa

40

<210> 124

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> adh1ter-pUC18-R

<400> 124

agagtcgacc tgcagagcag aacatgagtc tggaagcgag acac

44

<210> 125

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> adh1pro-adh1ter-R

<400> 125

taaaatgatt ttgttattta tttgtattaa ttgataatga

40

<210> 126

<211> 31

<212> ADN

<213> Nhân tạo

<220>

<223> adh1pro(o)-R

<400> 126

ttttgttatt tatttgtatt aattgataat g 31

<210> 127

<211> 30

<212> ADN

<213> Nhân tạo

<220>

<223> adh1ter(o)-F

<400> 127

tcatttttaac tacgcatttt catttttttac 30

<210> 128

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> adh1pro(o)-LifeTALEN-F

<400> 128

aaataaataa caaaaatggg aaaacctatt cctaatacctc tgct 44

<210> 129

<211> 44

<212> ADN

<213> Nhân tạo

<220>

<223> LifeTALEN-adhlter(o)-R

<400> 129

gcgtaattaa aatgatcaga agttgatctc gccgttggtg aact

44

<210> 130

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-upstr-F

<400> 130

cggccagtg c caagcaaaag aataagaaaa gatgtgtcag

40

<210> 131

<211> 32

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-upstr-R2

<400> 131

taaattaagg acttgagctt gaacgatgtc ag

32

<210> 132

<211> 33

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-downstr-F2

<400> 132

caagtcotta atttacaaat aataaatcat gtt

33

<210> 133

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-downstr-R

<400> 133

agagtcgacc tgcagaacga caaacatggc tatcaaggga

40

<210> 134

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-upstr-R

<400> 134

tgtgcaagtg aacaaaggac ttgagcttga acgatgtcag

40

<210> 135

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-downstr-F

<400> 135

gctcttgaca atgcataatt tacaaataat aaatcatgtt

40

<210> 136

<211> 25

<212> ADN

<213> Nhân tạo

<220>

<223> adelpro-R

<400> 136

tgcattgtca agagcgtgtc gcaac

25

<210> 137

<211> 29

<212> ADN

<213> Nhân tạo

<220>

<223> adelter-F

<400> 137

ttgttcactt gcacagcgtg atatgcaag

29

<210> 138

<211> 28

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-upstr-F2

<400> 138

aaaagaataa gaaaagatgt gtcaggac

28

<210> 139

<211> 27

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-downstr-R2-P

<400> 139

aacgacaaac atggctatca aggggaac

27

<210> 140

<211> 22

<212> ADN

<213> Nhân tạo

<220>

<223> ldhA-up

<400> 140

gaaactacag tattccctcg tg

22

<210> 141

<211> 21

<212> ADN

<213> Nhân tạo

<220>

<223> adel-15

<400> 141

tgcttccata tgtcaatagg c

21

<210> 142

<211> 40

<212> ADN

<213> Nhân tạo

<220>

<223> exol-adhter-R2

<400> 142

gcgtaattaa aatgattatc ttctttcatg agaaacacta

40

<210> 143

<211> 1749

<212> ADN

<213> Rhizopus delemar

<220>

<223> pdc3

<400> 143

atggttagta tcaaaattgg agattatctc attcaacgtc ttaaagaaac cggaatcgac	60
acgatttttg gtgttccagg tgattacaat atggtaagtt acgaagtgga tacagattta	120
tttgaatgtg tgtatatcct gaaaaacttt gcaatgtagc cgctattgga tttaattgaa	180
gatgattctg agctcatatg gggtaataat gcaaacgaac tcaatgcttc ctatgctgct	240
gatggctatg ctcgatcag aggccttggg gctgtagtta ccacgtttgg tgtaggggaa	300
ctttctgctg cagcagggat tgctggatct tattctgaaa aagtcccagt acttcatatt	360
gtcggcactc ctaatactaa atcccaggaa gctggagcta ttcttcatca cacgcttggg	420
aatggcaatt ttcagggtgt tggtgaaatg ttttccatga tcacgtgtgc ttctactcat	480
ttaaactttg acaatgccat ccgagaaaatt gatcgcgtca tccagcagac tatgattcga	540
aagcgacctg gttatatcgg catccccatt gatctgatta atgctgaggt tgctcttctc	600
agttctgaac ctctcaattt ttttgtcccc aaaaatccta ctcaaactca agatgtggct	660
cttaagggtg ttttggacgc tatttcacag gctaagcatc cgattatagt tgttgatgca	720
tgtgttcagc gacacaattt agttcaagag gctattgaat tcgtgaaacg taccggtttt	780
cccacttacg ttgcacctat gggtaaagggt attgttcccg aggatcttgt caattatcgt	840
ggttgctatg ctggtaatat taccatcgag ggtatcgcca gggaacttga gcaagctgat	900
ctggtcattg aacttgggtgc gattaaatcc gatttcaaca ctggtggctt cacctataaa	960
ctggatcctg ctagaacgat ctctcttcac tcatttggta ccagatatt ttatgccgac	1020
tatgacaaag tgggaatgac ggaatttctt cctcttttga ccaagtctct tcctcaaagg	1080
cctcgtgtat ttgatctagg cctcgtcat gagccagatc caattcaatc aggaactgaa	1140
ataacgcaca attatttctg gaataaggta ccagaataca tggatcctcg cgctgttgtt	1200

gttgctgaaa caggcacagc tgaatttgcc tcttttaatt taagagctcc caaagacgct 1260
 ctctttatatt ctcaagtact ctggggatcc atcggttttg ctgtcggatg tgctgtcggg 1320
 gctgcgttcg ctgatcgaga tcgacgagtt tatctttttg tgggtgatgg ctctttccag 1380
 gttacttgtc aagaaatttc agtctttttg catcaagggt tgacacctgt cattttcttg 1440
 ttgaacaatg atggatacct tatcgaaaaa ctcatlcatg ggccctcaccg ctcttataat 1500
 aactttcaga tgtggaacta cagcaaaact cttgactata tgggtggaca tcttcagcgc 1560
 aatctgtctg atgtttcgcc agctcaagtt ggtgttgaag ctcaagtccg tacgcgagat 1620
 gagtttgaga gggcaatgaa gacagtcaag gaggaacgta acaaaattca ttttattgaa 1680
 gttgtcatgc ctcaatttga tgctcctcgc gaattgatac tccaagttca aacttctgaa 1740
 aatcgtaa 1749

<210> 144

<211> 3075

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> LeftTALEN-pdc3

<400> 144

atgggaaaac ctattcctaa tcctctgctg ggccctggatt ctaccggagg catggcccct 60
 aagaaaaagc ggaagggtga cggcggagtg gacctgagaa cactgggata ttctcagcag 120
 cagcaggaga agatcaagcc caaggtgaga tctacagtgg cccagcacca cgaagccctg 180
 gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcaccg tgccgcctg 240
 ggaacagtgg ccgtgaaata tcaggatatg attgccgccc tgccctgaggc cacacacgaa 300
 gccattgtgg gagtgggaaa acgaggcgct ggagccagag ccctggaagc cctgctgaca 360
 gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt 420
 gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca 480
 ggagcccctc tgaacctgac ccccgaaacag gtggtggcca ttgccagcca cgacggcggc 540
 aagcaggccc tggaaaccgt gcagagactg ctgcccgtgc tgtgccaggc ccatggcctg 600

acacctgaac aggtggtggc tatcgctctt aacaacggag gaaaacaggc tctggaaaca	660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg	720
gctattgctt ccaacaacgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca	780
gtgctgtgcc aggtcacgg actgaccccc gaacaggtgg tggccattgc cagcaacatc	840
ggcggcaagc aggccctgga aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat	900
ggcctgacac ctgaacaggt ggtggctatc gcctctaata tcggaggaaa acaggctctg	960
gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag	1020
gtggtggcta ttgcttccaa cggcgggggg aaacaggccc tggaaactgt gcagcgctg	1080
ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtggtggc cattgccagc	1140
cacgacggcg gcaagcaggc cctggaaaacc gtgcagagac tgctgcccgt gctgtgccag	1200
gcccattggc tgacacctga acaggtggtg gctatcgctt ctaacaacgg aggaaaacag	1260
gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggtcacgg cttgactcca	1320
gaacaggtgg tggctattgc ttccaatatt ggggggaaac aggccctgga aactgtgcag	1380
cgctgctgc cagtgtgtg ccaggctcac ggctcactc ccgaacaggt ggtggccatt	1440
gccagccacg acggcggcaa gcaggccctg gaaaccgtgc agagactgct gcccgctg	1500
tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctaa tatcggagga	1560
aaacaggctc tggaaacagt gcagcggtg ctgcctgtgc tgtgtcaggc tcacggcttg	1620
actccagaac aggtggtggc tattgcttcc cacgacgggg ggaaacaggc cctggaaact	1680
gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acaggtggtg	1740
gccattgcca gcaacaacgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc	1800
gtgctgtgcc aggcccatgg cctgacacct gaacaggtgg tggctatcgc ctctaatac	1860
ggaggaaaac aagcactcga gacagtgcag cggctgctgc ctgtgctgtg tcaggctcac	1920
ggcttgactc cagaacaggt ggtggctatt gcttccaacg gcggggggaa acaggccctg	1980
gaaactgtgc agcgctgct gccagtgtg tgccaggctc acgggctgac cccggaacag	2040
gtggtggcca ttgccagcaa cggcggcggc aagcaggccc tggaaaccgt gcagagactg	2100
ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtggtggc tatcgctctt	2160
aacggcggag gaaaacaggc tctggaaaca gtgcagcggc tgctgcctgt gctgtgtcag	2220
gctcacggct tgactccaca gcaggtcgtg gcaattgcta gcaacggcgg cggacggccc	2280
gccctggaga gcattgtggc ccagctgtct agacctgatc ctgccctggc cgccctgaca	2340
aatgatcacc tgggtggccct ggctgtctg ggaggcagac ctgccctgga tgccgtgaaa	2400

aaaggactgc ctcacgcccc tgccctgatac aagagaacaa atagaagaat ccccgagcgg 2460
 acctctcaca gagtggccgg atcccagctg gtgaaatctg agctggagga gaagaagtct 2520
 gagctgagac acaagctgaa gtacgtgcct caccagtaga tcgagctgat cgagatcgcc 2580
 agaaatagca cccaggatag aatcctggag atgaaggtga tggagttctt catgaaggtg 2640
 tacggctaca gaggaaagca cctgggagga agcagaaaaac ctgacggagc catttataca 2700
 gtgggcagcc ctatcgatta tggcgtgatac gtggatacaa aggcctacag cggaggctac 2760
 aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg 2820
 aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc 2880
 aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggcccagct gaccaggctg 2940
 aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga 3000
 gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagt caacaacggc 3060
 gagatcaact tctga 3075

<210> 145

<211> 3075

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> RightTALEN-pdc3

<400> 145

atgggaaaaac ctattcctaa tcctctgctg ggctggatt ctaccggagg catggcccct 60
 aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag 120
 cagcaggaga agatcaagcc caaggtgaga tctacagtgg cccagcacca cgaagccctg 180
 gtgggacacg gatttacaca cgcccacatt gtggccctgt ctgagcacc tgccgcctg 240
 ggaacagtgg ccgtgaaata tcaggatatg attgccgcc tgctgaggc cacacacgaa 300
 gccattgtgg gagtgggaaa acgaggcgct ggagccagag ccctggaagc cctgctgaca 360
 gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt 420
 gccaaaaggg gcggagtgc cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca 480

ggagcccctc tgaacctgac cccgaacag gtggtggcca ttgccagcaa caacggcggc	540
aagcaggccc tggaaaccgt gcagagactg ctgccgtgc tgtgccaggc ccatggcctg	600
acacctgaac aggtggtggc tatcgctct aacggcggag gaaaacaggc tctggaaaca	660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg	720
gctattgctt ccaatattgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcc	780
gtgctgtgcc aggtcacgg actgaccccc gaacaggtgg tggccattgc cagcaacatc	840
ggcggcaagc aggccctgga aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat	900
ggcctgacac ctgaacaggt ggtggctatc gcctctcacg acggaggaaa acaggtctg	960
gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag	1020
gtggtggcta ttgcttccaa cggcgggggg aaacaggccc tggaaactgt gcagcgctg	1080
ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtggtggc cattgccagc	1140
aacggcggcg gcaagcaggc cctggaaacc gtgcagagac tgctgccgt gctgtgccag	1200
gcccattggc tgacacctga acaggtggtg gctatcgct ctaatatcgg aggaaaacag	1260
gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggtcacgg cttgactcca	1320
gaacaggtgg tggctattgc ttcccacgac ggggggaaac aggccctgga aactgtgcag	1380
cgctgctgc cagtgtgtg ccaggctcac ggctcactc ccgaacaggt ggtggccatt	1440
gccagccacg acggcggcaa gcaggccctg gaaaccgtgc agagactgct gccctgctg	1500
tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctaa tatcggagga	1560
aaacaggctc tggaaacagt gcagcggctg ctgcctgtgc tgtgtcaggc tcacggcttg	1620
actccagaac aggtggtggc tattgcttcc aacggcgggg ggaaacaggc cctggaaact	1680
gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acaggtggtg	1740
gccattgcca gcaacatcgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc	1800
gtgctgtgcc aggcccatgg cctgacacct gaacaggtgg tggctatcgc ctctaacggc	1860
ggaggaaaaa aagcactcga gacagtgcag cggctgctgc ctgtgctgtg tcaggctcac	1920
ggcttgactc cagaacaggt ggtggctatt gcttccaacg gcggggggaa acaggccctg	1980
gaaactgtgc agcgctgct gccagtgtg tgccaggctc acgggctgac cccgaacag	2040
gtggtggcca ttgccagcaa caacggcggc aagcaggccc tggaaaccgt gcagagactg	2100
ctgccgtgc tgtgccaggc ccatggcctg acacctgaac aggtggtggc tatcgctct	2160
aacggcggag gaaaacaggc tctggaaaca gtgcagcggc tgctgcctgt gctgtgtcag	2220
gctcacggct tgactccaca gcaggtcgtg gcaattgcta gcaatatcgg cggacggccc	2280

gccctggaga gcattgtggc ccagctgtct agacctgac ctgccctggc cgccctgaca 2340
 aatgatcacc tgggtggccct ggccctgtctg ggaggcagac ctgccctgga tgccgtgaaa 2400
 aaaggactgc ctacgcccc tgccctgac aagagaacaa atagaagaat ccccgagcgg 2460
 acctctcaca gagtggccgg atcccagctg gtgaaatctg agctggagga gaagaagtct 2520
 gagctgagac acaagctgaa gtacgtgcct caccagtaga tcgagctgat cgagatcgcc 2580
 agaaatagca cccaggatag aatcctggag atgaagggtga tggagttctt catgaagggtg 2640
 tacggctaca gaggaaagca cctgggagga agcagaaaaac ctgacggagc catttataca 2700
 gtgggcagcc ctatcgatta tggcgtgac gtggatacaa aggcctacag cggaggctac 2760
 aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg 2820
 aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc 2880
 aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggcccagct gaccaggctg 2940
 aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga 3000
 gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagt caacaacggc 3060
 gagatcaact tctga 3075

<210> 146

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của LeftTALEN-pdc3

<400> 146

ccggaatcga caccgtttt

19

<210> 147

<211> 19

<212> ADN

<213> Rhizopus delemar

<220>

<223> Đích của RightTALEN-pdc3

<400> 147

cgtaacttac catattgta

19

<210> 148

<211> 963

<212> ADN

<213> Rhizopus oryzae

<220>

<223> ldhA

<400> 148

atggtattac actcaaaggt cgccatcggt ggagctgggtg cagtaggagc ctccactgct	60
tatgcactta tgtttaaaaa catttggtaca gaaatcatta ttgttgatgt taatcctgac	120
atcggttcaag ctcaagtcct tgaccttgca gatgctgccca gtataagtca cacgccccatc	180
cgagcaggta gcgcagagga ggcagggcag gcagatattg ttgtcatcac ggccgggtgcg	240
aaacaaaggg aaggtgagcc tcggacaaaag ctcatggaac gaaacttcag agtggtgcaa	300
agtatcattg gtggcatgca acccattcga ccagacgcag tcactcttgg gtagcaaat	360
ccagtcgata tcttgacaca cattgcaaag accctctctg gactgcctcc aaaccaggtc	420
attggctccg gtacctacct tgacacgacc cgtcttcgcg tccatcttgg cgatgtcttt	480
gatgtcaatc ctcaatcggt ccatgctttt gtcttgggtg aacatgggga ttcccagatg	540
atcgcttggg aggctgcttc gattgggtggc cagccgttga caagtttccc ggaattcgca	600
aagctggata aaacagcaat ttcaaaaagcg atatcaggta aagcgatgga gatcattcgt	660
ttgaaaggag ccacgtttta tggaaattgg gcctgtgcag cggatttagt gcacactatc	720
atgttgaata ggaaatcagt acatccagtt tctgtttatg ttgaaaagta tggagccact	780
ttttctatgc ctgctaaaact tggatggaga ggtgttgaac agatctatga agtaccactg	840

acggaagaag aagaagcggtt gcttgtaaaa tctgtagagg cattgaaatc agttgaatat 900
 tcatctacaa aagttccaga aaaaaagggtt catgctactt ccttttctaa aagtagctgt 960
 tga 963

<210> 149

<211> 3075

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> LeftTALEN-ldhA

<400> 149

atgggaaaaac ctattcctaa tcctctgctg ggcttggatt ctaccggagg catggccctt 60
 aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag 120
 cagcaggaga agatcaagcc caaggtgaga tctacagtgg ccagcacca cgaagccctg 180
 gtgggacacg gatttacaca cgccacatt gtggccctgt ctcagcacc tgccgccctg 240
 ggaacagtgg ccgtgaaata tcaggatatg attgccgcc tgctgaggc cacacacgaa 300
 gccattgtgg gagtgggaaa acagtggctt ggagccagag ccttggaagc cctgctgaca 360
 gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt 420
 gccaaaaggg gcggagtgac cgcggtggaa gccgtgcacg cctggagaaa tgccctgaca 480
 ggagcccctc tgaacctgac ccccgaaacag gtgggtggcca ttgccagcaa caacggcggc 540
 aagcaggccc tggaaaccgt gcagagactg ctgcccgtgc tgtgccaggc ccatggcctg 600
 acacctgaac aggtggtggc tatcgctct cagcagggag gaaaacaggc tctggaaaca 660
 gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg 720
 gctattgctt ccaatattgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca 780
 gtgctgtgcc aggtcacgg actgaccccc gaacaggtgg tggccattgc cagcaacaac 840
 ggccggcaagc aggccctgga aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat 900
 ggctgacac ctgaacaggt ggtggctatc gcctctaata tcggaggaaa acaggctctg 960
 gaaacagtgc agcggctgct gcctgtgctg tgtcaggctc acggcttgac tccagaacag 1020

gtggtggcta ttgcttccaa cggcgggggg aaacaggccc tggaaactgt gcagcgctg 1080
 ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtggtggc cattgccagc 1140
 aacaacggcg gcaagcaggc cctggaaacc gtgcagagac tgctgcccgt gctgtgccag 1200
 gcccatggcc tgacacctga acaggtggtg gctatcgctt ctcacgacgg aggaaaacag 1260
 gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggtcacgg cttgactcca 1320
 gaacaggtgg tggctattgc ttccaacggc ggggggaaac aggccctgga aactgtgcag 1380
 cgctgctgc cagtgtgtg ccaggctcac ggctcactc ccgaacagg ggtggccatt 1440
 gccagcaaca acggcggcaa gcaggccctg gaaaccgtgc agagactgct gccctgctg 1500
 tgccaggccc atggcctgac acctgaacag gtggtggcta tcgcctctca cgacggagga 1560
 aaacaggctc tggaaacagt gcagcggctg ctgcctgtgc tgtgtcaggc tcacggcttg 1620
 actccagaac aggtggtggc tattgcttcc cacgacgggg gaaaacaggc cctggaaact 1680
 gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acaggtggtg 1740
 gccattgcca gcaacatcgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc 1800
 gtgctgtgcc agggccatgg cctgacacct gaacaggtgg tggctatcgc ctctaacaac 1860
 ggaggaaaac aagcactcga gacagtgcag cggctgctgc ctgtgctgtg tcaggctcac 1920
 ggcttgactc cagaacagg ggtggctatt gcttccaacg gcggggggaa acaggccctg 1980
 gaaactgtgc agcgctgct gccagtgtg tgccaggctc acgggctgac cccgaacag 2040
 gtggtggcca ttgccagcaa catcggcggc aagcaggccc tggaaaccgt gcagagactg 2100
 ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtggtggc tatcgctct 2160
 aacggcggag gaaaacaggc tctggaaaca gtgcagcggc tgctgcctgt gctgtgtcag 2220
 gctcacggct tgactccaca gcaggctgtg gcaattgcta gcaatatcgg cggacggccc 2280
 gccctggaga gcattgtggc ccagctgtct agacctgatc ctgccctggc cgccctgaca 2340
 aatgatcacc tgggtggcct ggctgtctg ggaggcagac ctgccctgga tgccgtgaaa 2400
 aaaggactgc ctcacgcccc tgccctgatc aagagaacaa atagaagaat cccgagcgg 2460
 acctctcaca gagtggcgg atcccagctg gtgaaatctg agctggagga gaagaagtct 2520
 gagctgagac acaagctgaa gtacgtgcct cacgagtaca tcgagctgat cgagatcgcc 2580
 agaaatagca cccaggatag aatcctggag atgaagggtg tggagttctt catgaagggtg 2640
 tacggctaca gaggaagca cctgggagga agcagaaaaac ctgacggagc catttataca 2700
 gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac 2760
 aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg 2820

aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc 2880
aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggcccagct gaccaggctg 2940
aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga 3000
gagatgatta aggccggaac actgacactg gaggaggtga gaagaaagtt caacaacggc 3060
gagatcaact tctga 3075

<210> 150

<211> 3075

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> RightTALEN-ldhA

<400> 150

atgggaaaaac ctattcctaa tcctctgctg ggcttggatt ctaccggagg catggccctt 60
aagaaaaagc ggaaggtgga cggcggagtg gacctgagaa cactgggata ttctcagcag 120
cagcaggaga agatcaagcc caaggtgaga tctacagtgg ccagcacca cgaagccctg 180
gtgggacacg gatttacaca cgcccacatt gtggccctgt ctcagcacc tgcgcctt 240
ggaacagtgg ccgtgaaata tcaggatatg attgccgccc tgcttgaggc cacacacgaa 300
gccattgtgg gagtgggaaa acagtggctt ggagccagag ccctggaagc cctgctgaca 360
gtggccggag aactgagagg acctcctctg cagctggata caggacagct gctgaagatt 420
gccaaaagg gcggtgtgac cgcgttgga gccgtgcacg cctggagaaa tgccctgaca 480
ggagcccctc tgaacctgac ccccgaaacag gtggtggcca ttgccagcca cgacggcggc 540
aagcaggccc tggaaaccgt gcagagactg ctgcccgtgc tgtgccaggc ccatggcctg 600
acacctgaac aggtggtggc tatcgctct cagcagggag gaaaacaggc tctggaaaca 660
gtgcagcggc tgctgcctgt gctgtgtcag gctcacggct tgactccaga acaggtggtg 720
gctattgctt ccaacggcgg ggggaaacag gccctggaaa ctgtgcagcg cctgctgcca 780
gtgctgtgcc aggtcacgg actgaccccc gaacaggtgg tggccattgc cagccacgac 840
ggcggcaagc aggccttgg aaccgtgcag agactgctgc ccgtgctgtg ccaggcccat 900

ggcctgacac ctgaacaggt ggtggctatc gcctctaacg gcggaggaaa acaggctctg	960
gaaacagtgc agcggtctgt gcctgtgctg tgtcaggctc acggcttgac tccagaacag	1020
gtggtggcta ttgcttccaa caacgggggg aaacaggccc tggaaactgt gcagcgctg	1080
ctgccagtgc tgtgccaggc tcacgggctg acccccgaac aggtgggtggc cattgccagc	1140
cacgacggcg gcaagcaggc cctggaaaacc gtgcagagac tgctgcccgt gctgtgccag	1200
gcccattggcc tgacacctga acagggtgtg gctatcgctt ctaacaacgg agggaaaacag	1260
gctctggaaa cagtgcagcg gctgctgcct gtgctgtgtc aggtcacgg cttgactcca	1320
gaacagggtg tggctattgc ttcccacgac ggggggaaac aggccctgga aactgtgcag	1380
cgctgtctgc cagtgtgtg ccaggctcac ggctcactc ccgaacaggt ggtggccatt	1440
gccagcaacg gcggcgggcaa gcaggccctg gaaaccgtgc agagactgct gcccgctgctg	1500
tgccaggccc atggcctgac acctgaacag gtggtggcta tcgctctaa tatcggagga	1560
aaacaggctc tggaaacagt gcagcggtg ctgctgtgc tgtgtcaggc tcacggcttg	1620
actccagaac aggtgggtggc tattgcttcc cacgacgggg ggaaacaggc cctggaaact	1680
gtgcagcgcc tgctgccagt gctgtgccag gctcacggac tgacccccga acagggtgtg	1740
gccattgcca gccacgacgg cggcaagcag gccctggaaa ccgtgcagag actgctgccc	1800
gtgctgtgcc agggccatgg cctgacacct gaacagggtg tggctatcgc ctctaacggc	1860
ggaggaaaaac aagcactcga gacagtgcag cggctgtgc ctgtgctgtg tcaggctcac	1920
ggcttgactc cagaacaggt ggtggctatt gcttccaaca acggggggaa acaggccctg	1980
gaaactgtgc agcgctgtg gccagtgtg tgccaggctc acgggctgac cccgaacag	2040
gtggtggcca ttgccagcca cgacggcggc aagcaggccc tggaaaccgt gcagagactg	2100
ctgcccgtgc tgtgccaggc ccatggcctg acacctgaac aggtgggtggc tatcgctct	2160
aacggcggag gaaaacaggc tctggaaaca gtgcagcggc tgctgcctgt gctgtgtcag	2220
gctcacggct tgactccaca gcaggctgtg gcaattgcta gccacgacgg cggacggccc	2280
gccctggaga gcattgtggc ccagctgtct agacctgatc ctgccctggc cgccctgaca	2340
aatgatcacc tgggtggccct ggctgtctg ggaggcagac ctgccctgga tgccgtgaaa	2400
aaaggactgc ctcacgcccc tgccctgatc aagagaacaa atagaagaat ccccgagcgg	2460
acctctcaca gagtggccgg atcccagctg gtgaaatctg agctggagga gaagaagtct	2520
gagctgagac acaagctgaa gtacgtgcct cacgagtaca tcgagctgat cgagatcgcc	2580
agaaaatagca cccaggatag aatcctggag atgaaggatg tggagtctct catgaagggtg	2640
tacggctaca gaggaaagca cctgggagga agcagaaaac ctgacggagc catttataca	2700

gtgggcagcc ctatcgatta tggcgtgatc gtggatacaa aggcctacag cggaggctac	2760
aatctgccta ttggacaggc cgatgagatg cagagatacg tggaggagaa ccagaccagg	2820
aacaagcaca tcaaccctaa cgagtgggtg aaggtgtacc cttctagcgt gaccgagttc	2880
aagttcctgt ttgtgagcgg ccacttcaag ggcaattata aggccagct gaccaggctg	2940
aaccacatca caaattgtaa tggcgccgtg ctgtctgtgg aggaactgct gattggagga	3000
gagatgatta aggcgggaac actgacactg gaggagggtga gaagaaagtt caacaacggc	3060
gagatcaact tctga	3075

<210> 151

<211> 19

<212> ADN

<213> Rhizopus oryzae

<220>

<223> Đích của LeftTALEN-ldhA

<400> 151

tgcagatgct gccagtata	19
----------------------	----

<210> 152

<211> 19

<212> ADN

<213> Rhizopus oryzae

<220>

<223> Đích của RightTALEN-ldhA

<400> 152

tcctctgcgc tacctgctc	19
----------------------	----