



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

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

(11)



1-0030944

(51)⁷ A01H 5/00; C12N 15/82

(13) B

(21) 1-2015-02230

(22) 21/11/2014

(86) PCT/US2014/066772 21/11/2014

(87) WO 2015/077538 A2 28/05/2015

(30) 61/908,444 25/11/2013 US

(45) 25/02/2022 407

(43) 26/09/2016 342A

(73) BANGLADESH JUTE RESEARCH INSTITUTE (BD)

Manik Mia Avenue, Dhaka, HI 1207 (BD) Bangladesh

(72) ISLAM, Mohammed, Shahidul (BD); AHMED, Borhan (BD); HAQUE, Mohammed, Samiul (BD); ALAM, Mohammed, Monjurul (BD); ALAM, Maqsubul (BD).

(74) Công ty TNHH Trần Hữu Nam và Đồng sự (TRAN H.N & ASS.)

(54) POLYNUCLEOTIT MÃ HÓA PROTEIN KHÓA KÉO HOMEBOX-LOXIN HAT22 (PROTEIN HD-ZIP 22) CÓ NGUỒN GỐC TỪ HAI LOÀI CÂY CORCHORUS OLITORIUS VÀ CORCHORUS CAPSULARIS VÀ PHƯƠNG PHÁP SẢN XUẤT CÂY TRỒNG CHUYỂN GEN BIỂU HIỆN POLYNUCLEOTIT NÀY

(57) Sáng chế đề cập đến polynucleotit phân lập được mã hóa protein khóa kéo homeobox-loxin HAT22 (protein HD-ZIP 22) từ cây thuộc loài *Corchorus olitorius* và *Corchorus capsularis* và polypeptit tương ứng có nguồn gốc từ đó. Sáng chế cũng đề cập đến cây trồng có sự biểu hiện được điều biến của axit nucleic mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22 hoặc chất tương đồng của chúng, mà có khả năng để cải biến, tốt hơn là làm tăng/làm tăng cường chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng. Cụ thể hơn, sáng chế đề cập đến polypeptit có hoạt tính protein khóa kéo homeobox-loxin HAT22, polynucleotit mã hóa cho polypeptit này, và phương pháp tạo ra và phương pháp sử dụng các polynucleotit và polypeptit này. Sáng chế còn đề xuất vector, cấu trúc biểu hiện và tế bào chủ bao gồm và/hoặc có chứa các trình tự nucleotit của protein khóa kéo homeobox-loxin HAT22 (protein HD-ZIP 22). Sáng chế cũng đề xuất phương pháp sản xuất protein này và phương pháp cải biến protein này để cải thiện các đặc tính mong muốn. Protein theo sáng chế có thể được sử dụng theo nhiều cách, bao gồm việc làm tăng/làm cải thiện chiều dài sợi, chiều cao và sinh khối của cây trồng và năng suất sợi.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực sinh học phân tử, và phương pháp làm tăng hoặc tăng cường độ dài sợi, chiều cao và/hoặc sinh khối của cây trồng bằng cách điều biến sự biểu hiện của axit nucleic mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22. Cụ thể hơn, sáng chế đề xuất chất tương đồng protein khóa kéo homeobox-loxin HAT22 được phân lập từ hai loài cây đay, cụ thể là, *Corchorus olitorius* ("*C. olitorius*") và *Corchorus capsularis* ("*C. capsularis*"), và ứng dụng của chúng để làm tăng độ dài sợi, chiều cao cây và/hoặc sinh khối cây ở cây đay bằng cách điều chỉnh sự biểu hiện của gen mã hóa cho protein khóa kéo homeobox-loxin HAT22.

Tình trạng kỹ thuật của sáng chế

Cây đay được trồng làm thương phẩm chủ yếu được sản xuất từ hai loài, cụ thể là, *Corchorus olitorius* (thường gọi là "cây đay tossa") và *Corchorus capsularis* (thường gọi là "cây đay trắng"). Cả hai loài này đều là lưỡng bội ($2n = 14$), và chỉ hai loài này được dùng trong nông nghiệp. Hai loài này là cây hoa lợi quan trọng của các nước Đông Nam Á và Braxin, cung cấp sợi ligno-xenluloza có thể thoái hóa sinh học. Sợi này thường được dùng làm nguyên liệu bao gói cho sản phẩm công nghiệp, và để sản xuất các sản phẩm có giá trị gia tăng khác, chẳng hạn như chỉ và vải dệt, dây thừng, dây bện và màn, sợi không dệt, giấy mỏng, bột nhão và giấy, vải địa chất, vật liệu kết hợp và vải dùng cho gia đình. Do tính chất thân thiện với môi trường và có thể thoái hóa sinh học của sợi đay, thị trường thế giới có nhu cầu ngày càng tăng về chúng. Để bắt kịp nhu cầu toàn cầu gia tăng, cần phải cải tiến hơn nữa việc sản xuất sợi đay. Sợi đay thu được từ vỏ của thân và các sợi được nhóm lại thành bó. Sợi được sắp xếp dưới dạng các cạnh tam giác của tế bào sợi mô cứng xen kẽ với tia lõi của các mô mềm khác. Các tế bào sợi này khu biệt với libe và phát triển trong bốn giai đoạn theo thứ tự thời gian – giai đoạn khởi đầu, giai đoạn kéo dài, giai đoạn làm đầy thành tế bào

thứ cấp, và giai đoạn trưởng thành. Có ít hiểu biết về việc kiểm soát di truyền của sự khởi đầu và kéo dài sợi dây. Tuy nhiên, độ dài và tính đồng đều sợi là yêu cầu chung đối với hầu hết các ứng dụng công nghiệp. Ví dụ, sợi dài và tế bào đồng đều là lý tưởng để sản xuất sợi mịn cho ngành công nghiệp dệt. Trong ngành công nghiệp bột nhão, đặc điểm độ bền của bột nhão được quyết định một phần bởi độ dài sợi, sản lượng bột nhão và sự tiêu thụ kiềm, do tính chất độ bền và tính chất bám dính của chúng. Để đáp ứng nhu cầu công nghiệp, sự phát triển của các giống dây có độ dài sợi cũng như là độ bền mong muốn là điều cần thiết. Do đó, sự tổng hợp sinh học của sợi và sinh học phân tử tham gia vào sự tổng hợp sinh học sợi có độ quan trọng đáng kể.

Sợi libe chủ yếu bắt nguồn từ mô phân sinh đỉnh và sợi này từ từ kéo dài ra, chủ yếu là thông qua sự sinh trưởng xâm nhập, đến khi bắt đầu phát triển thành thứ cấp. Sự kéo dài tế bào sợi là quá trình bắt nguồn từ sự tương tác giữa áp suất trương bên trong và độ bền cơ học của thành tế bào, mà được kiểm soát từ bên trong. Tuy nhiên, cơ chế và các gen tham gia vào sự kéo dài tế bào sợi chưa được xác định hoàn toàn.

Một vài gen ứng viên tham gia vào sự kéo dài và sự tạo thành của sợi bông đã được xác định. Ví dụ, năm gen của cây bông mà được biểu hiện đặc hiệu ở giai đoạn kéo dài sợi bông đã được xác định bằng phương pháp sàng lọc và bộc lộ khác biệt [Patent Mỹ số 5,880,100 và các Patent Mỹ số 5,932,713, 6,225,536 và 6,166,294], nhưng không có báo cáo nào như vậy trong lĩnh vực cây dây và/hoặc sự khởi đầu và sự kéo dài sợi libe.

Protein khóa kéo homeobox-loxin (HD-Zip) là yếu tố phiên mã duy nhất ở thực vật mà được mã hóa bằng nhiều bản sao của gen trong hệ gen thực vật. Ví dụ, hệ gen *Arabidopsis thaliana* có hơn 25 gen thuộc họ này. Trong số các gen này, protein khóa kéo homeobox-loxin HAT22 (protein HD-ZIP 22) tham gia vào sự kéo dài của tế bào sợi mạch.

Protein HD-ZIP 22 được biểu hiện ở vùng gian bó mà trong đó các sợi biệt hóa, mà phù hợp với vai trò của nó trong việc kiểm soát sự biệt hóa sợi gian bó. Ngoài ra, nó cũng được biểu hiện ở libe trong pha sinh trưởng thứ cấp (Torner et al. 1996), cho thấy vai trò có thể có của nó trong việc điều chỉnh sự tạo thành mô mạch (Zhong and Ye, 1999). Không giống như các gen homeobox thực vật khác, protein HD-ZIP 22 không được biểu hiện ở mô phân sinh, và chức năng của protein HD-ZIP 22 dường

như là tham gia vào sự điều chỉnh của sự đồng đều và/hoặc sự hoạt động của mô libe trong pha thứ cấp của sự phát triển mạch (Tornerio et al. 1996).

Dựa trên thực tế rằng protein khóa kéo homeobox-loxin HAT22 có thể đóng vai trò quan trọng trong sự kiểm soát không gian của sự khu biệt sợi libe, điều mong muốn là tiến hành nghiên cứu di truyền để hiểu sự tổng hợp sinh học của sợi ở thực vật bằng cách khảo sát và sử dụng thông tin sinh học phân tử và di truyền về protein khóa kéo homeobox-loxin HAT22. Bên cạnh đó, vì con đường tổng hợp sợi và sự sai khác cấu trúc di truyền của mỗi loài thực vật thường khác nhau, nghiên cứu đặc hiệu loài cũng được ưu tiên để tối ưu hóa năng suất của sợi từ cây đay, và thu được kết quả phù hợp để có thể sử dụng trong công nghiệp.

Bản chất kỹ thuật của sáng chế

Theo các phương án nhất định, sáng chế đề cập đến polynucleotit và polypeptit tham gia vào việc xác định hoặc tạo hình dạng chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng, và phương pháp sử dụng chúng.

Sáng chế cũng đề cập đến sự xác định gen tham gia vào quy trình có sự can thiệp của protein khóa kéo homeobox-loxin HAT22, và việc sử dụng gen và protein này để làm thay đổi các đặc tính ở cây trồng.

Sáng chế đề xuất protein khóa kéo homeobox-loxin HAT22 có nguồn gốc từ hai loài cây đay, *C. olitorius* và *C. capsularis*, phương pháp này bao gồm và/hoặc có chứa bước điều biến sự biểu hiện của polypeptit khóa kéo homeobox-loxin HAT22 ở tế bào thực vật, và phương pháp điều biến chiều dài sợi và/hoặc chiều cao và/hoặc sinh khối đã được tăng lên ở cây trồng.

Một khía cạnh của sáng chế là đề xuất gen mã hóa cho protein, có nguồn gốc từ *C. olitorius* và *C. capsularis*, mà tham gia vào việc tạo ra chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng tăng lên, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22. Cụ thể hơn, sáng chế đề xuất gen có khả năng cảm ứng và/hoặc làm tăng việc tạo ra chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng ở cây trồng; và protein được mã hóa bởi chúng, và các phương pháp sử dụng chúng.

Mục đích khác của sáng chế là đề xuất thông tin di truyền liên quan đến protein khóa kéo homeobox-loxin HAT22 để được khai thác/được sử dụng để gây cảm ứng,

cải thiện, hoặc làm tăng việc tạo ra chiều dài sợi và/hoặc chiều cao cây và/hoặc sinh khối ở cây trồng thuộc loài *C. olitorius* và *C. capsularis*, cũng như là ở cây trồng sản xuất sợi libre khác.

Mục đích khác nữa của sáng chế là để thu được cây trồng chuyển gen thuộc loài *C. olitorius* và *C. capsularis* có chiều dài sợi và/hoặc chiều cao cây/sinh khối tăng lên bằng cách điều biến sự tổng hợp sinh học của protein khóa kéo homeobox-loxin HAT22 ở cây trồng.

Mục đích khác nữa của sáng chế là đề xuất polynucleotit phân lập được, mà có thể hữu dụng để làm thuận tiện cho sự thực hiện phương pháp được bộc lộ trong bản mô tả này và sự phát triển của cây trồng *C. olitorius* và *C. capsularis* chuyển gen.

Mục đích khác nữa của sáng chế là đề xuất phương pháp khả thi về mặt thương mại để làm tăng chiều dài sợi và/hoặc chiều cao cây cho sản phẩm được tạo ra từ sợi đay.

Sáng chế còn đề xuất gen được phân lập từ *C. olitorius* mã hóa cho protein mà có trình tự axit amin như nêu trong SEQ ID NO: 3, mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây và/hoặc sinh khối, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22. Sáng chế cũng đề xuất gen mã hóa cho protein mà có trình tự axit amin được cải biến bằng cách thêm hoặc làm khuyết một hoặc một số axit amin và/hoặc thay thế bằng axit amin khác trong trình tự axit amin như nêu trong SEQ ID NO: 3, mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22. Sáng chế còn đề xuất gen mà lai hóa axit nucleic như nêu trong SEQ ID NO: 1, cụ thể là ADN của nó hoặc một phần của nó, và mã hóa cho protein mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối và có trình tự giống như protein khóa kéo homeobox-loxin HAT22.

Sáng chế còn đề xuất gen được phân lập từ *C. capsularis* mã hóa cho protein mà có trình tự axit amin như nêu trong SEQ ID NO: 6, mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22. Sáng chế cũng đề xuất gen mã hóa cho protein mà có trình tự axit amin được cải biến bằng cách thêm hoặc làm

khuyết một hoặc một số axit amin và/hoặc thay thế bằng axit amin khác trong trình tự axit amin như nêu trong SEQ ID NO: 6, mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22. Sáng chế còn đề xuất gen mà lai hóa với axit nucleic như nêu trong SEQ ID NO: 4, cụ thể là ADN của nó hoặc một phần của nó, và mã hóa cho protein mà tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối và có trình tự giống như protein khóa kéo homeobox-loxin HAT22.

Theo một phương án của sáng chế, cây *C. olitorius* là giống O-4, và cây *C. capsularis* là giống CVL-1.

Phương án khác của sáng chế bộc lộ cấu trúc gen tái tổ hợp bao gồm và/hoặc có chứa polynucleotit có trình tự nucleotit nêu trong SEQ ID NO: 2 và/hoặc SEQ ID NO: 5, trong đó polynucleotit có thể biểu hiện được ở tế bào chủ để sản xuất chất tương đồng của protein khóa kéo homeobox-loxin HAT22 lần lượt ở cây trồng *C. olitorius* và *C. capsularis*.

Phương án khác của sáng chế đề xuất thể biến nạp bao gồm và/hoặc có chứa cấu trúc gen tái tổ hợp có khả năng biểu hiện polynucleotit có trình tự nucleotit nêu trong SEQ ID NO: 2 và/hoặc SEQ ID NO: 5 để sản xuất chất tương đồng của protein khóa kéo homeobox-loxin HAT22.

Sáng chế cũng đề xuất phương pháp sản xuất protein tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối, và có trình tự giống như protein khóa kéo homeobox-loxin HAT22 bằng cách nuôi cấy và/hoặc trồng vật chủ nêu trên.

Sáng chế cũng đề xuất phương pháp gây cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối của cây trồng hoặc tế bào thực vật. Ngoài ra, sáng chế cũng bộc lộ phương pháp bao gồm và/hoặc có chứa bước đưa gen nêu trên vào cây trồng hoặc tế bào thực vật, và làm biểu hiện gen này.

Người có hiểu biết trung bình trong lĩnh vực sẽ dễ dàng nhận ra rằng sáng chế cũng thích hợp để thực hiện các mục đích và thu được kết quả và ưu điểm đã nêu,

cũng như là các kết quả và ưu điểm vốn có của nó. Các phương án được mô tả trong bản mô tả này không được dự định là làm giới hạn phạm vi của sáng chế.

Các dấu hiệu, khía cạnh và ưu điểm này và các dấu hiệu, khía cạnh và ưu điểm khác của sáng chế sẽ được hiểu rõ hơn khi tham khảo đến phần mô tả và yêu cầu bảo hộ sau đây.

Mô tả vắn tắt hình vẽ

Fig. 1 thể hiện cây phát sinh loài so sánh SEQ ID NO: 3 từ *C. olitorius* và SEQ ID NO: 6 từ *C. capsularis* với trình tự axit amin khác, mà sản xuất protein khóa kéo homeobox-loxin HAT22.

Mô tả chi tiết sáng chế

Sáng chế có thể được hiểu đầy đủ hơn từ phần mô tả chi tiết dưới đây và hình vẽ kèm theo mà tạo thành một phần của bản mô tả sáng chế này.

Các định nghĩa và/hoặc phương pháp được đề xuất trong bản mô tả này xác định sáng chế và hướng dẫn người có hiểu biết trung bình trong lĩnh vực thực hành sáng chế. Trừ khi có chỉ dẫn khác, các thuật ngữ cần được hiểu theo cách dùng thông thường bởi người có hiểu biết trung bình trong lĩnh vực liên quan. Ở chừng mực mà các định nghĩa và/hoặc phương pháp bất kỳ được phát hiện thấy là không thống nhất với định nghĩa và/hoặc phương pháp bất kỳ nêu trong tài liệu tham khảo sáng chế hoặc không sáng chế bất kỳ nêu trong bản mô tả này hoặc trong tài liệu tham khảo bất kỳ được tìm thấy ở nơi khác, cần hiểu rằng định nghĩa và/hoặc phương pháp mà đã được nêu /được thừa nhận rõ ràng trong bản mô tả này sẽ được áp dụng trong bản mô tả này. Các thuật ngữ số ít bao gồm cả nghĩa số nhiều trừ khi nội dung chỉ ra rõ ràng theo nghĩa khác. Tương tự, từ "hoặc" được dự định là bao gồm cả nghĩa "và" trừ khi nội dung chỉ ra rõ ràng nghĩa khác. Do đó, "bao gồm A hoặc B" có nghĩa là bao gồm A, hoặc B, hoặc A và B. Cũng cần hiểu rằng tất cả các kích thước bazơ hoặc kích thước axit amin, và tất cả các giá trị khối lượng phân tử hoặc số khối phân tử, được nêu đối với axit nucleic hoặc polypeptit là giá trị xấp xỉ, và được nêu để mô tả. Mặc dù phương pháp và nguyên liệu tương tự hoặc tương đương với phương pháp và nguyên liệu được mô tả trong bản mô tả này có thể được sử dụng để thực hiện hoặc thử nghiệm sáng chế này, phương pháp và nguyên liệu thích hợp được mô tả dưới đây.

Sáng chế bao gồm polynucleotit phân lập được mã hóa cho protein khóa kéo homeobox-loxin HAT22 được chiết từ *C. olitorius* và *C. capsularis*, và polypeptit tương ứng có nguồn gốc từ đó. Cụ thể hơn, sáng chế đề xuất chất tương đồng protein khóa kéo homeobox-loxin HAT22 và ứng dụng của chúng trong việc làm cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối ở *C. olitorius* và *C. capsularis*, cũng như là cây *C. olitorius* và *C. capsularis* chuyển gen của chúng. Trình tự hệ gen theo sáng chế mà mã hóa cho enzym đã được xác định chủ yếu bằng cách so sánh các trình tự nucleotit của ADN hệ gen *C. olitorius* và *C. capsularis* với các trình tự nucleotit của gen enzym đã biết của cây trồng khác. Trước sáng chế này, các trình tự nucleotit của các gen, khung đọc mở, vị trí của exon và intron, cấu trúc của enzym của *C. olitorius* và *C. capsularis*, và sự hữu dụng tiềm năng của chúng trong sự phát triển của chất lượng sợi cao chưa được biết tới, đặc biệt là đối với chiều dài sợi, chiều cao /sinh khối cây trồng và năng suất tiềm năng của cây đay.

Phân tích trình tự hệ gen của các loài đay được trồng thương mại, *C. olitorius* và *C. capsularis*, phát hiện ra rằng cả hai loài này có gen đơn lẻ mã hóa cho enzym có hoạt tính xúc tác trong việc làm cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây ở cây đay. Các trình tự nucleotit được chú thích ban đầu bằng chương trình phần mềm, chẳng hạn như Augustus, Semi-HMM-based Nucleic Acid Parser (SNAP), và Geneid (Genome BioInformatics Research Lab), mà có thể xác định giả định vùng mã hóa, intron, và mối nối ghép. Sự điều chỉnh tự động và bằng tay thêm nữa của các trình tự nucleotit được thực hiện để cải tiến và thiết lập sự mô tả đặc điểm chính xác của vùng mã hóa và đặc tính gen khác.

Trên 30.096 cADN từ *C. olitorius* và 37.031 cADN từ *C. capsularis* đã được đọc trình tự một phần hoặc toàn bộ. Từ đó, chúng tôi đã phát triển cADN đơn lẻ từ mỗi loài mã hóa cho enzym mới, có vai trò giả định, hoạt tính xúc tác được ưu tiên trong việc làm cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây ở cây đay.

Khung đọc mở (ORF) được phân tích sau khi đọc trình tự toàn bộ hoặc một phần của các dòng của thư viện cADN có nguồn gốc từ mARN *C. olitorius* và *C. capsularis*, và được phân tích thêm bằng cách sử dụng phần mềm phân tích trình tự, và

bằng cách xác định sự tương đồng với các trình tự đã biết trong cơ sở dữ liệu (công cộng/tư nhân).

Trong nội dung bản mô tả này, một số thuật ngữ được sử dụng trong toàn bộ bản mô tả có nghĩa như được chỉ ra trừ khi được chỉ ra rõ ràng là có nghĩa khác.

Như được sử dụng trong bản mô tả này, “polynucleotit” là trình tự nucleotit chẳng hạn như mảnh axit nucleic. Polynucleotit có thể là polyme của ARN hoặc ADN mà là mạch đơn hoặc mạch đôi, tùy ý có chứa bazơ nucleotit tổng hợp, không tự nhiên hoặc đã được làm thay đổi. Polynucleotit ở dạng polyme của ADN có thể có chứa và/hoặc gồm có một hoặc nhiều đoạn của cADN, ADN hệ gen, ADN tổng hợp, hoặc hỗn hợp/dạng kết hợp của chúng. Polynucleotit phân lập được theo sáng chế có thể bao gồm ít nhất một trong số 150 nucleotit liên tiếp (cả ngược chiều và xuôi chiều) có nguồn gốc từ SEQ ID NO: 1 và, SEQ ID NO: 4, hoặc dạng bổ sung của các trình tự này.

"Polypeptit" như được sử dụng trong bản mô tả này, là mạch thẳng đơn lẻ của các axit amin được liên kết với nhau bằng liên kết peptit, và thường có trình tự có chiều dài lớn hơn 100 axit amin.

"Được phân lập" có nghĩa là được làm thay đổi "bởi bàn tay con người" từ trạng thái tự nhiên. Nếu hợp phần hoặc chất có trong tự nhiên, nó đã "được phân lập" nếu nó được thay đổi hoặc được lấy ra từ môi trường ban đầu của nó, hoặc cả hai. Ví dụ, polynucleotit hoặc polypeptit có trong tự nhiên ở thực vật hoặc động vật sống không "được phân lập", nhưng cùng polynucleotit hoặc polypeptit đó được tách ra từ nguyên liệu đồng tồn tại của trạng thái tự nhiên của nó "được phân lập", như thuật ngữ được sử dụng trong bản mô tả này.

Thuật ngữ “gen”, như được sử dụng trong bản mô tả này, dùng để chỉ trình tự hệ gen của cây *C. olitorius* và *C. capsularis*, cụ thể là trình tự polynucleotit mã hóa cho trình tự polypeptit của enzym khóa kéo homeobox-loxin HAT22 tham gia vào hoạt tính xúc tác được ưu tiên trong việc làm tăng chiều dài sợi và/hoặc chiều cao cây ở cây đay. Thuật ngữ này có thể còn bao gồm phân tử axit nucleic bao gồm các trình tự nucleotit ngược chiều, xuôi chiều, và/hoặc intron.

"Trình tự mã hóa" hoặc "vùng mã hóa" dùng để chỉ phân tử axit nucleic có thông tin trình tự cần để sản xuất sản phẩm gen, chẳng hạn như axit amin hoặc polypeptit, khi trình tự này được biểu hiện. Trình tự mã hóa có thể có chứa và/hoặc bao gồm trình tự không dịch mã (bao gồm intron hoặc các vùng không dịch mã 5' hoặc 3') ở trong vùng dịch mã, hoặc có thể thiếu trình tự không dịch mã xen kẽ (ví dụ, như trong cADN).

Thuật ngữ "oligonucleotit", như được sử dụng trong bản mô tả này, là polynucleotit ngắn hoặc một phần của polynucleotit mà có thể có chiều dài tốt hơn là 10-1000, tốt nhất là từ 12 đến 50 nucleotit. Theo một phương án của sáng chế, nucleotit chứa trong oligonucleotit có thể là chất tương tự hoặc dẫn xuất của nucleotit có trong tự nhiên.

Thuật ngữ "đoạn mồi", như được sử dụng trong bản mô tả này, là oligonucleotit có khả năng liên kết với trình tự axit nucleic đích và làm mồi cho sự tổng hợp axit nucleic. Oligonucleotit khuếch đại như được xác định trong bản mô tả này có thể tốt hơn là có chiều dài từ 10 đến 50, tốt nhất là từ 15 đến 25 nucleotit. Ngoài ra, oligonucleotit khuếch đại theo sáng chế có thể được tổng hợp hóa học và oligonucleotit này là axit nucleic không có trong tự nhiên.

Từ viết tắt được sử dụng trong toàn bộ bản mô tả để chỉ axit nucleic bao gồm và/hoặc có chứa các trình tự nucleotit là các chữ viết tắt một kí tự thông thường. Do đó, khi được bao gồm trong axit nucleic, nucleotit mã hóa có trong tự nhiên được viết tắt như sau: adenin (A), guanin (G), xytosin (C), tymin (T) và uraxil (U). Ngoài ra, trừ khi có chỉ dẫn khác, trình tự axit nucleic được thể hiện trong bản mô tả này là theo chiều 5' → 3'.

Như được sử dụng trong bản mô tả này, thuật ngữ "bổ sung" và biến thể của nó được sử dụng để chỉ sự bắt cặp của axit nucleic theo quy tắc đã biết là A bắt cặp với T hoặc U và C bắt cặp với G. Sự bổ sung có thể là "một phần" hoặc "hoàn toàn". Trong sự bổ sung một phần, chỉ một số bazơ axit nucleic được bắt cặp theo quy tắc bắt cặp bazơ; trong khi đó trong sự bổ sung hoàn toàn hoặc toàn bộ, tất cả các bazơ được bắt cặp theo quy tắc bắt cặp này. Mức độ bổ sung giữa các sợi axit nucleic có thể có ảnh hưởng đáng kể lên hiệu quả và độ bền của sự lai hóa giữa các sợi axit nucleic như đã

biết trong lĩnh vực. Hiệu quả và độ bền của sự lai hóa này phụ thuộc vào phương pháp phát hiện.

Thuật ngữ "tế bào chủ", như được sử dụng trong bản mô tả này, bao gồm dạng tế bào bất kỳ mà dễ bị biến nạp, chuyển nạp, tải nạp, biểu hiện và quy trình tương tự bằng cấu trúc axit nucleic hoặc vector biểu hiện bao gồm và/hoặc có chứa polynucleotit của sáng chế. Tế bào chủ thích hợp bao gồm tế bào nấm và/hoặc tế bào thực vật, cụ thể là tế bào thực vật sản xuất sợi libe.

Thuật ngữ "được liên kết có thể điều khiển được" trong bản mô tả này thường dùng để chỉ cấu hình trong đó trình tự kiểm soát được đặt ở vị trí thích hợp đối với trình tự mã hóa của trình tự polynucleotit sao cho trình tự kiểm soát định hướng sự biểu hiện của trình tự mã hóa của polypeptit. Ví dụ, trình tự mở đầu có thể được liên kết có thể điều khiển được với trình tự mã hóa khi nó ảnh hưởng đến sự biểu hiện của trình tự mã hóa đó, tức là trình tự mã hóa đó chịu sự kiểm soát phiên mã của trình tự mở đầu.

"Vector" thường dùng để chỉ đơn vị sao chép, chẳng hạn như plasmit, phagơ, cosmit, nấm men hoặc virus mà đoạn axit nucleic khác có thể được chèn có thể điều khiển được vào đó sao cho dẫn đến sự sao chép hoặc sự biểu hiện của đoạn này. Thuật ngữ "vector" cũng nhằm để chỉ phân tử axit nucleic có khả năng vận chuyển axit nucleic khác đến nơi mà nó được liên kết. Một loại vector là "plasmit", mà dùng để chỉ vòng ADN mạch kép tròn mà đoạn ADN khác có thể được nối vào. Một loại vector khác là vector virus, trong đó đoạn ADN khác có thể được nối vào hệ gen virus. Các vector nhất định có khả năng tự sao chép trong tế bào chủ mà chúng được đưa vào (ví dụ, vector vi khuẩn có điểm mở đầu sao chép vi khuẩn và vector động vật có vú thuộc thể bổ sung). Các vector khác có thể được hợp nhất vào hệ gen của tế bào chủ khi đưa vào trong tế bào chủ, và bằng cách đó chúng được sao chép cùng với hệ gen vật chủ. Ngoài ra, các vector nhất định có khả năng định hướng sự biểu hiện của gen mà chúng được liên kết có thể điều khiển được với. Các vector này được đề cập đến trong bản mô tả này là "vector biểu hiện tái tổ hợp" (hoặc đơn giản là "vector biểu hiện"). Nhìn chung, vector biểu hiện để sử dụng trong kỹ thuật ADN tái tổ hợp thường ở dạng plasmit. Trong bản mô tả này, "plasmit" và "vector" có thể được sử dụng thay thế lẫn nhau vì plasmit là dạng thường được sử dụng nhất của vector. Tuy nhiên, sáng chế được dự

định bao gồm các dạng vector biểu hiện khác, chẳng hạn như vector virus (ví dụ, retrovirus, adenovirus và virus kết hợp adeno không có khả năng sao chép), mà có chức năng tương đương.

Thuật ngữ "cấu trúc axit nucleic" hoặc "cấu trúc ADN" đôi khi được dùng để chỉ trình tự hoặc các trình tự mã hóa được liên kết có thể điều khiển được với trình tự điều hòa thích hợp và được chèn vào vector để làm biến nạp tế bào. Thuật ngữ này có thể được sử dụng thay thế lẫn nhau với thuật ngữ "ADN biến nạp" hoặc "gen chuyển".

Thuật ngữ "promoter (vùng khởi đầu)", như được sử dụng trong bản mô tả này, dùng để chỉ trình tự axit nucleic mà có chức năng định hướng sự phiên mã của gen xuôi chiều. Promoter thường thích hợp với tế bào chủ trong đó gen đích sẽ được biểu hiện. Promoter cùng với trình tự axit nucleic điều hòa phiên mã và dịch mã khác (còn được gọi là “trình tự kiểm soát”) là cần thiết để biểu hiện gen nhất định. Nhìn chung, trình tự điều hòa phiên mã và dịch mã bao gồm, nhưng không giới hạn ở, promoter, vị trí liên kết ribosom, trình tự khởi đầu và kết thúc phiên mã, trình tự khởi đầu và kết thúc dịch mã, và các trình tự tăng cường hoặc hoạt hóa.

Như được sử dụng trong bản mô tả này, “polypeptit khóa kéo homeobox-loxin HAT22” có nghĩa là polypeptit có sự hoạt động khóa kéo homeobox-loxin HAT22, tức là, tham gia vào việc cảm ứng, cải thiện, hoặc làm tăng chiều dài sợi và/hoặc chiều cao cây/sinh khối ở cây trồng. Sự hoạt động protein khóa kéo homeobox-loxin HAT22 kích thích sự sinh trưởng tế bào sợi và/hoặc chiều cao cây/sinh khối. Tính chất liên kết ADN của protein khóa kéo homeobox-loxin HAT22 cho thấy chúng tương tác với ADN dưới dạng homodimer và nhận diện hai trình tự pseudopalindromic 9 bp (xấp xỉ) riêng biệt, CAAT(G/C)ATTG, như được xác định bằng cách chọn lọc các vị trí liên kết ái lực cao từ ADN trình tự - ngẫu nhiên. Thuật ngữ này cũng bao gồm mảnh, biến thể, và chất tương đồng, có một chức năng bất kỳ trong số các chức năng nêu trên.

Thuật ngữ "chất tương đồng", như được sử dụng trong bản mô tả này, dùng để chỉ protein bao hàm peptit, oligopeptit, polypeptit, protein và enzym có sự thể, làm khuyết và/hoặc cải xen axit amin so với protein chưa được cải biến đang quan tâm và có hoạt tính sinh học và chức năng tương tự như protein chưa được cải biến mà chúng bắt nguồn từ đó.

Xóa dùng để chỉ sự loại bỏ một hoặc nhiều axit amin từ protein.

Chèn dùng để chỉ một hoặc nhiều gốc axit amin được đưa vào vị trí được xác định trước trong protein. Chèn có thể bao gồm sự dung hợp đầu N và/hoặc đầu C cũng như là chèn nội trình tự của một hoặc nhiều axit amin. Thông thường, chèn ở trong trình tự axit amin sẽ nhỏ hơn dung hợp đầu N hoặc đầu C, ở mức khoảng từ 1 đến 10 gốc.

Thế dùng để chỉ sự thay thế axit amin của protein bằng axit amin khác có tính chất tương tự (chẳng hạn như tính kỵ nước, tính ưa nước, tính kháng nguyên, xu hướng tạo thành hoặc phá vỡ cấu trúc xoắn α hoặc cấu trúc tấm β tương tự). Sự thế axit amin thường là ở gốc đơn lẻ, nhưng có thể được nhóm lại tùy thuộc vào sự ràng buộc chức năng đặt lên polypeptit và có thể nằm trong khoảng từ 1 đến 10 axit amin; sự cài xen thường ở mức từ khoảng 1 đến 10 gốc axit amin. Sự thế axit amin tốt hơn là sự thế axit amin bảo toàn. Bảng sự thế bảo toàn là đã biết trong lĩnh vực (ví dụ, xem tài liệu Creighton (1984) *Proteins*. W.H. Freeman and Company (Eds) và Bảng 1 dưới đây).

Bảng 1: Ví dụ về sự thế axit amin bảo toàn

Gốc	Sự thế bảo toàn	Gốc	Sự thế bảo toàn
Ala	Ser	Leu	Ile; Val
Arg	Lys	Lys	Arg; Gln
Asn	Gln; His	Met	Leu; Ile
Asp	Glu	Phe	Met; Leu; Tyr
Gln	Asn	Ser	Thr; Gly
Cys	Ser	Thr	Ser; Val
Glu	Asp	Trp	Tyr
Gly	Pro	Tyr	Trp; Phe
His	Asn; Gln	Val	Ile; Leu
Ile	Leu, Val		

Sự thế, xóa và/hoặc chèn axit amin có thể dễ dàng được tạo ra bằng cách sử dụng kỹ thuật tổng hợp peptit đã biết trong lĩnh vực, chẳng hạn như tổng hợp peptit

pha rắn và/hoặc kỹ thuật tổng hợp khác bất kỳ, hoặc bằng thao tác ADN tái tổ hợp. Phương pháp thao tác trình tự ADN để tạo ra biến thể thể, cài xen hoặc làm khuyết của protein là đã biết trong lĩnh vực. Ví dụ, kỹ thuật tạo ra đột biến thể ở vị trí được xác định trước trong ADN là đã biết đối với người có hiểu biết trung bình trong lĩnh vực và bao gồm gây đột biến M13, gây đột biến *in vitro* T7-Gen (USB, Cleveland, OH), gây đột biến định hướng vị trí thay đổi nhanh (Stratagene, San Diego, CA), gây đột biến định hướng vị trí qua PCR hoặc các quy trình gây đột biến định hướng vị trí khác.

Như được sử dụng trong bản mô tả này, “phần có hoạt tính sinh học” có thể dùng để chỉ mảnh của protein khóa kéo homeobox-loxin HAT22 có hoạt tính sinh học để xúc tác cho sự khởi đầu, sự tạo thành, sự tăng cường, hoặc sự thay đổi trong hợp phần của sợi libe ở cây trồng *C. olitorius* hoặc *C. capsularis*. Phần có hoạt tính sinh học của protein khóa kéo homeobox-loxin HAT22 bao gồm peptit hoặc polypeptit bao gồm trình tự axit amin có đủ độ tương đồng với hoặc có nguồn gốc từ trình tự axit amin của protein khóa kéo homeobox-loxin HAT22, ví dụ, trình tự axit amin như nêu trong SEQ ID NO: 3 hoặc SEQ ID NO: 6, mà bao gồm ít axit amin hơn so với protein khóa kéo homeobox-loxin HAT22 chiều dài đầy đủ, và thể hiện ít nhất một hoạt tính của protein khóa kéo homeobox-loxin HAT22. Thông thường, phần có hoạt tính sinh học có chứa miền hoặc motif có ít nhất một hoạt tính của protein khóa kéo homeobox-loxin HAT22. Phần có hoạt tính sinh học của protein khóa kéo homeobox-loxin HAT22 có thể là polypeptit mà có chiều dài là, ví dụ, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, hoặc 267 axit amin.

Protein khóa kéo homeobox-loxin HAT22 có thể có trình tự axit amin nêu trong SEQ ID NO: 3 hoặc SEQ ID NO: 6. Theo phương án khác, protein khóa kéo homeobox-loxin HAT22 về cơ bản tương đồng với SEQ ID NO: 3 hoặc SEQ ID NO: 6, và giữ lại hoạt tính chức năng của protein của SEQ ID NO: 3 hoặc SEQ ID NO: 6, vẫn khác biệt ở trình tự axit amin do sự thay đổi alen tự nhiên hoặc sự gây đột biến. Theo phương án khác, protein khóa kéo homeobox-loxin HAT22 có chứa trình tự axit amin có độ tương đồng ít nhất là khoảng 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9% hoặc lớn hơn so với SEQ ID NO: 3 hoặc SEQ ID NO: 6.

Thuật ngữ "miền", như được sử dụng trong bản mô tả này, dùng để chỉ tập hợp axit amin được bảo toàn ở các vị trí cụ thể dọc theo sự sắp thẳng hàng của các trình tự của các protein có liên quan về mặt tiến hóa. Mặc dù axit amin ở các vị trí khác có thể thay đổi giữa các chất tương đồng, axit amin mà được bảo toàn cao ở các vị trí cụ thể cho thấy axit amin này dường như là thiết yếu trong cấu trúc, độ ổn định hoặc chức năng của protein. Được xác định bằng mức độ bảo toàn cao của chúng trong các trình tự được sắp thẳng hàng của họ chất tương đồng protein, chúng có thể được sử dụng làm chất nhận dạng để xác định liệu polypeptit bất kỳ đang quan tâm có thuộc về họ polypeptit đã được xác định từ trước hay không.

Thuật ngữ "motif", như được sử dụng trong bản mô tả này, dùng để chỉ vùng được bảo toàn ngắn trong trình tự của các protein có liên quan về mặt tiến hóa. Motif thường là các phần được bảo toàn cao của miền, nhưng cũng có thể chỉ bao gồm một phần của miền, hoặc nằm ở bên ngoài miền được bảo toàn (nếu tất cả các của axit amin của motif nằm ngoài miền đã được xác định).

Có các cơ sở dữ liệu chuyên môn để xác định các miền, ví dụ, SMART (Schultz et al. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic et al. (2002) Nucleic Acids Res 30, 242-244), InterPro (Mulder et al., (2003) Nucl. Acids. Res. 31, 315-318), Prosite (Bucher and Bairoch (1994), A generalized profile syntax for biomolecular sequences motifs and its function in automatic sequence interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., pp53-61, AAAI Press, Menlo Park; Hulo et al., Nucl. Acids. Res. 32:D134-D137, (2004)), hoặc Pfam (Bateman et al., Nucleic Acids Research 30(1): 276-280 (2002)). Tập hợp công cụ để phân tích *in silico* các trình tự protein có sẵn trên máy chủ proteomics ExPASy (Viện Thông tin sinh học Thụy Sĩ (Gasteiger et al., ExPASy: The Proteomics Server for In-Depth Protein Knowledge and Analysis, Nucleic Acids Res. 31:3784-3788(2003)). Miền hoặc motif cũng có thể được xác định bằng cách sử dụng kỹ thuật thông thường, chẳng hạn như bằng cách sắp thẳng hàng trình tự.

Nhằm đạt được mục đích của sáng chế, "chuyển gen", "gen chuyển" hoặc "tái tổ hợp" có nghĩa là đối với, ví dụ, trình tự axit nucleic, bằng biểu hiện, cấu trúc gen hoặc vector bao gồm và/hoặc có chứa trình tự axit nucleic hoặc sinh vật được biến nạp

bằng trình tự axit nucleic, bằng biểu hiện hoặc vector theo sáng chế, tất cả các các cấu trúc này được tạo ra bằng phương pháp tái tổ hợp trong đó: (a) trình tự axit nucleic mã hóa cho protein hữu dụng trong phương pháp theo sáng chế, hoặc (b) (các) trình tự kiểm soát di truyền mà được liên kết có thể điều khiển được với trình tự axit nucleic theo sáng chế, ví dụ trình tự mở đầu, hoặc (c) a) và b) không nằm ở môi trường tự nhiên của chúng hoặc đã được cải biến bằng phương pháp tái tổ hợp. Sự cải biến có thể ở dạng, ví dụ, sự thế, sự thêm, sự làm khuyết, sự đảo ngược hoặc sự cài xen của một hoặc nhiều gốc nucleotit. Môi trường di truyền tự nhiên được hiểu là vị trí trong hệ gen hoặc nhiễm sắc thể tự nhiên ở cây trồng ban đầu hoặc sự có mặt trong thư viện hệ gen. Trong trường hợp thư viện hệ gen, môi trường di truyền tự nhiên của trình tự axit nucleic tốt hơn là được giữ lại, ít nhất là một phần. Môi trường tấn công ở bên sườn trình tự axit nucleic ít nhất ở một phía và có chiều dài trình tự bằng khoảng 50 bp, tốt hơn là bằng khoảng 500 bp. Bảng biểu hiện có trong tự nhiên - ví dụ dạng kết hợp có trong tự nhiên của trình tự mở đầu tự nhiên của trình tự axit nucleic với trình tự axit nucleic tương ứng mã hóa cho polypeptit hữu dụng trong phương pháp theo sáng chế, như xác định ở trên - trở thành bảng biểu hiện chuyển gen khi bảng biểu hiện này được cải biến bằng phương pháp tổng hợp ("nhân tạo"), không tự nhiên chẳng hạn như, ví dụ, xử lý gây đột biến. Phương pháp thích hợp được mô tả, ví dụ, trong tài liệu US 5,565,350 hoặc WO 00/15815.

Thực vật chuyển gen để thực hiện mục đích của sáng chế được hiểu là bao gồm các cây mà trong đó axit nucleic được sử dụng trong phương pháp theo sáng chế không ở locut tự nhiên của chúng trong hệ gen của thực vật này, và do đó axit nucleic có thể được biểu hiện tương đồng hoặc dị tương đồng. Tuy nhiên, chuyển gen cũng có nghĩa là, mặc dù axit nucleic theo sáng chế hoặc được sử dụng trong phương pháp theo sáng chế nằm ở locut tự nhiên của chúng trong hệ gen của thực vật này, trình tự này đã được cải biến so với trình tự tự nhiên, và/hoặc trình tự điều hòa của các trình tự tự nhiên đã được cải biến. Chuyển gen được hiểu là sự biểu hiện của axit nucleic theo sáng chế ở vị trí không tự nhiên trong hệ gen, trong khi sự biểu hiện tương đồng hoặc, tốt hơn là, dị tương đồng của axit nucleic này diễn ra.

Thuật ngữ "đưa vào" hoặc "biến nạp", như được sử dụng trong bản mô tả này, dùng để chỉ sự chuyển polynucleotit ngoại sinh vào tế bào chủ, bằng bất kể phương

pháp nào được sử dụng để chuyển. Mô thực vật có khả năng nhân dòng tiếp theo, bằng cách phát sinh cơ quan hoặc phát sinh phôi, có thể được biến nạp bằng cấu trúc di truyền theo sáng chế và cây trồng nguyên vẹn được tái tạo từ đó. Mô cụ thể được chọn sẽ thay đổi tùy thuộc vào hệ thống nhân dòng có sẵn, và thích hợp nhất, đối với loài cụ thể cần được biến nạp. Các đích mô ví dụ bao gồm đĩa lá, phần, phôi, lá mầm, trụ dưới lá mầm, thể giao tử cái, mô sẹo, mô phân sinh hiện hữu (ví dụ, mô phân sinh đỉnh, chồi nách, và mô phân sinh rễ), và mô phân sinh cảm ứng (ví dụ, mô phân sinh lá mầm và mô phân sinh trụ dưới lá mầm). Polynucleotit có thể được đưa vào tạm thời hoặc ổn định trong tế bào chủ và có thể được duy trì không hợp nhất, ví dụ, dưới dạng plasmid. Theo cách khác, nó có thể được hợp nhất vào hệ gen vật chủ. Sau đó, tế bào thực vật biến nạp thu được có thể được sử dụng để tạo ra cây trồng biến nạp theo phương pháp đã biết đối với người có hiểu biết trung bình trong lĩnh vực.

Việc chuyển gen lạ vào hệ gen của cây trồng được gọi là biến nạp.

Sự biến nạp của loài thực vật hiện nay là kỹ thuật khá thông thường. Điều có lợi là, phương pháp biến nạp bất kỳ trong số một vài phương pháp biến nạp có thể được sử dụng để đưa gen mong muốn vào tế bào ông bà thích hợp. Phương pháp được mô tả để biến nạp và tái tạo cây trồng từ mô thực vật hoặc tế bào thực vật có thể được sử dụng để biến nạp tạm thời hoặc ổn định. Phương pháp biến nạp bao gồm việc sử dụng liposom, đục lỗ điện và hóa chất mà làm tăng sự hấp thụ ADN tự do, tiêm ADN trực tiếp vào cây trồng, súng bắn hạt, và biến nạp bằng cách sử dụng virus hoặc phần và vi chiếu ảnh. Cây trồng chuyển gen, bao gồm cây trồng mùa vụ chuyển gen, tốt hơn là được sản xuất thông qua sự biến nạp qua trung gian *Agrobacterium*. Phương pháp biến nạp có lợi là biến nạp trong giới thực vật (Sajib et. al. 2008).

Thông thường sau khi biến nạp, tế bào thực vật hoặc nhóm tế bào được chọn lọc về sự có mặt của một hoặc nhiều gen đánh dấu mà được mã hóa bởi gen có thể biểu hiện ở cây trồng được đồng biến nạp với gen mong muốn, sau đó nguyên liệu được biến nạp được tái tạo vào cây trồng nguyên vẹn. Để chọn lọc cây trồng được biến nạp, nguyên liệu cây trồng thu được trong quy trình biến nạp, về nguyên tắc, được đưa vào điều kiện chọn lọc sao cho cây trồng được biến nạp có thể phân biệt được đối với cây trồng không được biến nạp. Ví dụ, hạt thu được theo phương thức được mô tả ở trên có thể được gieo và, sau giai đoạn phát triển ban đầu, được đưa đi chọn lọc thích

hợp bằng cách phun. Khả năng khác nữa bao gồm việc trồng hạt, nếu thích hợp sau khi tiệt trùng, trên đĩa aga bằng cách sử dụng chất chọn lọc thích hợp sao cho chỉ hạt được biến nạp có thể phát triển thành cây trồng. Theo cách khác, cây trồng được biến nạp được sàng lọc về sự có mặt của gen đánh dấu chọn lọc chẳng hạn như các gen được mô tả ở trên.

Sau khi truyền ADN và tái tạo, giả định rằng cây trồng được biến nạp cũng có thể được đánh giá, ví dụ bằng cách sử dụng phân tích Southern, về sự có mặt của gen mong muốn, số lượng bản sao và/hoặc sự tổ chức hệ gen. Theo cách khác hoặc ngoài ra, mức độ biểu hiện của ADN mới được đưa vào có thể được kiểm tra bằng cách sử dụng phân tích Northern và/hoặc Western, cả hai kỹ thuật này là đã biết đối với người có hiểu biết trung bình trong lĩnh vực.

Cây trồng được biến nạp được tạo ra có thể được nhân lên bằng nhiều phương pháp, chẳng hạn như bằng cách nhân dòng hoặc kỹ thuật gây giống cổ điển. Ví dụ, cây trồng được biến nạp thế hệ thứ nhất (hay T1) có thể được chọn lọc và thế biến nạp thế hệ thứ hai đồng hợp tử (hay T2) có thể được chọn lọc, và sau đó cây trồng T2 có thể được tiếp tục nhân lên bằng kỹ thuật gây giống cổ điển. Sinh vật biến nạp được tạo ra có thể có nhiều dạng. Ví dụ, chúng có thể là thể khảm của tế bào được biến nạp và tế bào không được biến nạp; thể biến nạp đơn dòng (ví dụ, tất cả các tế bào được biến nạp đều có chứa băng biểu hiện); cảnh ghép của mô được biến nạp và không được biến nạp (ví dụ, ở cây trồng, thân rễ được biến nạp được ghép vào chồi không được biến nạp).

Thuật ngữ "sự biểu hiện tăng" hoặc "sự biểu hiện quá mức" như được sử dụng trong bản mô tả này dùng để chỉ dạng bất kỳ của sự biểu hiện mà tăng thêm so với mức độ biểu hiện kiểu dại ban đầu.

Phương pháp làm tăng sự biểu hiện của gen hoặc sản phẩm gen đã được ghi nhận trong tài liệu trong lĩnh vực và bao gồm, ví dụ, sự biểu hiện quá mức được làm cho diễn ra bằng trình tự mở đầu thích hợp, bằng cách sử dụng gen tăng cường phiên mã hoặc gen tăng cường dịch mã. Axit nucleic được phân lập mà đóng vai trò làm trình tự mở đầu hoặc yếu tố tăng cường có thể được đưa vào vị trí thích hợp (thường là ngược chiều) của dạng không khác loại của polynucleotit sao cho điều hòa tăng sự biểu hiện của axit nucleic mã hóa cho polypeptit mong muốn. Ví dụ, trình tự mở đầu

nội sinh có thể được làm thay đổi *in vivo* bằng đột biến, sự làm khuyết, và/hoặc sự thể (xem tài liệu, Kmiec, US 5,565,350; Zarling et al., WO93/22443), hoặc trình tự mở đầu được phân lập có thể được đưa vào tế bào thực vật theo hướng thích hợp và ở cách gen theo sáng chế một khoảng sao cho kiểm soát sự biểu hiện của gen.

Nếu sự biểu hiện polypeptit là điều mong muốn, thường mong muốn bao gồm vùng polyadenyl hóa ở đầu 3' của vùng mã hóa polynucleotit. Vùng polyadenyl hóa có thể là có nguồn gốc từ gen tự nhiên, từ nhiều loại gen thực vật khác, hoặc từ T-ADN. Trình tự đầu 3' cần được bổ sung vào có thể có nguồn gốc từ, ví dụ, gen nopalín syntaza hoặc gen octopín syntaza, hoặc theo cách khác là từ gen thực vật khác, hoặc kém hơn là từ gen sinh vật nhân chuẩn khác bất kỳ.

Trình tự intron cũng có thể được bổ sung vào vùng không dịch mã 5' (UTR) hoặc trình tự mã hóa của một phần trình tự mã hóa để làm tăng lượng thông tin trưởng thành tích lũy trong dung dịch bào tương. Sự bao gồm của chỗ nối làm cho intron trong đơn vị phiên mã ở cả cấu trúc biểu hiện thực vật và động vật được thể hiện để làm tăng sự biểu hiện gen ở cả cấp độ mRNA và protein lên đến 1000 lần (Buchman and Berg (1988) Mol. Cell biol. 8: 4395-4405; Callis et al. (1987) Genes Dev 1:1183-1200). Sự tăng cường intron này của sự biểu hiện gen thường lớn nhất khi đặt ở gần đầu 5' của đơn vị phiên mã. Việc sử dụng các intron ngô như intron Adh1-S 1, 2, và 6, intron Bronze-1 là đã biết trong lĩnh vực. Thông tin chung về chúng xem trong tài liệu: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, N.Y. (1994).

Để xác định sự tương đồng theo tỷ lệ phần trăm của hai trình tự axit amin hoặc của hai trình tự axit nucleic, các trình tự có thể được sắp thẳng hàng nhằm mục đích so sánh tối ưu (ví dụ, chỗ trống có thể được đưa vào một hoặc cả hai trình tự axit amin hoặc trình tự axit nucleic thứ nhất hoặc thứ hai để sắp thẳng hàng tối ưu và các trình tự không tương đồng có thể bị bỏ qua khi so sánh). Chiều dài của trình tự tham chiếu được sắp thẳng hàng để so sánh có thể bằng ít nhất 95% chiều dài của trình tự tham chiếu này. Sau đó, gốc axit amin hoặc nucleotit ở vị trí axit amin hoặc vị trí nucleotit tương ứng có thể được so sánh. Khi vị trí ở trình tự thứ nhất được chiếm giữ bởi cùng gốc axit amin hoặc nucleotit như vị trí tương ứng ở trình tự thứ hai, thì các phân tử tương đồng ở vị trí đó (như được sử dụng trong bản mô tả này “sự tương đồng” axit amin hoặc axit nucleic là tương đương với “sự đồng đều” axit amin hoặc axit nucleic).

Sự tương đồng theo tỷ lệ phần trăm giữa hai trình tự là hàm của số lượng vị trí tương đồng giữa các trình tự, có tính đến số lượng chỗ trống, và chiều dài của mỗi chỗ trống, mà cần phải được đưa vào để sắp thẳng hàng tối ưu hai trình tự.

Việc so sánh các trình tự và xác định sự tương đồng theo tỷ lệ phần trăm giữa hai trình tự có thể được thực hiện bằng cách sử dụng thuật toán. Theo một phương án, sự tương đồng theo tỷ lệ phần trăm giữa hai trình tự axit amin có thể được xác định bằng cách sử dụng thuật toán Needleman và Wunsch (J. Mol. Biol. 48:444-453 (1970)) mà đã được kết hợp vào chương trình GAP trong gói phần mềm GCG (có sẵn ở <http://www.gcg.com>), bằng cách sử dụng ma trận Blosum 62 hoặc ma trận PAM250, và tải trọng chỗ trống bằng 16, 14, 12, 10, 8, 6, hoặc 4 và tải trọng chiều dài bằng 1, 2, 3, 4, 5, hoặc 6. Theo phương án được ưu tiên khác nữa, sự tương đồng theo tỷ lệ phần trăm giữa hai trình tự nucleotit có thể được xác định bằng cách sử dụng chương trình GAP trong gói phần mềm GCG (có sẵn ở <http://www.gcg.com>), bằng cách sử dụng ma trận NWSgapdna.CMP và tải trọng chỗ trống bằng 40, 50, 60, 70, hoặc 80 và tải trọng chiều dài bằng 1, 2, 3, 4, 5, hoặc 6. Theo phương án khác, sự tương đồng theo tỷ lệ phần trăm giữa hai axit amin hoặc các trình tự nucleotit có thể được xác định bằng cách sử dụng thuật toán của E. Meyers và W. Miller (Comput. Appl. Biosci. 4:11-17 (1988)) mà đã được kết hợp vào chương trình ALIGN (phiên bản 2.0 hoặc 2.0U), bằng cách sử dụng bảng gốc tải trọng PAM120, khoản trừ chiều dài chỗ trống bằng 12 và khoản trừ chỗ trống bằng 4.

Chương trình máy tính ví dụ mà có thể được sử dụng để xác định sự đồng đều giữa hai trình tự bao gồm, nhưng không giới hạn ở, đây các chương trình BLAST, *ví dụ*, BLASTN, BLASTX, và TBLASTX, BLASTP và TBLASTN, có thể truy cập công cộng ở www.ncbi.nlm.nih.gov/BLAST.

Nghiên cứu trình tự thường được thực hiện bằng cách sử dụng chương trình BLASTN, khi đánh giá trình tự axit nucleic nhất định đối với trình tự axit nucleic trong trình tự ADN GenBank và các cơ sở dữ liệu công khai khác. Chương trình BLASTX được ưu tiên để nghiên cứu trình tự axit nucleic mà được dịch mã trong tất cả các khung đọc so với trình tự axit amin trong các trình tự protein GenBank và các cơ sở dữ liệu công khai khác.

Sự sắp xếp trình tự của các trình tự được chọn để xác định "% tương đồng" giữa hai hoặc nhiều trình tự được thực hiện bằng cách sử dụng ví dụ, chương trình CLUSTAL-W.

Một phương án của sáng chế là polynucleotit phân lập được mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22 được phát hiện ở *C. olitorius* và *C. capsularis* lần lượt bao gồm và/hoặc có chứa trình tự nucleotit như nêu trong SEQ ID NO: 2 và SEQ ID NO: 5. Một cách tương ứng, polypeptit khóa kéo homeobox-loxin HAT22 tương ứng được mã hóa bởi các trình tự nucleotit sẽ mang trình tự axit amin như nêu trong SEQ ID NO: 3 và SEQ ID NO: 6. Theo một phương án theo sáng chế, SEQ ID NO: 3 dùng để chỉ trình tự polypeptit của chất tương đồng protein khóa kéo homeobox-loxin HAT22 có nguồn gốc từ *C. olitorius*, và SEQ ID NO: 6 dùng để chỉ trình tự polypeptit của chất tương đồng protein khóa kéo homeobox-loxin HAT22 có nguồn gốc từ *C. capsularis*. Cả hai enzym này có mặt trong con đường tổng hợp sinh học của sợi ở *C. olitorius* và *C. capsularis* để xúc tác cho việc tạo ra chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng tăng lên ở cây trồng.

Sáng chế cũng đề xuất gen trình tự mã hóa cho chất tương đồng protein khóa kéo homeobox-loxin HAT22 từ cây trồng *C. olitorius* và *C. capsularis*.

Theo một phương án, polynucleotit dài 1296 bp được minh họa trong SEQ ID NO: 1 là gen chiều dài đầy đủ được phân lập từ *C. olitorius*. Trình tự gen này bao gồm ít nhất 150 nucleotit liên tiếp từ cả ngược chiều và xuôi chiều của gen. Phương án này cũng đề xuất trình tự intron của gen.

Theo phương án khác, polynucleotit dài 1337 bp được minh họa trong SEQ ID NO: 4 là gen chiều dài đầy đủ được phân lập từ *C. capsularis*. Trình tự gen này bao gồm ít nhất 150 nucleotit liên tiếp từ cả ngược chiều và xuôi chiều của gen. Phương án này cũng đề xuất trình tự intron của gen.

Theo phương án khác nữa, sáng chế đề xuất polynucleotit phân lập được mã hóa cho polypeptit bao gồm và/hoặc có chứa trình tự nucleotit nêu trong SEQ ID NO: 2 và/hoặc SEQ ID NO: 5. SEQ ID NO: 2 dùng để chỉ trình tự polynucleotit của trình tự chất tương đồng protein khóa kéo homeobox-loxin HAT22 có nguồn gốc từ *C. olitorius* và SEQ ID NO: 5 dùng để chỉ trình tự polynucleotit của trình tự chất tương đồng protein khóa kéo homeobox-loxin HAT22 có nguồn gốc từ *C. capsularis*.

Theo phương án khác nữa, phân tử axit nucleic được phân lập mà có khả năng mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22, hoặc mảnh có hoạt tính sinh học của chúng, có chứa trình tự nucleotit mà có độ tương đồng ít nhất là khoảng 95%, 96%, 97%, 98%, 99%, 99,1%, 99,2%, 99,3%, 99,4%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9% hoặc lớn hơn so với toàn bộ chiều dài của trình tự nucleotit nêu trong SEQ ID NO: 2, SEQ ID NO: 5, hoặc dạng bổ sung bất kỳ của chúng.

Theo một phương án, polynucleotit dài 804 bp được minh họa trong SEQ ID NO: 2 là protein khóa kéo homeobox-loxin HAT22 được mã hóa bởi dòng cADN chiều dài đầy đủ thể hiện khung đọc mở mã hóa cho polypeptit có 267 axit amin, như trong SEQ ID NO: 3, có số khối phân tử tính được bằng khoảng 29,99 kD. Thông qua phân tích SMART của SEQ ID NO: 3, đã phát hiện ra sự có mặt của miền homeobox trong trình tự. Đây là yếu tố liên kết ADN, mà tham gia vào sự điều hòa phiên mã của các quy trình phát triển then chốt của cây trồng. Đây là protein khóa kéo homeobox-loxin HAT22 tham gia vào việc tạo ra chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng tăng lên ở thực vật.

Theo một phương án, polynucleotit dài 807 bp được minh họa trong SEQ ID NO: 5 là protein khóa kéo homeobox-loxin HAT22 được mã hóa bởi dòng cADN chiều dài đầy đủ thể hiện khung đọc mở mã hóa cho polypeptit có 268 axit amin, như trong SEQ ID NO: 6, có số khối phân tử tính được bằng khoảng 30,10 kD. Thông qua phân tích SMART của SEQ ID NO: 6, đã phát hiện ra sự có mặt của miền homeobox trong trình tự. Đây là yếu tố liên kết ADN, mà tham gia vào sự điều hòa phiên mã của các quy trình phát triển then chốt của cây trồng. Đây là protein khóa kéo homeobox-loxin HAT22 tham gia vào việc tạo ra chiều dài sợi, chiều cao cây, và/hoặc sinh khối cây trồng tăng lên ở thực vật.

Theo một phương án của sáng chế, polynucleotit phân lập được minh họa trong SEQ ID NO: 2 có thể thu được bằng cách khuếch đại PCR của vùng được bảo toàn của gen bằng cách sử dụng ARN tổng số được phân lập từ cây *C. olitorius*, và SEQ ID NO: 5 có thể thu được bằng cách khuếch đại PCR của vùng được bảo toàn của gen này bằng cách sử dụng ARN tổng số được phân lập từ cây *C. capsularis*. Như nêu trong phần mô tả trên đây, cây *C. olitorius* được sử dụng là giống O-4, và *C. capsularis* được sử dụng là giống CVL-1.

Phương án khác của sáng chế đề xuất cấu trúc gen tái tổ hợp bao gồm và/hoặc có chứa polynucleotit có trình tự nucleotit nêu trong SEQ ID NO: 2 và/hoặc SEQ ID NO: 4, trong đó polynucleotit có thể biểu hiện được ở tế bào chủ, và có thể dịch mã để sản xuất chất tương đồng của protein khóa kéo homeobox-loxin HAT22 ở cây trồng *C. olitorius* và *C. capsularis*. Quy trình khuếch đại, tách dòng và đọc trình tự khóa kéo homeobox-loxin HAT22 từ cây *C. olitorius* và *C. capsularis* được mô tả chi tiết hơn ở ví dụ 2. Theo các phương án nhất định, cấu trúc gen tái tổ hợp còn có chứa vùng trình tự mở đầu được liên kết có thể điều khiển được để tăng cường sự biểu hiện của khuôn polynucleotit. Sau đó, trong sự kiểm soát phiên mã của trình tự mở đầu cụ thể, sự biểu hiện của vùng mã hóa ở trong cấu trúc gen tái tổ hợp có chứa polynucleotit theo SEQ ID NO: 2 và/hoặc SEQ ID NO: 4 có thể được tăng cường, dẫn đến năng suất cao hơn của protein khóa kéo homeobox-loxin HAT22.

Theo một khía cạnh của sáng chế, sự biểu hiện được điều biến là sự biểu hiện hoặc hoạt tính tăng, ví dụ sự biểu hiện quá mức của phân tử axit nucleic mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22, ví dụ phân tử axit nucleic mã hóa cho SEQ ID NO: 2 và SEQ ID NO: 5. Phương pháp làm tăng sự biểu hiện của axit nucleic hoặc gen, hoặc sản phẩm gen, đã được ghi nhận trong tài liệu trong lĩnh vực và các ví dụ được nêu trong phần định nghĩa.

Sáng chế cũng đề xuất phương pháp sản xuất cây trồng chuyển gen có chiều dài sợi và năng suất được tăng cường so với cây trồng đối chứng, bao gồm và/hoặc có chứa bước đưa vào và biểu hiện ở cây trồng axit nucleic bất kỳ mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22 như được xác định ở trên trong bản mô tả này.

Cụ thể hơn, sáng chế đề xuất phương pháp sản xuất cây trồng chuyển gen có năng suất sợi được tăng cường so với cây trồng đối chứng trồng, mà phương pháp này bao gồm bước:

(i) đưa vào và biểu hiện ở cây trồng hoặc tế bào thực vật axit nucleic mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22 hoặc cấu trúc di truyền bao gồm và/hoặc có chứa axit nucleic mã hóa cho polypeptit khóa kéo homeobox-loxin HAT22; và

(ii) cấy tế bào thực vật trong điều kiện thúc đẩy sự sinh trưởng và sự phát triển tế bào sợi.

Khía cạnh khác của sáng chế đề cập đến polynucleotit phân lập được mà mã hóa cho protein khóa kéo homeobox-loxin HAT22 và có nguồn gốc từ cây trồng *C. olitorius*, có chứa phân tử axit nucleic được chọn từ nhóm gồm có:

a) phân tử axit nucleic bao gồm trình tự nucleotit nêu trong SEQ ID NO: 2, hoặc dạng bổ sung của chúng; và

b) phân tử axit nucleic bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 2, hoặc dạng bổ sung của chúng.

Theo các phương án nhất định, cây *C. olitorius* là giống O-4.

Khía cạnh khác của sáng chế đề cập đến polynucleotit phân lập được mà mã hóa cho protein khóa kéo homeobox-loxin HAT22 và có nguồn gốc từ cây trồng *C. capsularis*, có chứa phân tử axit nucleic được chọn từ nhóm gồm có:

a) phân tử axit nucleic bao gồm trình tự nucleotit nêu trong SEQ ID NO: 5, hoặc dạng bổ sung của chúng; và

b) phân tử axit nucleic bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 5, hoặc dạng bổ sung của chúng.

Theo các phương án nhất định, cây *C. capsularis* là giống CVL-1.

Khía cạnh khác của sáng chế đề cập đến polypeptit khóa kéo homeobox-loxin HAT22 được phân lập có chứa trình tự axit amin như nêu trong SEQ ID NO: 3, hoặc mảnh có hoạt tính sinh học của chúng, polypeptit này có một hoặc nhiều chức năng được chọn từ nhóm gồm có xúc tác cho sự khởi đầu, sự tạo thành, sự cải thiện, sự tăng cường, và sự thay đổi, để bằng cách đó cải biến chiều dài sợi, chiều cao cây, sinh khối, hoặc dạng kết hợp bất kỳ của chúng ở cây trồng *C. olitorius*.

Theo các phương án nhất định, cây *C. olitorius* là giống O-4.

Theo các phương án nhất định, polypeptit này có độ tương đồng trình tự ít nhất là 95% so với trình tự axit amin nêu trong SEQ ID NO: 3.

Khía cạnh khác của sáng chế đề cập đến protein polypeptit khóa kéo homeobox-loxin HAT22 được phân lập có chứa trình tự axit amin như nêu trong SEQ

ID NO: 6, hoặc mảnh có hoạt tính sinh học của chúng, polypeptit này có một hoặc nhiều chức năng được chọn từ nhóm gồm có xúc tác cho sự khởi đầu, sự tạo thành, sự cải thiện, sự tăng cường, và sự thay đổi, để bằng cách đó cải biến chiều dài sợi, chiều cao cây, sinh khối, hoặc dạng kết hợp bất kỳ của chúng ở cây trồng *C. capsularis*.

Theo các phương án nhất định, cây *C. capsularis* là giống CVL-1.

Theo các phương án nhất định, polypeptit này có độ tương đồng trình tự ít nhất là 95% so với trình tự axit amin nêu trong SEQ ID NO: 6.

Khía cạnh khác của sáng chế đề cập đến cấu trúc gen tái tổ hợp bao gồm polynucleotit được chọn từ nhóm gồm có:

a) phân tử axit nucleic bao gồm trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; và

b) phân tử axit nucleic bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; trong đó polynucleotit có thể biểu hiện được ở tế bào chủ để sản xuất chất tương đồng của polypeptit khóa kéo homeobox-loxin HAT22 ở cây trồng *C. olitorius* và *C. capsularis*.

Theo các phương án nhất định, cấu trúc gen tái tổ hợp còn có chứa vùng trình tự mở đầu được liên kết có thể điều khiển được với phân tử axit nucleic nêu trong a) hoặc b), trong đó trình tự mở đầu này làm tăng cường sự phiên mã hoặc sự biểu hiện của polynucleotit.

Khía cạnh khác của sáng chế đề cập đến thể biến nạp bao gồm cấu trúc gen tái tổ hợp có khả năng biểu hiện polynucleotit được chọn từ nhóm gồm có:

a) phân tử axit nucleic bao gồm trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; và

b) phân tử axit nucleic bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; trong đó thể biến nạp này sản xuất chất tương đồng của polypeptit khóa kéo homeobox-loxin HAT22.

Khía cạnh khác của sáng chế đề cập đến phương pháp sản xuất cây trồng hoặc cây trồng chuyển gen bao gồm một hoặc nhiều đặc tính được chọn từ nhóm gồm có năng suất sợi, chiều dài sợi, và chiều cao cây được làm tăng, được tăng cường, hoặc được cải biến, so với cây trồng đối chứng, phương pháp này bao gồm các bước:

a) đưa vào tế bào thực vật cấu trúc gen tái tổ hợp bao gồm polynucleotit được chọn từ nhóm gồm có: i) phân tử axit nucleic bao gồm trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; và ii) phân tử axit nucleic bao gồm trình tự nucleotit có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc SEQ ID NO: 5, hoặc dạng bổ sung của chúng; và

b) biểu hiện polypeptit được chọn từ nhóm gồm có: i) polypeptit bao gồm trình tự axit amin nêu trong SEQ ID NO: 3 hoặc SEQ ID NO: 6, hoặc mảnh có hoạt tính sinh học của chúng; và ii) polypeptit có độ tương đồng trình tự ít nhất là 95% so với trình tự axit amin nêu trong SEQ ID NO: 3 hoặc SEQ ID NO: 6; và

c) cấy tế bào thực vật trong điều kiện để thúc đẩy sự sinh trưởng và sự phát triển của cây trồng.

Các trình tự được đề xuất bởi sáng chế cũng có thể được sử dụng làm nguyên liệu chuẩn bị để cải biến hoặc thiết kế kết hợp lý enzym mới có các đặc tính mà làm cho enzym có khả năng hoạt động tốt hơn trong các quy trình khó khăn.

Bản mô tả này bao gồm nội dung như nêu trong các yêu cầu bảo hộ kèm theo, cũng như là trong phần mô tả trên đây. Mặc dù sáng chế này được mô tả ở các phương án ưu tiên của nó ở mức độ cụ thể nhất định, cần hiểu rằng việc bộc lộ các phương án ưu tiên được tạo ra chỉ bằng cách lấy ví dụ và nhiều thay đổi về chi tiết của cấu trúc và dạng kết hợp và sự sắp xếp của các phần có thể được viện đến mà không nằm ngoài phạm vi của sáng chế và yêu cầu bảo hộ.

Nhiều tài liệu tham khảo khác nhau được viện dẫn trong bản mô tả này, việc viện dẫn này hoàn toàn chỉ nhằm mục đích tham khảo.

Ví dụ thực hiện sáng chế

Các ví dụ sau đây nhằm để minh họa thêm cho sáng chế, mà không dự định làm sáng chế bị giới hạn ở các phương án cụ thể được mô tả trong bản mô tả này.

Ví dụ 1: Thiết kế và tổng hợp đoạn mồi

Các đoạn mồi được sử dụng trong nghiên cứu này được thiết kế từ hệ phiên mã được điều chỉnh bằng tay và sáng chế bị giới hạn ở các phương án cụ thể của *C. olitorius* và *C. capsularis*, bằng cách lựa chọn các trình tự bằng tay có ORF hoàn chỉnh hoặc bằng cách sử dụng cơ sở dữ liệu trong đó gen tương tự đã được phân lập thành công từ cây trồng khác. Phân tích thông tin sinh học so sánh của các trình tự nucleotit thu được từ hệ phiên mã được thực hiện bằng cách sử dụng NCBI BLAST, BLASTP, RPS-BLAST, BLASTX và PSI-BLAST để nhận diện chất tương đồng của gen liên quan và để nhận diện thích hợp đối với gen. Sự sắp thẳng hàng trình tự nucleotit được thực hiện bằng clustalW phiên bản 1,82 bất cứ khi nào các trình tự phức hợp được tìm thấy từ “bể gen”. Sau đó, sự sắp thẳng hàng được hiệu chỉnh. Đoạn mồi đặc hiệu gen (cả xuôi và ngược) được chọn lọc bằng tay hoặc thông qua công cụ cộng Đoạn mồi 3 và đoạn mồi được tổng hợp tùy chỉnh.

Tất cả các oligonucleotit được sử dụng trong nghiên cứu này được tổng hợp và được tinh chế HPLC bởi nhà cung cấp và được mua của hãng Integrated DNA Technologies (IDT). Dung dịch dự trữ khoảng 100 pmol được điều chế trong ddH₂O đã được thanh trùng và được lưu trữ ở nhiệt độ khoảng -20°C, trong các phân phân ước để sử dụng.

Các trình tự oligonucleotit được dùng làm đoạn mồi để PCR

Tên đoạn mồi	SEQ ID NO	Trình tự oligonucleotit	Sản phẩm được khuếch đại từ cADN
COLHLZ F	1	TCTTGCTCGCTACCAACATT	883
COLHLZ R	1	GCACTGTTCCTATGGCCTCA	
CCAHLZ F	4	ACCCATTTTCCTTGTTTCCTCCTT	926
CCAHLZ R	4	GCACTGTTCCTATGGCCTCAT	

Ví dụ 2: Khuếch đại, tách dòng và đọc trình tự của khóa kéo homeobox-loxin HAT22 từ *C. olitorius* và *C. capsularis*

ARN tổng số được phân lập từ hạt ba ngày tuổi sinh trưởng trên môi trường MS như đã được mô tả từ trước bởi Chomczynski P và Sacchi N, phương pháp một bước để phân lập ARN bằng cách chiết axit guanidinium thioxyanat-phenol-clorofom. (Anal Biochem 1987, 162: 156-159). Chất lượng hoặc tính toàn vẹn của ARN được kiểm tra bằng sắc ký trên gel agarosa và được định lượng bằng cách sử dụng Thermo Scientific Nano Drop 2000 theo quy trình tiêu chuẩn. Mạch thứ nhất cADN được tổng hợp bằng cách sử dụng transcriptaza ngược SuperScript III (Invitrogen) theo hướng dẫn của nhà sản xuất. Gen được khuếch đại từ cADN bằng PCR bằng cách sử dụng đoạn môi đặc hiệu gen. Phản ứng PCR (50 μ l) có chứa 1 μ l cADN, 20 pmol của mỗi đoạn môi, 5 μ l 10X chất đệm PCR, 5 μ l hỗn hợp 2,5 mM dNTP và 1,0 đơn vị ADN polymeraza PfuTaq. PCR được tiến hành trong Thermal Cycler (Applied Biosystems) bằng cách sử dụng các điều kiện sau: làm biến tính ban đầu trong thời gian khoảng 5 phút ở nhiệt độ xấp xỉ 95°C sau đó là 35 chu kỳ biến tính ở nhiệt độ xấp xỉ 95°C trong thời gian khoảng 30 giây, ủ ở nhiệt độ xấp xỉ 58°C trong thời gian khoảng 30 giây và kéo dài ở nhiệt độ xấp xỉ 72°C trong thời gian khoảng 1 phút, với sự kéo dài cuối cùng ở nhiệt độ xấp xỉ 72°C trong thời gian khoảng 7 phút. Sản phẩm PCR được phân tích bằng gel agarosa 1% bằng cách sử dụng 1X chất đệm TAE và sản phẩm khuếch đại được rửa giải từ gel bằng cách sử dụng bộ kit chiết gel QIAGEN theo hướng dẫn của nhà sản xuất. Sản phẩm PCR đã tinh sạch được nối vào bộ kit tách dòng pCR®8/GW/TOPO® TA (Invitrogen) và được biến nạp vào tế bào E. coli có đủ khả năng (Invitrogen). Plasmid được phân lập từ các khuẩn lạc giả định bằng cách sử dụng bộ kit QIAprep Spin Miniprep (QIAGEN) sau khi chiếu tia sản xuất cây trồng từ hạt già sinh trưởng trên môi trường MS như đã được mô tả từ trước bởi Choe, đoạn môi đặc hiệu gen và plasmid dương tính được đưa đi đọc trình tự.

Ví dụ 3: Phân tích trình tự

Trình tự nucleotit và trình tự axit amin lần lượt được phân tích bằng các chương trình BLASTN và BLASTP. Các trình tự được báo cáo từ cây trồng khác được sắp thẳng hàng bằng ClustalW. Phân tích phát sinh loài được tiến hành bằng cách sử dụng Neighbour Joining (NJ).

Ví dụ 4: Xây dựng con đường sinh tổng hợp sợi

Vai trò hiển thị xây dựng lại con đường chuyển hóa tự động của khóa kéo homeobox-loxin HAT22 trong sự sinh tổng hợp sợi được xây dựng bằng cách xác định các đồng đẳng thích hợp từ protein được dự đoán *C. olitorus* và *C. capsularis* so với hệ gen *Arbidopsis*. Phản ứng enzym được xúc tác bằng khóa kéo homeobox-loxin HAT22 được mã hóa ở trong hệ gen *C. olitorus* và *C. capsularis* được xây dựng bằng cách sử dụng các phản ứng enzym có sẵn trong cơ sở dữ liệu Resnet-Plant 3.0 đối với Pathway Studio cũng như là từ cơ sở dữ liệu con đường chuyển hóa.

Tài liệu tham khảo được viện dẫn

Tất cả các bằng sáng chế Mỹ, các đơn sáng chế đã công bố Mỹ, và các đơn PCT đã công bố được viện dẫn trong bản mô tả này chỉ nhằm mục đích tham khảo.

Dạng tương đương

Mặc dù một số phương án của sáng chế đã được mô tả và được minh họa trong bản mô tả này, người có hiểu biết trung bình trong lĩnh vực sẽ dễ dàng hình dung ra nhiều phương tiện và/hoặc cấu trúc khác để thực hiện các chức năng và/hoặc thu được kết quả và/hoặc một hoặc nhiều ưu điểm được mô tả trong bản mô tả này, và mỗi sự thay đổi và/hoặc cải biến này được cho là thuộc phạm vi của sáng chế. Chỉ cần sử dụng các thí nghiệm thông thường, người có hiểu biết trung bình trong lĩnh vực vẫn có thể nhận diện, hoặc có khả năng xác định nhiều dạng tương đương đối với các phương án cụ thể của sáng chế được mô tả trong bản mô tả này. Do đó, cần hiểu rằng các phương án nêu trên được thể hiện chỉ nhằm làm ví dụ và, nằm trong phạm vi của các yêu cầu bảo hộ kèm theo và các dạng tương đương của chúng; sáng chế có thể được thực hiện theo các khác với cách được mô tả và được yêu cầu bảo hộ cụ thể.

Các trình tự SEQ ID NO:1 đến SEQ ID NO:6

SEQ ID NO : 1
 CHIỀU DÀI : 1296 bp
 LOẠI : ADN
 TÊN DẤU HIỆU/CHÌA KHÓA : Intron, exon bao gồm UTR 5' 150 bp và UTR 3' 150 bp
 SINH VẬT : Cây đay, *C. olitorius*

CACCCCCCTAATTAACACTCCTTCCTCTCATTTCTCTCTAAGAATTTTTTTTTTAAAAAAATTACAACCCTTTT
 CCTTGTTTCCTCCTTAGCATCAACATTTTCTTGCTCGCTACCAACATTTTCTCAGGTTTGAAGTTTGAAACTACC
 ATGGGTTTTGATGATATTTGTAACACTGGCCTTGGTCTTGGACTAGGCTGCCTTGTGAAGCAAGAAAAGTATTCTG
 CAGTCTGATCATCAGCAGCAAAAGAAGAAGAAATTGTTCTTGAAAGATGATCAGTTTTTCCCATCTCTTACTTTA
 GGTCCATCAGATGACATATACCAATTATCTCCTGCTAATAAAATTCATGGGGAATCAACTGATTTGCACCAACAA
 CAACAAGCCTCTTCTATAAGTGGAGTATCTTCATTTTCTAACTCTAGTGTCAAGAAGGAGAAAGATATTTGTGGT
 GAAGAGGTAGAATTGGAAAGAATGTCTTCAAGGGTCAGTGATGAAGATGATGAAGGCAGCCCCAGGAAGAACTT
 CGGCTTAGCAAAGAACAAGCTGCCATTTTAGAAGACAGCTTCAAAGAACACAGCACTCTCAATCCTGTAAATCAT
 CTTCCTTGTCCTTTTAATATGTTATTTTCATGTTTTATATACAATTTTATTCATATAATCTTAATATTTACTTG
 TTTCTATTGTTAACAGAAACAGAAGCAGGCTTTAGCAGAACATCTGAATCTTAGGCCACGGCAAGTGGAAGTATG
 GTTCCAAAACAGGAGAGCCAGGTTAATTTTCCTTTCAATTTTGACCCTATATTCATAGCTTGGCCTTATATGAGT
 TTATTCCCTTACTTTGGATTCTCTTTTTGAAACAACAGGACAAAGCTGAAACAGACTGAAGTAGATTGTGAGTTA
 CTGAAGAAATGTTGTGAGACATTAACAGAAGAGAACAACGGTTGCAGAAGGAAGTGCAGGAGCTTAAATCATTG
 AAATTGACAGCATCTTATTATATGCAGCTGCCGGCGGCCACCCCTACCATGTGCCCTTCTGTGAGAGGGTTGCC
 GGCGCCGGTGAAGGTCCTTCCAGTACAAGCCCTTTTACAATTGGACAGAAATCTCACTTCTTCAACCCCTTCACT
 CACCCATCTGCAGCATGCTAGGCAACTTGTGTGATGAGGCCATAGGAACAGTGCTATATTTTTGTTCTAGGCTTTA
 TTTATGTATGTCAAAATTATATATATGATCAGTATGGAGTAATTAAATCTTAAAGTATAAGTAGGCTAATATCTT
 TATATTTCTTGTATAAATTAA

SEQ ID NO : 2
 CHIỀU DÀI : 804 bp
 LOẠI : ADN
 SINH VẬT : Cây đay, *C. olitorius*
 TÊN DẤU HIỆU/CHÌA KHÓA : CDS
 VỊ TRÍ : (1) (804)

ATGGGTTTTGATGATATTTGTAACACTGGCCTTGGTCTTGGACTAGGCTGCCTTGTGAAGCAAGAAAAGTATTCTG
 CAGTCTGATCATCAGCAGCAAAAGAAGAAGAAATTGTTCTTGAAAGATGATCAGTTTTTCCCATCTCTTACTTTA
 GGTCCATCAGATGACATATACCAATTATCTCCTGCTAATAAAATTCATGGGGAATCAACTGATTTGCACCAACAA
 CAACAAGCCTCTTCTATAAGTGGAGTATCTTCATTTTCTAACTCTAGTGTCAAGAAGGAGAAAGATATTTGTGGT
 GAAGAGGTAGAATTGGAAAGAATGTCTTCAAGGGTCAGTGATGAAGATGATGAAGGCAGCCCCAGGAAGAACTT
 CGGCTTAGCAAAGAACAAGCTGCCATTTTAGAAGACAGCTTCAAAGAACACAGCACTCTCAATCCTAAACAGAAG

CAGGCTTTAGCAGAACATCTGAATCTTAGGCCACGGCAAGTGGAAAGTATGGTTCCAAAACAGGAGAGCCAGGACA
 AAGCTGAAACAGACTGAAGTAGATTGTGAGTTACTGAAGAAATGTTGTGAGACATTAACAGAAGAGAACAAACGG
 TTGCAGAAGGAACTGCAGGAGCTTAAATCATTGAAATTGACAGCATCTTATTATATGCAGCTGCCGGCGGCCACC
 CTCACCATGTGCCCTTCCTGTGAGAGGGTTGCCGGCGCCGGTGAAGGTCCTTCCAGTACAAGCCCTTTTACAATT
 GGACAGAAATCTCACTTCTTCAACCCCTTCACTCACCCATCTGCAGCATGCTAG

SEQ ID NO : 3
 CHIỀU DÀI : 267
 LOẠI : PRT
 SINH VẬT : Cây đay, *C. olitorius*

MGFDDICNTGLGLGLGCLVKQEKYSQSDHQQQKKKFLKDDQFFPSLTLPSSDDIYQLSPANKIHGESTDLHQQQQASSISG
 VSSFSNSSVKKEKDICGEEVELERMSSRVSEDEDEGSPRKKLRLSKEQAILEDSEFKEHSTLNPKQKQALAEHLNLRPRQVEV
 WFQNRRTKRLKQTEVDCELLKKCCETLTEENKRLQKELQELKSLKLTASYMQLPAAATLTMCPSCERVAGAGEGPSSTSPFT
 IGQKSHFFNPFTHPAAC

SEQ ID NO : 4
 CHIỀU DÀI : 1337 bp
 LOẠI : ADN
 TÊN DẤU HIỆU/CHÌA KHÓA : Intron, exon bao gồm UTR 5' 150 bp và UTR 3' 150 bp
 SINH VẬT : Cây đay, *C. capsularis*

CTCAAACAAGTTCCTAATTAACACTCCTTCCTCTCATTTCTCTCTAAGAAAAAAAAAATACAACCCATTTCTCTT
 GTTTCCTCCTTAGCATCAAGATTAATTTTCTTGCTAGCTACCAACATTTTCTAGGTTTGAAGTTTGAACTACC
 ATGGGTTTTGATGATATTTGTAACACTGGCCTTAATCTTGGACTCGGATGCCTTGTGAAGCAAGAAAAGTATTGC
 CCGTCCGATCATCAGCAGCAAAAGAAGAAGAAATTGTTGTTGAAAGATGATCAGTTTTTCCCATCTCTTACTTTA
 GGTCCATCAGATGACATATACCAATTATCTGCTGCTAATAAATTAGCTCATGGGGAATCAATTGATTTGCACCAA
 CAACAACAAGCCTCTTCTATCAGTGGAGTATCTTCATTTTCTAAGTCTAGTGTCAAGAAGGAGAAAGATATTTGT
 GGTGAAGAGGTAGAATTGGAGAGAATGTCTTCAAGGGTCAGTGATGAAGATGATGAAGGCAGCCCCAGGAAGAAA
 CTTTCGGCTTAGCAAAGAACAAGCTGCCATTTTAGAAGACAGCTTCAAAGAACACAGCACTCTCAATCCTGTAAAT
 CATCTTCTAATTAATTCCTTGTCTTTTTCTATATCATTTTCATGTTTTATATTAATTTTATTGCACATAATCTTA
 ATATTTACTTGTTTCTCATGTTAACAGAAACAGAAGCAGGCTTTAGCAGAACATCTGAATCTTAGGCCACGGCAA
 GTGGAAGTATGGTTCCAAAACAGGAGAGCCAGGTTAATTTTCTTTTCAATTTTGACCATATATTATATACATAGC
 TTGGCCTTGTATGATTTTATATTATTCTTCTTACTTTGGATTCTCTTTTTGAAACAACAGGACAAAGCTGAAACAG
 ACTGAAGTAGATTGTGAGTTACTGAAGAAATGTTGTGAGACATTAACAGAAGAGAACAACGGTTGCAGAAGGAA
 CTGCAGGAGCTTAAATCATTGAAATTGACAACATCTTATTATATGCAGCTGCCGGCTGCCACCCTCACCATGTGC
 CCTTCCTGTGAGAGGGTTGCCGGCACCGGTGAAGGCCCTTCCAGTACAAGCCCTTTTACAATTGGACAGAAATCT
 CACTTCTTCAACCCCTTCACTCACCCATCAGCCGCATGCTAGGCAACTTGTTGATGAGGCCATAGGAACAGTGCT

ATTTTTGTTCTAGGCTTTTTATTTATGTCAAATATATATATGATCAGTATGGAGTAATTAAATCCTGAAGTATTA
TAAGTAGGCTAATATCTTTATATTTCTTGTATAAAATTAAGAT

SEQ ID NO : 5
CHIỀU DÀI : 807
LOẠI : ADN
SINH VẬT : Cây đay, *C. capsularis*
TÊN DẤU HIỆU/CHÌA KHÓA : CDS
VỊ TRÍ : (1) (807)

ATGGGTTTTGATGATATTTGTAACACTGGCCTTAATCTTGGACTCGGATGCCTTGTGAAGCAAGAAAAGTATTCGCCGTCCGA
TCATCAGCAGCAAAAGAAGAAGAAATTGTTGTTGAAAGATGATCAGTTTTTCCCATCTCTTACTTTAGGTCCATCAGATGACA
TATACCAATTATCTGCTGCTAATAAATTAGCTCATGGGGAATCAATTGATTTGCACCAACAACAACAAGCCTCTTCTATCAGT
GGAGTATCTTCATTTTCTAACTCTAGTGTCAAGAAGGAGAAAGATATTTGTGGTGAAGAGGTAGAATTGGAGAGAATGTCTTC
AAGGGTCAGTGATGAAGATGATGAAGGCAGCCCCAGGAAGAACTTCGGCTTAGCAAAGAACAAGCTGCCATTTTAGAAGACA
GCTTCAAAGAACACAGCACTCTCAATCCTAAACAGAAGCAGGCTTTAGCAGAACATCTGAATCTTAGGCCACGGCAAGTGGAA
GTATGGTTCCAAAACAGGAGAGCCAGGACAAAGCTGAAACAGACTGAAGTAGATTGTGAGTTACTGAAGAAATGTTGTGAGAC
ATTAACAGAAGAGAACAACGGTTGCAGAAGGAAGTGCAGGAGCTTAAATCATTGAAATTGACAACATCTTATTATATGCAGC
TGCCGGCTGCCACCCTCACCATGTGCCCTTCCTGTGAGAGGGTTGCCGGCACCGGTGAAGGCCCTTCCAGTACAAGCCCTTTT
ACAATTGGACAGAAATCTCACTTCTTCAACCCTTCACTCACCCATCAGCCGCATGCTAG

SEQ ID NO : 6
CHIỀU DÀI : 268
LOẠI : PRT
SINH VẬT : Cây đay, *C. capsularis*

MGFDDICNTGLNLGLGCLVKQEKYSPSDHQQQKKKLLKDDQFFPSLTLGPSDDIYQLSAANKLAHGESIDLHQQQQASSIS
GVSSFSNSSVKKEKDICGEEVELERMSSRVSDDEDEGSPRKKLRLSKEQAAILLEDSFKEHSTLNPQKQALAEHLNLRPRQVE
VWFQNRRTKRLKQTEVDCELLKKCCETLTEENKRLQKELQELKSLKLTTSSYMQLPAAATLTMCPSCERVAGTGEGPSSTSPF
TIGQKSHFFNPFTHPAAC

YÊU CẦU BẢO HỘ

1. Polynucleotit bao gồm cADN có độ tương đồng trình tự ít nhất là 95% so với trình tự nucleotit nêu trong SEQ ID NO: 2 hoặc 5, trong đó polynucleotit này mã hóa protein HAT22 khoá kéo homeobox-loxin.
2. Polynucleotit theo điểm 1, trong đó cADN có độ tương đồng trình tự ít nhất là 98% so với trình tự nucleotit nêu trong SEQ ID NO: 2.
3. Polynucleotit theo điểm 1, trong đó cADN có độ tương đồng trình tự ít nhất là 98% so với trình tự nucleotit nêu trong SEQ ID NO: 5.
4. Cấu trúc gen tái tổ hợp bao gồm polynucleotit theo điểm 1, trong đó polynucleotit có thể biểu hiện được ở tế bào chủ để sản xuất polypeptit HAT22 khoá kéo homeobox-loxin ở *Corchorus olitorius* và *Corchorus capsularis*.
5. Cấu trúc gen tái tổ hợp theo điểm 4, còn bao gồm vùng trình tự mở đầu được liên kết có thể điều khiển được với phân tử axit nucleic được nêu trong a) hoặc b), trong đó trình tự mở đầu này làm tăng cường sự phiên mã hoặc biểu hiện của polynucleotit.
6. Thẻ biến nạp chứa cấu trúc gen tái tổ hợp theo điểm 4, trong đó thẻ biến nạp này có khả năng biểu hiện polynucleotit và thẻ biến nạp này tạo ra polypeptit HAT22 khoá kéo homeobox-loxin.
7. Phương pháp sản xuất cây trồng chuyển gen, phương pháp này bao gồm các bước:
 - a) đưa vào tế bào thực vật cấu trúc gen tái tổ hợp theo điểm 4;
 - b) biểu hiện polypeptit có độ tương đồng trình tự ít nhất là 95% so với trình tự axit amin nêu trong SEQ ID NO: 3 hoặc SEQ ID NO: 6; và

c) cấy tế bào thực vật trong các điều kiện để thúc đẩy sự sinh trưởng và phát triển của cây trồng.

8. Phương pháp theo điểm 7, trong đó polypeptit này có độ tương đồng trình tự ít nhất là 98% so với trình tự axit amin này.

9. Phương pháp theo điểm 7, trong đó polypeptit này có độ tương đồng trình tự ít nhất là 99% so với trình tự axit amin này.

10. Cấu trúc biểu hiện bao gồm polynucleotit theo điểm 1, trong đó polynucleotit này được liên kết có thể điều khiển được với phân tử axit nucleic điều hòa có chứa các tín hiệu điều hòa phiên mã hoặc dịch mã, hoặc cả hai, kiểm soát sự biểu hiện của polynucleotit trong tế bào chủ.

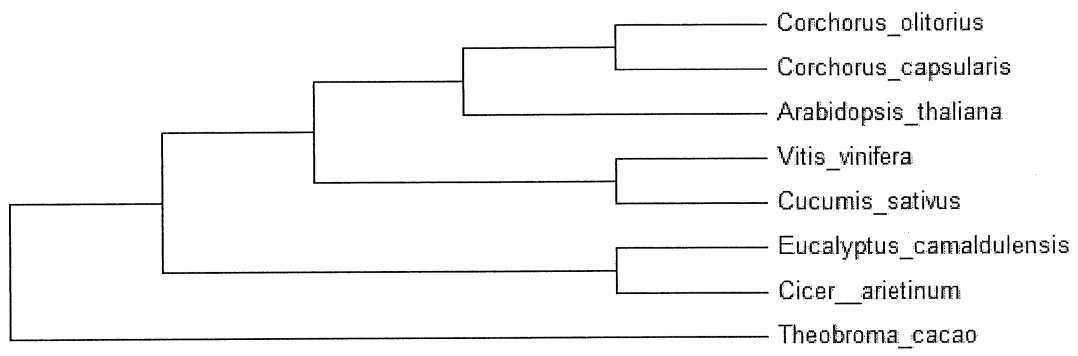
11. Chất được phân lập từ thể biến nạp theo điểm 6, trong đó chất này bao gồm polynucleotit.

12. Hạt từ thể biến nạp theo điểm 6, trong đó hạt này bao gồm polynucleotit.

13. Phương pháp gây ra, khởi tạo, cải thiện hoặc tăng cường sự phát triển của sợi thực vật và chiều cao của thực vật ở cây đay, phương pháp này bao gồm bước đưa vào cây đay polynucleotit phân lập được theo điểm 1.

14. Polynucleotit theo điểm 1, trong đó cADN nêu trên có độ tương đồng trình tự là 100% so với trình tự nucleotit này.

15. Phương pháp theo điểm 7, trong đó polypeptit này có độ tương đồng trình tự là 100% so với axit amin này.

**Fig. 1**

DANH MỤC TRÌNH TỰ

<110> BANGLADESH JUTE RESEARCH INSTITUTE

<120> POLYNUCLEOTIT MÃ HÓA PROTEIN KHÓA KÉO HOMEBOX-LOXIN HAT22 (PROTEIN HD-ZIP 22) CÓ NGUỒN GỐC TỪ HAI LOÀI CÂY *Corchorus olitorius* VÀ *Corchorus capsularis* VÀ PHƯƠNG PHÁP SẢN XUẤT CÂY TRỒNG CHUYỂN GEN BIỂU HIỆN POLYNUCLEOTIT NÀY

<130> JGX-009.25

<140> PCT/US2014/066772

<141> 2014-11-21

<150> 61/908,444

<151> 2013-11-25

<160> 10

<170> PatentIn phiên bản 3.5

<210> 1

<211> 1296

<212> ADN

<213> *Corchorus olitorius*

<400> 1

caccccccta attaacactc ctctctctca ttctctctct aagaattttt tttaaaaaa 60

aattacaacc ctttccttg ttctctctt agcatcaaca tttcttgct cgctaccaac 120

atctctcag gttgaagtt tgaaactacc atgggttttg atgatattg taacttggc 180

cttggtcttg gactaggctg cttgtgaag caagaaaagt atcgagtc tgatcatcag 240

cagcaaaaga agaagaaatt gtcttgaaa gatgatcagt tttcccatc tcttacttta 300

ggatcatcag atgacatata ccaattatct cctgctaata aaatcatgg ggaatcaact 360

gattgcacc aacaacaaca agcctctctc ataagtggag tatcttcatt tttaactct 420

agtgtaaga aggagaaaga tattgtggt gaagaggtag aattggaaag aatgtctca 480

agggtcagtg atgaagatga tgaaggcagc cccaggaaga aacttcggct tagcaaagaa 540
 caagctgcc tttagaaga cagcttcaaa gaacacagca ctctcaatcc tgtaaactat 600
 ctcccttgt ccttttaata tgtatttca tgtttatat acaattttat tgcataaat 660
 cttaatattt acttgtttct attgttaaca gaaacagaag caggctttag cagaacatct 720
 gaactctagg ccacggcaag tggaagtatg gttccaaaac aggagagcca ggttaatttt 780
 cctttcaatt ttgacctat attcatagct tggecttata tgagtttatt cccttacttt 840
 ggattctctt ttgaaacaa caggacaaag ctgaaacaga ctgaagtaga ttgtgagtta 900
 ctgaagaaat gttgtgagac attaacagaa gagaacaaac ggttgcagaa ggaactgcag 960
 gagcttaaat cattgaaatt gacagcatct tattatatgc agctgccggc ggccaccctc 1020
 accatgtgcc ctctctgtga gaggggtgcc ggcgccgggtg aaggctctc cagtacaagc 1080
 ccttttaca ttgacagaa atctcacttc ttcaaccct tcaactaccc atctgcagca 1140
 tgctaggcaa cttgtgatg aggccatagg aacagtgcta ttttttgtt ctaggcttta 1200
 tttatgtatg tcaaaattat atatatgatc agtatggagt aattaaatct taaagtataa 1260
 gtaggcta atctttatat ttcttgata aattaa 1296

<210> 2

<211> 804

<212> ADN

<213> Corchorus olitorius

<400> 2

atgggttttg atgatattg taacactggc ctgggtcttg gactaggctg cttgtgaag 60
 caagaaaagt attcgcagtc tgatcatcag cagcaaaaga agaagaaatt gttcttgaaa 120
 gatgatcagt tttcccatc tcttacttta ggtccatcag atgacatata ccaattatct 180
 cctgctaata aaattcatgg ggaatcaact gatttgcacc aacaacaaca agcctcttct 240
 ataagtggag tatcttcatt ttctaactct agtgtcaaga aggagaaaga tatttgtggt 300
 gaagaggtag aattggaaag aatgtcttca agggtcagtg atgaagatga tgaaggcagc 360
 cccaggaaga aacttcggct tagcaaagaa caagctgcc tttagaaga cagcttcaaa 420

gaacacagca ctctcaatcc taaacagaag caggctttag cagaacatct gaatcttagg 480
 ccacggcaag tggaagtatg gttccaaaac aggagagcca ggacaaagct gaaacagact 540
 gaagtagatt gtgagttact gaagaaatgt tgtgagacat taacagaaga gaacaaacgg 600
 ttgcagaagg aactgcagga gcttaaatca ttgaaattga cagcatctta ttatatgcag 660
 ctgccggcgg ccaccctcac catgtgccct tcctgtgaga gggttgccgg cgccggtgaa 720
 ggtccttcca gtacaagccc tttacaatt ggacagaaat ctcacttctt caacccttc 780
 actcacccat ctgcagcatg ctag 804

<210> 3

<211> 267

<212> PRT

<213> *Corchorus olitorius*

<400> 3

Met Gly Phe Asp Asp Ile Cys Asn Thr Gly Leu Gly Leu Gly Leu Gly
 1 5 10 15

Cys Leu Val Lys Gln Glu Lys Tyr Ser Gln Ser Asp His Gln Gln Gln
 20 25 30

Lys Lys Lys Lys Leu Phe Leu Lys Asp Asp Gln Phe Phe Pro Ser Leu
 35 40 45

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

Ile His Gly Glu Ser Thr Asp Leu His Gln Gln Gln Gln Ala Ser Ser
 65 70 75 80

Ile Ser Gly Val Ser Ser Phe Ser Asn Ser Ser Val Lys Lys Glu Lys
 85 90 95

Asp Ile Cys Gly Glu Glu Val Glu Leu Glu Arg Met Ser Ser Arg Val
 100 105 110

Ser Asp Glu Asp Asp Glu Gly Ser Pro Arg Lys Lys Leu Arg Leu Ser
 115 120 125

Lys Glu Gln Ala Ala Ile Leu Glu Asp Ser Phe Lys Glu His Ser Thr
 130 135 140

Leu Asn Pro Lys Gln Lys Gln Ala Leu Ala Glu His Leu Asn Leu Arg
 145 150 155 160

Pro Arg Gln Val Glu Val Trp Phe Gln Asn Arg Arg Ala Arg Thr Lys
 165 170 175

Leu Lys Gln Thr Glu Val Asp Cys Glu Leu Leu Lys Lys Cys Cys Glu
 180 185 190

Thr Leu Thr Glu Glu Asn Lys Arg Leu Gln Lys Glu Leu Gln Glu Leu
 195 200 205

Lys Ser Leu Lys Leu Thr Ala Ser Tyr Tyr Met Gln Leu Pro Ala Ala
 210 215 220

Thr Leu Thr Met Cys Pro Ser Cys Glu Arg Val Ala Gly Ala Gly Glu
 225 230 235 240

Gly Pro Ser Ser Thr Ser Pro Phe Thr Ile Gly Gln Lys Ser His Phe
 245 250 255

Phe Asn Pro Phe Thr His Pro Ser Ala Ala Cys
 260 265

<210> 4

<211> 1317

<212> ADN

<213> *Corchorus capsularis*

<400> 4

ctcaaacaag ttctaatta acactccttc ctctcatttc ctctctaaga aaaaaaaaaat 60
 acaaccatt tcttgtttc ctcttagca tcaagattaa tttcttget agtaccac 120
 attttctag gttgaagtt tgaaactacc atgggttttg atgatattg taactggc 180
 cttaatcttg gactcggatg ccttgtgaag caagaaaagt attcgccgtc cgatcatcag 240
 cagcaaaaga agaagaaatt gttgtgaaa gatgatcagt tttcccatc tcttacttta 300
 ggtccatcag atgacatata ccaattatct gctgctaata aattagctca tggggaatca 360
 attgattgc accaacaaca acaagcctct tctatcagtg gagtatcttc atttctaac 420
 tctagtgtca agaaggagaa agatattgt ggtgaagagg tagaattgga gagaatgtct 480
 tcaagggtca gtgatgaaga tgatgaaggc agccccagga agaaactcg gcttagcaaa 540
 gaacaagctg ccattttaga agacagcttc aaagaacaca gcactctcaa tcctgtaa 600
 catcttctaa ttaattccct tgtcttttcc tatatcatt catgtttat attaatTTTA 660
 ttgcacataa tcttaatat tcttgtttc tcatgttaac agaaacagaa gcaggttta 720
 gcagaacatc tgaatcttag gccacggcaa gtggaagtat ggttcaaaa caggagagcc 780
 aggttaattt tccttcaat ttgaccata tattatatac atagcttggc cttgtatgat 840
 tttatattat tctcttactt tggattctct tttgaaaca acaggacaaa gctgaaacag 900
 actgaagtag attgtgagtt actgaagaaa tgttgtgaga cattaacaga agagaacaaa 960
 cgggtgcaga aggaactgca ggagcttaaa tcattgaaat tgacaacatc ttattatag 1020
 cagctgccgg ctgccacct caccatgtgc cttctctgtg agagggttgc cggcaccggt 1080
 gaaggccctt ccagtacaag ccttttaca attggacaga aatctcactt ctcaacccc 1140
 ttcactcacc catcagccgc atgctaggca actgttgat gaggccatag gaacagtgt 1200
 attttgttc taggttttt atttatgtca aatatatata tgatcagtat ggagtaatta 1260
 aatcctgaag tattataagt aggctaatat ctttatattt cttgtataaa ttaagat 1317

<210> 5

<211> 807

<212> ADN

<213> *Corchorus capsularis*

<400> 5

atgggttttg atgatatttg taacactggc cttaatcttg gactcggatg ccttgtgaag 60
 caagaaaagt attcgccgtc cgatcatcag cagcaaaaga agaagaaatt gttgttgaaa 120
 gatgatcagt tttcccatc tcttacttta ggtccatcag atgacatata ccaattatct 180
 gctgctaata aattagctca tggggaatca attgattgc accaacaaca acaagcctct 240
 tctatcagtg gagtatcttc attttctaac tctagtgtca agaaggagaa agatatttgt 300
 ggtgaagagg tagaattgga gagaatgtct tcaagggtca gtgatgaaga tgatgaaggc 360
 agccccagga agaaacttcg gcttagcaaa gaacaagctg ccattttaga agacagcttc 420
 aaagaacaca gcactctcaa tctaaacag aagcaggctt tagcagaaca tctgaatctt 480
 aggccacggc aagtggaagt atggttccaa aacaggagag ccaggacaaa gctgaaacag 540
 actgaagtag attgtgagtt actgaagaaa tgttgtgaga cattaacaga agagaacaaa 600
 cggttgcaga aggaactgca ggagcttaaa tcattgaaat tgacaacatc ttattatatg 660
 cagctgccgg ctgccaccct caccatgtgc ccttctgtg agagggttgc cggcaccggt 720
 gaaggccctt ccagtacaag cccttttaca attggacaga aatctcactt cttcaacccc 780
 ttcactcacc catcagccgc atgctag 807

<210> 6

<211> 268

<212> PRT

<213> *Corchorus capsularis*

<400> 6

Met Gly Phe Asp Asp Ile Cys Asn Thr Gly Leu Asn Leu Gly Leu Gly
 1 5 10 15

Cys Leu Val Lys Gln Glu Lys Tyr Ser Pro Ser Asp His Gln Gln Gln
 20 25 30

Lys Lys Lys Lys Leu Leu Leu Lys Asp Asp Gln Phe Phe Pro Ser Leu
 35 40 45

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

Leu Ala His Gly Glu Ser Ile Asp Leu His Gln Gln Gln Gln Ala Ser
 65 70 75 80

Ser Ile Ser Gly Val Ser Ser Phe Ser Asn Ser Ser Val Lys Lys Glu
 85 90 95

Lys Asp Ile Cys Gly Glu Glu Val Glu Leu Glu Arg Met Ser Ser Arg
 100 105 110

Val Ser Asp Glu Asp Asp Glu Gly Ser Pro Arg Lys Lys Leu Arg Leu
 115 120 125

Ser Lys Glu Gln Ala Ala Ile Leu Glu Asp Ser Phe Lys Glu His Ser
 130 135 140

Thr Leu Asn Pro Lys Gln Lys Gln Ala Leu Ala Glu His Leu Asn Leu
 145 150 155 160

Arg Pro Arg Gln Val Glu Val Trp Phe Gln Asn Arg Arg Ala Arg Thr
 165 170 175

Lys Leu Lys Gln Thr Glu Val Asp Cys Glu Leu Leu Lys Lys Cys Cys
 180 185 190

Glu Thr Leu Thr Glu Glu Asn Lys Arg Leu Gln Lys Glu Leu Gln Glu
 195 200 205

Leu Lys Ser Leu Lys Leu Thr Thr Ser Tyr Tyr Met Gln Leu Pro Ala
 210 215 220

Ala Thr Leu Thr Met Cys Pro Ser Cys Glu Arg Val Ala Gly Thr Gly
 225 230 235 240

Glu Gly Pro Ser Ser Thr Ser Pro Phe Thr Ile Gly Gln Lys Ser His
 245 250 255

Phe Phe Asn Pro Phe Thr His Pro Ser Ala Ala Cys
 260 265

<210> 7

<211> 20

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn mỗi tổng hợp

<400> 7

tcttgctcgc taccaacatt 20

<210> 8

<211> 20

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn mỗi tổng hợp

<400> 8

gcactgttcc tatggcctca 20

<210> 9

<211> 23

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn mỗi tổng hợp

<400> 9

accatttcc ttgttcctc ctt 23

<210> 10

<211> 21

<212>ADN

<213> Trình tự nhân tạo

<220>

<223> Mô tả trình tự nhân tạo: đoạn môi tổng hợp

<400> 10

gcactgttcc tatggcctca t

21