



(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)



1-0025356

(51)⁷ C12N 15/09; C12N 1/19; C12N 9/00;
C12N 1/15; C12N 1/21

(13) B

(21) 1-2014-00442

(22) 13/07/2012

(86) PCT/CA2012/000656 13/07/2012

(87) WO 2013/006953 17/01/2013

(30) 61/507,331 13/07/2011 US

(45) 25/09/2020 390

(43) 27/10/2014 319A

(73) 1. National Research Council of Canada (CA)

1200 Montreal Road, Ottawa, Ontario K1A 0R6, Canada

2. University of Saskatchewan (CA)

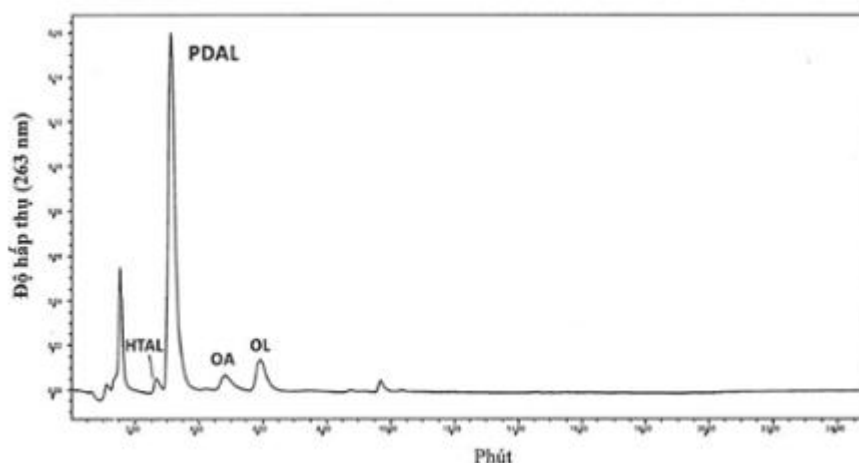
Industry Liaison Office, University of Saskatchewan, Suite 501, 121 Research Drive,
Saskatoon, Saskatchewan, S7N 1K2, Canada

(72) PAGE, Jonathan E. (CA); STOUT, Jason M. (CA).

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

(54) PHÂN TỬ AXIT NUCLEIC, VECTƠ CHỨA PHÂN TỬ AXIT NUCLEIC NÀY,
QUY TRÌNH TỔNG HỢP ALKANOYL-COA VÀ QUY TRÌNH BIẾN ĐỔI HÀM
LƯỢNG CỦA HỢP CHẤT CANABINOIT

(57) Sáng chế đề cập đến phân tử axit nucleic được phân lập hoặc được tinh chế. Ngoài ra, sáng chế còn đề cập đến polypeptit được phân lập, vectơ chứa phân tử axit nucleic, hệ biểu hiện chứa phân tử axit nucleic, cấu trúc chứa phân tử axit nucleic, tế bào chủ được biến nạp, quy trình tổng hợp alkanoyl-CoA, quy trình biến đổi hàm lượng của hợp chất canabinoit và quy trình tổng hợp hợp chất canabinoit.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phân tử axit nucleic và protein có liên quan đến quá trình tổng hợp alkanoyl-CoA thioeste, và các ứng dụng của phân tử axit nucleic và protein này để biến đổi di truyền quá trình sinh tổng hợp cannabinoids ở cây, vi sinh vật hoặc trong hệ thống không có tế bào và để tạo ra cây gai dầu với hàm lượng cannabinoids được tăng cường hoặc làm giảm.

Tình trạng kỹ thuật của sáng chế

Cannabis sativa L. (cannabis, hemp, marijuana) là một trong số các cây trồng được trồng lâu nhất và nhiều tác dụng nhất, mà hiện nay được sử dụng làm nguồn sản phẩm thuốc, thực phẩm, mỹ phẩm và sản phẩm công nghiệp. Cũng đã biết rõ việc sử dụng cây trồng này làm thuốc trái phép do hàm lượng cannabinoids tác động đến tâm thần của nó (ví dụ, Δ^9 -tetrahydrocannabinol, Δ^9 -THC). Cannabinoids và các thuốc khác mà tác động thông qua thụ thể cannabinoids ở động vật có vú đang được nghiên cứu để điều trị các tình trạng bệnh khác nhau như chứng đau mạn tính, bệnh xơ cứng rải rác và bệnh động kinh.

Cannabinoids có nguồn gốc sinh tổng hợp của nó ở cả cơ chế chuyển hóa polyketit và terpenoit và được gọi là terpenophenolic hoặc polyketit được prenyl hóa (Page J., Nagel J. (2006) Biosynthesis of terpenophenolics in hop and cannabis. In JT Romeo, ed, Integrative Plant Biochemistry, Vol. 40. Elsevier, Oxford, pp 179-210.). Quá trình sinh tổng hợp cannabinoids xảy ra chủ yếu trong các túi lông tuyến bao phủ hoa cái ở mật độ cao. Cannabinoids được tạo thành bởi quá trình sinh tổng hợp gồm ba bước: tạo thành polyketit, prenyl hóa thơm và đóng vòng (xem Fig.1).

Bước enzym thứ nhất trong quá trình sinh tổng hợp cannabinoids là sự tạo thành axit olivetolic bởi enzym polyketit syntaza xúc tác sự ngưng tụ của hexanoyl-coenzym (CoA) với ba phân tử malonyl-CoA. Các cannabinoids chính,

kể cả axit Δ^9 -tetrahydrocannabinolic và axit canabidiolic, được tạo thành từ tiền chất hexanoyl-CoA, tiền chất này là axyl-CoA béo mạch trung bình (xem Fig.1). Các cannabinoid khác với các mạch bên biến đổi được tạo thành từ các CoA béo có chiều dài khác nhau (ví dụ, axit Δ^9 -tetrahydrocannabivarinic được tạo thành từ đoạn mỗi n-butyryl-CoA).

Hexanoyl-CoA và các axyl-CoA thioeste khác ở cây được tổng hợp bởi enzym hoạt hóa axyl (AAE, còn được gọi là axyl-CoA synthetaza), enzym này xúc tác sự hoạt hóa của cơ chất axit carboxylic sử dụng ATP. Các enzym này tác động trên nhiều axit cacboxylat khác nhau bao gồm axit béo mạch ngắn, trung bình, dài và rất dài, tiền chất jasmonat, axit có nguồn gốc từ phenylpropanoid (ví dụ, axit xinnamic) và các axit hữu cơ khác như malonat, axetat và xitrat. Rất ít axyl CoA synthetaza mạch trung bình được nhận diện trước đây trong tự nhiên. Ở cây, ba enzym từ *A. thaliana*, AAE7, At4g05160 và At5g63380 đã cho thấy là tạo thành hexanoyl-CoA từ hexanoat (Schneider K et al. (2005) A new type of peroxisomal acyl-coenzyme A synthetase from *Arabidopsis thaliana* has the catalytic capacity to activate biosynthetic precursors of jasmonic acid. The Journal of Biological Chemistry 280:13962-72; Shockey JM, Fulda MS, Browse J (2003) *Arabidopsis* contains a large superfamily of acyl-activating enzymes. Phylogenetic and biochemical analysis reveals new class of acyl-coenzyme a synthetases. Plant Physiology 132:1065-76.) Acyl-CoA synthetaza từ *Pseudomonas* spp. đã cho thấy là tác động trên các axit béo mạch trung bình như hexanoat (Fernandez-Valverde M, Reglero A, Martinez-Blanco H, Luengo JM (1993) Purification of *Pseudomonas putida* acyl coenzyme A ligase active with a range of aliphatic and aromatic substrates. Applied Environmental Microbiology 59:1149-1154.)

Canabinoid là sản phẩm tự nhiên có giá trị. Gen mã hóa cho các enzym có liên quan đến quá trình sinh tổng hợp cannabinoid sẽ hữu dụng trong việc biến đổi di truyền sự trao đổi chất của cây gai dầu để tạo ra cây chứa mức rất thấp, hoặc mức zero, của THCA và các cannabinoid khác thông qua sự phát sinh đột biến có đích (ví dụ, TILLING) hoặc các kỹ thuật bất hoạt gen khác. Các gen

này cũng có thể chứng tỏ là hữu dụng để tạo ra, thông qua sự chọn lọc được hỗ trợ bởi chất đánh dấu, các giống cây gai dầu đặc trưng để sản xuất được phẩm trên cơ sở canabinoit, hoặc để khôi phục lại quá trình sinh tổng hợp canabinoit trong các sinh vật khác loài như vi khuẩn hoặc nấm men, hoặc để sản xuất canabinoit trong hệ thống không có tế bào mà sử dụng các protein tái tổ hợp.

Gen mã hóa cho các enzym của quá trình sinh tổng hợp canabinoit cũng có thể được sử dụng trong tổng hợp các chất tương tự canabinoit và tổng hợp các chất tương tự của tiền chất canabinoit. Các chất tương tự canabinoit trước đây đã được tổng hợp và có thể hữu ích dùng làm sản phẩm được.

Vẫn có nhu cầu trong lĩnh vực kỹ thuật này để nhận diện các enzym có liên quan đến quá trình tổng hợp polyketit thơm, và trình tự nucleotit mã hóa cho các enzym này.

Bản chất kỹ thuật của sáng chế

Hai gen mới từ cây gai dầu hiện nay đã được phát hiện, hai gen này mã hóa alkanoyl-CoA synthetaza chưa được biết trước đây. Hai alkanoyl Co-A synthetaza mới này được gọi là *Cannabis sativa* hexanoyl-CoA synthetaza 1 (CsHCS1) và *Cannabis sativa* hexanoyl-CoA synthetaza 2 (CsHCS2).

Như vậy, theo khía cạnh thứ nhất, sáng chế đề xuất phân tử axit nucleic được phân lập hoặc được tinh chế chứa trình tự nucleotit có ít nhất 75% độ đồng nhất trình tự với SEQ ID NO: 1, hoặc trình tự thoái hóa codon của nó.

Theo khía cạnh thứ hai, sáng chế đề xuất phân tử axit nucleic được phân lập hoặc được tinh chế chứa trình tự nucleotit có ít nhất 75% độ đồng nhất trình tự với SEQ ID NO: 3, hoặc trình tự thoái hóa codon của nó.

Theo khía cạnh thứ ba, sáng chế đề xuất polypeptit được phân lập hoặc được tinh chế chứa trình tự axit amin có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 2, hoặc trình tự axit amin được thể bảo toàn của nó.

Theo khía cạnh thứ tư, sáng chế đề xuất polypeptit được phân lập hoặc được tinh chế chứa trình tự axit amin có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 4, hoặc trình tự axit amin được thể bảo toàn của nó.

Theo khía cạnh thứ năm, sáng chế đề xuất vectơ, cấu trúc hoặc hệ biểu hiện chứa phân tử axit nucleic theo sáng chế.

Theo khía cạnh thứ sáu, sáng chế đề xuất tế bào chủ được biến nạp với phân tử axit nucleic theo sáng chế.

Theo khía cạnh thứ bảy, sáng chế đề xuất quy trình tổng hợp alkanoyl-CoA trong sự có mặt của enzym theo sáng chế.

Theo khía cạnh thứ tám, sáng chế đề xuất quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bao gồm sử dụng phân tử axit nucleic theo sáng chế, hoặc một phần của nó, để làm bất hoạt ở sinh vật, tế bào hoặc mô này gen mã hóa enzym xúc tác sự tổng hợp alkanoyl-CoA.

Theo khía cạnh thứ chín, sáng chế đề xuất quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bao gồm làm đột biến gen ở sinh vật, tế bào hoặc mô này, và sử dụng phân tử axit nucleic theo sáng chế để lựa chọn sinh vật, tế bào hoặc mô chứa các thể đột biến hoặc các biến thể của gen mã hóa enzym xúc tác sự tổng hợp alkanoyl-CoA.

Theo khía cạnh thứ mười, sáng chế đề xuất quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bao gồm biểu hiện hoặc biểu hiện quá mức phân tử axit nucleic theo sáng chế ở sinh vật, tế bào hoặc mô này so với giống sinh vật, tế bào hoặc mô tương tự sinh trưởng trong các điều kiện tương tự nhưng không có sự biểu hiện hoặc sự biểu hiện quá của phân tử axit nucleic này.

Theo khía cạnh thứ mười một, sáng chế đề xuất quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bao gồm biểu hiện hoặc biểu hiện quá mức phân tử axit nucleic mã hóa cho polypeptit theo sáng chế ở sinh vật, tế bào hoặc mô này so với giống sinh vật, tế bào hoặc mô tương tự sinh trưởng trong các điều kiện tương tự nhưng không có sự biểu hiện hoặc sự biểu hiện quá của phân tử axit nucleic này.

Theo khía cạnh thứ mười hai, sáng chế đề xuất quy trình tổng hợp hợp chất canabinoit có trong tự nhiên hoặc chất tương tự không có trong tự nhiên

của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bao gồm biểu hiện phân tử axit nucleic theo sáng chế ở sinh vật, tế bào hoặc mô này trong điều kiện có mặt axit carboxylic và CoA.

Theo khía cạnh thứ mười ba, sáng chế đề xuất quy trình tổng hợp alkanoyl-CoA trong phản ứng không có tế bào *in vitro*, quy trình này bao gồm: cho axit carboxylic với phản ứng coenzym A trong sự có mặt của enzym theo sáng chế.

Các polypeptit là enzym xúc tác sự tổng hợp alkanoyl-CoA, và trình tự nucleotit mã hóa cho các enzym này, hiện nay đã được nhận diện và xác định đặc điểm. Các trình tự nucleotit này có thể được sử dụng để tạo ra, thông qua quá trình nhân giống, chọn lọc hoặc biến đổi di truyền, cây gai dầu sản xuất quá mức hoặc sản xuất dưới mức hợp chất canabinoit, chất tương tự của hợp chất canabinoit hoặc hỗn hợp của chúng. Các trình tự nucleotit này cũng có thể được sử dụng, một mình hoặc kết hợp với các gen mã hóa cho các bước khác trong con đường tổng hợp canabinoit, để biến đổi di truyền quá trình sinh tổng hợp canabinoit ở các cây khác hoặc ở các vi sinh vật (ví dụ, nấm men, vi khuẩn, nấm) hoặc các sinh vật không có nhân điển hình hoặc có nhân điển hình khác hoặc trong hệ thống không có tế bào. Ngoài ra, việc phong bế hoặc làm giảm sự biểu hiện của các gen này trong cây gai dầu có thể được sử dụng để phong bế quá trình sinh tổng hợp canabinoit và bằng cách đó làm giảm sự sản xuất canabinoit.

Các đặc điểm khác nữa của sáng chế sẽ được mô tả hoặc sẽ trở nên rõ ràng từ phần mô tả chi tiết sau đây.

Mô tả vắn tắt các hình vẽ

Để hiểu sáng chế một cách rõ ràng hơn, các phương án của sáng chế bây giờ sẽ được mô tả chi tiết bằng cách ví dụ, dựa vào các hình vẽ kèm theo, trong đó:

Fig.1 mô tả con đường được đề xuất dẫn đến tạo ra các loại canabinoit chính trong *Cannabis sativa*. Các từ viết tắt: THCA syntaza là Δ^9 -

tetrahydrocannabinolic axit syntaza; CBDA syntaza là canabidiolic axit syntaza; CBCA syntaza là canabichromenic axit syntaza.

Fig. 2A-2F mô tả phân tích khối phổ-sắc ký lỏng/khối phổ (LC-MS/MS) hoạt tính enzym của *Cannabis sativa* hexanoyl-CoA syntaza. Mỗi trong số các Fig. 2A-2F cho thấy độ dư ion (m/z 866>359) trên trục tung và thời gian (phút) trên trục hoành. Fig.2A và 2B mô tả thời gian lưu của chuẩn hexanoyl-CoA tin cậy. Fig.2C mô tả thử nghiệm của CsHCS1 protein, CoA, $MgCl_2$, natri hexanoat, ATP, và đệm HEPES, trong đó hexanoyl-CoA được sản xuất và được phát hiện. Fig.2D mô tả thử nghiệm của CsHCS2 protein, CoA, $MgCl_2$, natri hexanoat, ATP, và đệm HEPES, trong đó hexanoyl-CoA được sản xuất. Fig.2E mô tả thử nghiệm trong đó CsHCS1 protein đã được làm bất hoạt trước bằng cách đun sôi ở nhiệt độ 95°C trong 15 phút, CoA, natri hexanoat, ATP, và đệm HEPES, trong đó không có hexanoyl-CoA được sản xuất. Fig 2F mô tả thử nghiệm trong đó CsHCS2 protein đã được làm bất hoạt trước bằng cách đun sôi ở nhiệt độ 95°C trong 15 phút, CoA, natri hexanoat, ATP, và đệm HEPES, trong đó không có hexanoyl-CoA được sản xuất.

Fig.3 mô tả hai đồ thị minh họa cơ chất axit carboxylic được sử dụng bởi các enzym theo sáng chế. Fig.3A mô tả cơ chất axit carboxylic được sử dụng bởi CsHCS1. Fig.3B mô tả cơ chất axit carboxylic được sử dụng bởi CsHCS2.

Fig.4 mô tả phân tích sắc ký lỏng tính năng cao của các sản phẩm được sản xuất bởi thử nghiệm enzym liên hợp gồm *Cannabis sativa* hexanoyl-CoA synthetaza CsHCS2, malonyl-CoA synthetaza (MCS), *Cannabis sativa* “olivetol syntaza”/polyketit syntaza, và *Cannabis sativa* olivetolic axit syntaza. Hợp chất đã được rửa giải được phát hiện bởi độ hấp thụ ở 263 nm và được nhận diện cả do có cùng thời gian lưu như các chuẩn đã được phân lập, và do khối lượng của chúng sử dụng bộ dò khối lượng bốn cực đơn. Việc phát hiện olivetol và axit olivetolic cho thấy rằng CsHCS2 có khả năng tạo ra đủ cơ chất hexanoyl-CoA cho sự tổng hợp axit olivetolic. Các thử nghiệm không có CsHCS2, CoA, hoặc hexanoat thì không sản xuất bất kỳ sản phẩm polyketit

nào. HTAL = hexanoyltriactetic lacton, PDAL = pentyldiaxetic lacton, OA = axit olivetolic, OL = olivetol.

Fig.5 mô tả đồ thị thể hiện sự sản xuất axit olivetolic ở tế bào nấm men được biến đổi di truyền để sản xuất axit olivetolic bằng cách sử dụng CsHCS1 và CsHCS2 để tổng hợp hexanoyl-CoA, và sự dung hợp “olivetol syntaza”/polyketit syntaza (PKS) và olivetolic axit syntaza (OAS) của cây gai dầu để tạo thành axit olivetolic.

Fig.6 mô tả phân tích qRT-PCR sự biểu hiện của *CsHCS1*, *CsHCS2* và *CBDA* Syntaza ở các mô khác nhau của cây gai dầu ‘Finola’. Các giá trị biểu hiện gen so với *actin* được vẽ đồ thị dưới dạng số lần chênh lệch so với lá, trong đó sự biểu hiện ở lá được gán giá trị bằng 1. Các bản đồ lồng mô tả sự biểu hiện gen ở hoa cái có và không có các túm lông, trong đó các giá trị cũng được biểu thị dưới dạng số lần chênh lệch so với lá. R, rễ; S, thân; L, lá; FF+, hoa cái có các túm lông; FF-, hoa cái được loại bỏ các túm lông bởi phương pháp Beadbeater; T, các túm lông; MF, hoa đực. Các giá trị là giá trị trung bình $\pm$ SD, n=3.

Mô tả chi tiết sáng chế

Thư viện cADN đặc hiệu túm lông cây gai dầu deom được giải trình tự để tạo ra 9157 thẻ gắn trình tự rõ (“EST”) mà lắp ráp thành 4113 trình tự đơn nhất (1227 contig, 2886 singleton). Các unigen được chú thích bởi sự so sánh với cơ sở dữ liệu protein UniProt™ sử dụng tra cứu trực tuyến và công cụ so sánh được gọi là blastx. Các protein enzym hoạt hóa axyl ở cây gai dầu được nhận diện bằng cách sử dụng trình tự enzym hoạt hóa axyl của *Arabidopsis* để truy vấn các EST ở cây gai dầu đã được lắp ráp sử dụng tra cứu trực tuyến và công cụ so sánh được gọi là tblastn. 11 enzym hoạt hóa axyl được nhận diện và được đặt tên theo độ dồi dào phiên mã của chúng trong thư viện cADN. CsHCS1 là enzym hoạt hóa axyl dồi dào nhất dựa trên mức phiên mã (42 EST); CsHCS2 có độ dồi dào thấp hơn (5 EST). Dựa trên mức phiên mã cao của nó trong các túm lông và sự định vị của CsHCS1 ở bào lập, dường như enzym này là enzym hoạt hóa axyl có liên quan đến việc cung cấp hexanoyl-CoA cho con

đường canabinoit. CsHCS2, được định vị ở peroxisom, có lẽ không có liên quan đến sự tạo thành canabinoit. Tuy nhiên, các tính chất động lực học của nó làm cho enzym này là enzym hữu dụng để tổng hợp hexanoyl-CoA trong các vật chủ khác loài hoặc trong hệ thống không có tế bào.

Trình tự của gen CsHCS1 là như sau:

Cannabis sativa CsHCS1 – 2163 bp (SEQ ID NO: 1)

ATGGGTAAGAATTACAAGTCCCTGGACTCTGTTGTGGCCTCTGACTTCATAGCCCT
AGGTATCACCTCTGAAGTTGCTGAGACACTCCATGGTAGACTGGCCGAGATCGTG
TGTAATTATGGCGCTGCCACTCCCCAAACATGGATCAATATTGCCAACCATATTCT
GTCGCCTGACCTCCCCTTCTCCCTGCACCAGATGCTCTTCTATGGTTGCTATAAAG
ACTTTGGACCTGCCCCCTCCTGCTTGGATACCCGACCCGGAGAAAGTAAAGTCCAC
CAATCTGGGCGCACTTTTGGAGAAGCGAGGAAAAGAGTTTTTGGGAGTCAAGTAT
AAGGATCCCATTTCAAGCTTTTCTCATTTCCAAGAATTTTCTGTAAGAAACCCTGA
GGTGTATTGGAGAACAGTACTAATGGATGAGATGAAGATAAGTTTTTCAAAGGAT
CCAGAATGTATATTGCGTAGAGATGATATTAATAATCCAGGGGGTAGTGAATGGC
TTCCAGGAGGTTATCTTAACCTCAGCAAAGAATTGCTTGAATGTAAATAGTAACAA
GAAATTGAATGATACAATGATTGTATGGCGTGATGAAGGAAATGATGATTTGCCT
CTAAACAAATTGACACTTGACCAATTGCGTAAACGTGTTTGGTTAGTTGGTTATG
CACTTGAAGAAATGGGTTTGGAGAAGGGTTGTGCAATTGCAATTGATATGCCAAT
GCATGTGGATGCTGTGGTTATCTATCTAGCTATTGTTCTTGCGGGATATGTAGTTG
TTTCTATTGCTGATAGTTTTTCTGCTCCTGAAATATCAACAAGACTTCGACTATCA
AAAGCAAAAGCCATTTTTACACAGGATCATATTATTCGTGGGAAGAAGCGTATTC
CCTTATACAGTAGAGTTGTGGAAGCCAAGTCTCCCATGGCCATTGTTATTCCTTGT
AGTGGCTCTAATATTGGTGCAGAATTGCGTGATGGCGATATTTCTTGGGATTACTT
TCTAGAAAGAGCAAAAGAGTTTAAAAATTGTGAATTTACTGCTAGAGAACAACC
AGTTGATGCCTATACAAACATCCTCTTCTCATCTGGAACAACAGGGGAGCCAAAG
GCAATTCCATGGACTCAAGCAACTCCTTTAAAAGCAGCTGCAGATGGGTGGAGCC
ATTTGGACATTAGGAAAGGTGATGTCATTGTTTGGCCCACTAATCCTTGGTTGGAT
GATGGGTCCTTGGCTGGTCTATGCTTCACTCCTTAATGGGGCTTCTATTGCCTTGT
ATAATGGATCACCCTTGTCTGGCTTTGCCAAATTTGTGCAGGATGCTAAAGTA
ACAATGCTAGGTGTGGTCCCTAGTATTGTTTCGATCATGGAAAAGTACCAATTGTG
TTAGTGGCTATGATTGGTCCACCATCCGTTGCTTTTCTCTTCTGGTGAAGCATCT
AATGTAGATGAATACCTATGGTTGATGGGGAGAGCAAACTACAAGCCTGTTATCG
AAATGTGTGGTGGCACAGAAATTGGTGGTGCATTTTCTGCTGGCTCTTTCTTACAA
GCTCAATCATTATCTTCATTTAGTTCACAATGTATGGGTTGCACTTTATACATACT

TGACAAGAATGGTTATCCAATGCCTAAAAACAAACCAGGAATTGGTGAATTAGC
 GCTTGGTCCAGTCATGTTTGGAGCATCGAAGACTCTGTTGAATGGTAATCACCAT
 GATGTTTATTTTAAGGGAATGCCTACATTGAATGGAGAGGTTTTAAGGAGGCATG
 GGGACATTTTTGAGCTTACATCTAATGGTTATTATCATGCACATGGTCGTGCAGAT
 GATACAATGAATATTGGAGGCATCAAGATTAGTTCCATAGAGATTGAACGAGTTT
 GTAATGAAGTTGATGACAGAGTTTTTCGAGACAACTGCTATTGGAGTGCCACCTTT
 GGGCGGTGGACCTGAGCAATTAGTAATTTTCTTTGTATTAAAAGATTCAAATGAT
 ACAACTATTGACTTAAATCAATTGAGGTTATCTTTCAACTTGGGTTTACAGAAGA
 AACTAAATCCTCTGTTCAAGGTCACCTCGTGTGTGCCTCTTTCATCACTTCCGAGA
 ACAGCAACCAACAAGATCATGAGAAGGGTTTTGCGCCAGCAATTTTCTCACTTTG
 AATGA

Trình tự của gen CsHCS2 là như sau:

Cannabis sativa CsHCS2 – 1547 bp (SEQ ID NO: 3)

ATGGAGAAATCTTTTTTCAGAACTCATCTTCATACCCACAAAAGCCAGCTCTCAT
 TGATTCCGAAACCAACCAAAATACTCTCCTTTTCCCACTTCAAATCTACGGTTATCA
 AGGTCTCCCATGGCTTTCTCAATCTGGGTATCAAGAAAAACGACGTCGTTCTCAT
 CTACGCCCTAATTCTATCCACTTCCCTGTTTGTTCCTGGGAATTATAGCCTCTG
 GAGCCATTGCCACTACCTCAAATCCTCTCTACACAGTTTCCGAGCTTCCAAACAG
 GTCAAGGATTCCAATCCCAAACCTCATTATCACCGTTCCTCAACTCTTGAAAAAG
 TAAAGGGTTTCAATCTCCCCACGATTCTAATTGGTCCTGATTCTGAACAAGAATCT
 TCTAGTGATAAAGTAATGACCTTTAACGATTTGGTCAACTTAGGTGGGTCGTCG
 GCTCAGAATTTCCAATTGTTGATGATTTTAAGCAGAGTGACACTGCTGCGCTATTG
 TACTCATCTGGCACAACGGGAATGAGTAAAGGTGTGGTTTTGACTCACAAAACT
 TCATTGCCTCTTCTTTAATGGTGACAATGGAGCAAGACCTAGTTGGAGAGATGGA
 TAATGTGTTTCTATGCTTTTTGCCAATGTTTCATGTATTGTTTGGCTATCATCAC
 CTATGCTCAGTTGCAGAGAGGAAACACTGTTATTTCAATGGCGAGATTTGACCTT
 GAGAAGATGTAAAAGATGTGAAAAAGTATAAAGTTACCCATTTGTGGGTTGTGC
 CTCCTGTGATACTGGCTCTGAGTAAGAACAGTATGGTGAAGAAGTTAATCTTTC
 TTCTATAAAGTATATTGGCTCCGGTGCAGCTCCTTTGGGCAAAGATTTAATGGAG
 GAGTGCTCTAAGGTTGTTCTTATGGTATTGTTGCTCAGGGATATGGTATGACAG
 AAACCTTGTTGGGATTGTATCCATGGAGGATATAAGAGGAGGTAAACGAAATAGTG
 GTTCAGCTGGAATGCTGGCATCTGGAGTAGAAGCCCAGATAGTTAGTGTAGATAC
 ACTGAAGCCCTTACCTCCTAATCAATTGGGGGAGATATGGGTGAAGGGGCCTAAT
 ATGATGCAAGGTTACTTCAATAACCCACAGGCAACCAAGTTGACTATAGATAAGA
 AAGGTTGGGTACATACTGGTGATCTTGGATATTTTGATGAAGATGGACATCTTTA

TGTTGTTGACCGTATAAAAAGAGCTCATCAAATATAAAGGATTTTCAGGTTGCTCCT
 GCTGAGCTTGAAGGATTGCTTGTCTTCTCACCTGAAATACTCGATGCTGTTGTGAT
 TCCATTTCTGACGCTGAAGCGGGTGAAGTCCCAGTTGCTTATGTTGTGCGCTCTC
 CCAACAGTTCATTAACCGAAAATGATGTGAAGAAATTTATCGCGGGCCAGGTTGC
 ATCTTTCAAAAAGATTGAGAAAAGTAACATTTATAAACAGTGTCCCGAAATCTGCT
 TCGGGGAAAATCCTCAGAAGAGAACTCATTGAGAAAGTACGCTCCAACATGTGA

CsHCS1 và CsHCS2 được khuếch đại bằng PCR như được nêu trong Ví dụ 1 và hoạt tính alkanoyl-CoA synthetaza hoặc CoA-ligaza được xác định như nêu trong Ví dụ 2. Như được thể hiện trên Fig.2, CsHCS1 và CsHCS2 xúc tác sự sản xuất alkanoyl-CoA từ axit carboxylic và CoA.

Theo một số phương án, sáng chế đề cập đến phân tử axit nucleic được phân lập hoặc được tinh chế có SEQ ID NO: 1 hoặc có ít nhất 75%, ít nhất 76%, ít nhất 77%, ít nhất 78%, ít nhất 79%, ít nhất 80%, ít nhất 81%, ít nhất 82%, ít nhất 83%, ít nhất 84%, ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với SEQ ID NO: 1.

Theo một số phương án, sáng chế đề cập đến phân tử axit nucleic được phân lập hoặc được tinh chế có SEQ ID NO: 3 hoặc có ít nhất 75%, ít nhất 76%, ít nhất 77%, ít nhất 78%, ít nhất 79%, ít nhất 80%, ít nhất 81%, ít nhất 82%, ít nhất 83%, ít nhất 84%, ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với SEQ ID NO: 3.

Cũng được bao gồm là phân tử axit nucleic lai với trình tự axit nucleic đã được bộc lộ ở trên. Điều kiện lai hóa có thể là nghiêm ngặt trong đó sự lai hóa sẽ xảy ra nếu có ít nhất 90%, 95% hoặc 97% độ đồng nhất trình tự với phân tử axit nucleic mã hóa enzym theo sáng chế. Điều kiện nghiêm ngặt có thể bao gồm các điều kiện được sử dụng cho phép lai hóa Southern đã biết như, ví dụ, ủ qua đêm ở nhiệt độ 42°C trong dung dịch có formamit 50%, 5x SSC (NaCl 150 mM, trinati xitrat 15 mM), natri phosphat 50 mM (pH 7,6), 5x dung dịch

Denhardt, dextran sulfat 10%, và 20 microgam/mililit ADN tinh trùng cá hồi đã được cắt, được làm biến tính, sau đó là rửa giá lai hóa trong 0,1x SSC ở nhiệt độ khoảng 65°C. Các điều kiện lai hóa đã biết khác là đã được biết rõ và được mô tả trong Sambrook et al., *Molecular Cloning: Laboratory Manual*, Third Edition, Cold Spring Harbor, N.Y. (2001).

Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng các thay đổi không đáng kể trong trình tự axit nucleic không nhất thiết làm biến đổi trình tự axit amin của polypeptit đã được mã hóa. Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng các thay đổi trong độ đồng nhất của các nucleotit trong trình tự gen đặc hiệu mà thay đổi trình tự axit amin của polypeptit đã được mã hóa có thể dẫn đến tính hiệu quả giảm hoặc tăng của các gen và, trong một số ứng dụng (ví dụ, đối nghĩa, ức chế đồng thời, hoặc ARNi), các trình tự một phần thường hoạt động hiệu quả như các kiểu chiều dài đầy đủ. Các cách thức trong đó trình tự nucleotit có thể được làm thay đổi hoặc được làm ngắn là đã biết rõ với người có hiểu biết trung bình về lĩnh vực kỹ thuật này, như là các cách thử nghiệm tính hiệu quả của các gen đã được làm biến đổi. Theo các phương án nhất định, tính hiệu quả có thể được thử nghiệm một cách dễ dàng bởi, ví dụ, sắc ký khí thông thường. Do đó, tất cả các biến đổi như vậy của các gen đều được bao gồm là một phần của phần bộc lộ này.

Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng chiều dài của phân tử axit nucleic nêu trên đây sẽ tùy thuộc vào ứng dụng đã định. Ví dụ, nếu ứng dụng đã định là để làm đoạn môi hoặc mẫu dò, ví dụ, để khuếch đại bằng PCR hoặc để sàng lọc thư viện, chiều dài của phân tử axit nucleic sẽ nhỏ hơn trình tự chiều dài đầy đủ, ví dụ, 15-50 nucleotit. Theo các phương án này, đoạn môi hoặc mẫu dò có thể về cơ bản giống với vùng được bảo toàn cao của trình tự axit nucleic hoặc có thể về cơ bản giống với đầu 5' hoặc 3' của trình tự ADN. Trong một số trường hợp, các đoạn môi hoặc mẫu dò này có thể sử dụng các bazơ chung ở một số vị trí để 'về cơ bản giống' nhưng vẫn tạo ra tính linh hoạt trong việc nhận biết trình tự. Cần lưu ý rằng điều kiện lai hóa đoạn môi và mẫu dò thích hợp là đã biết rõ trong lĩnh vực kỹ thuật này.

Sáng chế còn bao gồm enzym CsHCS1. Trình tự axit amin của CsHCS1 (SEQ ID NO: 2) là:

MGKNYKSLDSVVASDFIALGITSEVAETLHGRLAEIVCNYGAATPQTWINIANHILSP
DLFSLHQMLFYGCYKDFGPAPPAWIPDPEKVKSTNLGALLEKRGKEFLGVKYKDPI
SSFSHFQEFSSVRNPEVYWRTVLMDEMKISFSKDPECILRRDDINNPGGSEWLPGGYLN
SAKNCLNVNSNKKLNDTMIVWRDEGNDDLPLNKLTL DQLRKRVLVGYALEEMGL
EKGCAIAIDMPMHVDAVVIYLAIVLAGYVVVSIADSFSAPEISTRRLSKAKAIFTQDH
IIRGKKRIPLYSRVVEAKSPMAIVIPCSGSNIGAELRDGDISWDYFLERAKEFKNCEFTA
REQPVDAITNLFSSGTTGEPKAIPWTQATPLKAAADGWSHLDIRKGDVIVWPTNLG
WMMGPWLVIYASLLNGASIALYNGSPLVSGFAKFVQDAKVTMLGVVPSIVRSWKST
NCVSGYDWSTIRCFSSSGEASNVEYLWLMGRANYKPVIEMCGGTEIGGAFSAGSFL
QAQSLSSFSSQCMGCTLYILDKNGYPMPKNKPGIGELALGPVMFGASKTLLNGNHHD
VYFKGMPTLNGEVLRRHGDIFELTSNGYYHAHGRADDTMNIGGIKISSIEIERVCNEV
DDR VFETTAIGVPPLGGGPEQLVIFVFLKDSNDTTIDLNQLRSLFNGLQKKLNPLFKV
TRVVPLSSLPRATATNKIMRRVLRQFSHFE

Một số phương án đề cập đến polypeptit được phân lập hoặc được tinh chế có SEQ ID NO. 2 hoặc có ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với trình tự axit amin như được nêu trong SEQ ID NO: 2.

Sáng chế còn bao gồm enzym CsHCS2. Trình tự axit amin của CsHCS2 (SEQ ID NO: 4) là:

MEKSGYGRDGIYRSLRPPLHLPNNNNLSMVSFLFRNSSSYQKPALIDSETNQILSFSH
FKSTVIKVSHGFLNLGIKKNDVVLIYAPNSIHFPVCFLGIASGAIATTSNPLYTVSELK
QVKDSNPCLIITVPQLLEKVKGFNLPTILIGPDSEQESSSDKVMTFNDLVNLGGSSGSE
FPIVDDFKQSDTAALLYSSGTTGMSKGVVLTHKNFIASSLMVTMEQDLVGEMDNVFL
CFLPMFHVFGLAITYAQLQRGNTVISMARFDLEKMLKDVEKYKVTHLWVPPVILA
LSKNSMVKKFNLSSIKYIGSGAAPLGKDLMEESKVVYPYGIVAQGYGMTETCGIVSM
EDIRGGKRNSGSAGMLASGVEAQIVSVDTLKPLPPNQLGEIWVKGPNMMQGYFNNP
QATKL TIDKKGWVHTGDLGYFDEDGHLYVVDRIKELIKYKGFQVAPAELEGLLVSH
EILDAVVIPFPDAEAGEVPVAYVVRSPNSSLTENDVKKFIAGQVASFKRLRKVTFINSV
PKSASGKILRRELIQKVRNM

Một số phương án đề cập đến polypeptit được phân lập hoặc được tinh chế có SEQ ID NO. 4 hoặc có ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với trình tự axit amin như được nêu trong SEQ ID NO: 4.

Một số phương án đề cập đến vector, cấu trúc hoặc hệ biểu hiện chứa polynucleotit được phân lập hoặc được tinh chế có trình tự của SEQ ID NO: 1 hoặc SEQ ID NO: 3, hoặc ít nhất 75%, ít nhất 76%, ít nhất 77%, ít nhất 78%, ít nhất 79%, ít nhất 80%, ít nhất 81%, ít nhất 82%, ít nhất 83%, ít nhất 84%, ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với SEQ ID NO: 1 hoặc SEQ ID NO: 3. Ngoài ra, sáng chế cũng đề xuất phương pháp điều chế vector, cấu trúc hoặc hệ biểu hiện bao gồm trình tự này, hoặc một phần của nó, để đưa trình tự hoặc một phần trình tự này theo hướng có nghĩa hoặc đối nghĩa, hoặc phần bổ sung của nó, vào tế bào.

Theo một số phương án, phân tử axit nucleic được phân lập và/hoặc được tinh chế, hoặc vector, cấu trúc hoặc hệ biểu hiện chứa phân tử axit nucleic được phân lập và/hoặc được tinh chế này, có thể được sử dụng để tạo ra sinh vật hoặc tế bào của sinh vật chuyển gen sản xuất các polypeptit xúc tác sự tổng hợp của polyketit thơm. Do vậy, theo một phương án, sáng chế đề cập đến sinh vật chuyển gen, tế bào hoặc mô phôi của sinh vật này chứa phân tử axit nucleic được phân lập và/hoặc được tinh chế có SEQ ID NO: 1 hoặc SEQ ID NO: 3

có ít nhất 75%, ít nhất 76%, ít nhất 77%, ít nhất 78%, ít nhất 79%, ít nhất 80%, ít nhất 81%, ít nhất 82%, ít nhất 83%, ít nhất 84%, ít nhất 85%, ít nhất 86%, ít nhất 87%, ít nhất 88%, ít nhất 89%, ít nhất 90%, ít nhất 91%, ít nhất 92%, ít nhất 93%, ít nhất 94%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98% hoặc ít nhất 99% độ đồng nhất với SEQ ID NO: 1 hoặc SEQ ID NO: 3.

Tốt hơn là, sinh vật là cây, vi sinh vật hoặc côn trùng. Cây tốt hơn nếu thuộc giống *Cannabis*, ví dụ *Cannabis sativa* L., *Cannabis indica* Lam. và

Cannabis ruderalis Janisch. Đặc biệt được ưu tiên là *Cannabis sativa*. Vi sinh vật tốt hơn nếu là vi khuẩn (ví dụ, *Escherichia coli*) hoặc nấm men (ví dụ *Saccharomyces cerevisiae*). Côn trùng tốt hơn nếu là *Spodoptera frugiperda*.

Sinh vật, tế bào và mô phôi theo phương án này có thể có hàm lượng hợp chất canabinoit thay đổi. Liên quan đến Fig.1, người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng sự biểu hiện hoặc biểu hiện quá mức của phân tử axit nucleic theo sáng chế sẽ dẫn đến sự biểu hiện hoặc biểu hiện quá mức của enzym xúc tác sự tổng hợp của hexanoyl-CoA, enzym này khi kết hợp với các enzym khác, có thể dẫn đến sự sản xuất hoặc sản xuất tăng hợp chất canabinoit như axit canabigerolic (CBGA), axit Δ^9 -tetrahydrocannabinolic (THCA), axit canabidiolic (CBDA), axit canabichromenic (CBCA), Δ^9 -tetrahydrocannabinol (THC), canabidiol (CBD), canabichromen (CBC), v.v.. Tương tự, phụ thuộc vào cơ chất được sử dụng, sự biểu hiện hoặc biểu hiện quá mức của phân tử axit nucleic theo sáng chế dẫn đến sự biểu hiện hoặc biểu hiện quá của enzym xúc tác sự tổng hợp của hexanoyl-CoA, có thể dẫn đến sự sản xuất hoặc sản xuất tăng chất tương tự của hợp chất canabinoit, hoặc chất tương tự của tiền chất của các hợp chất này.

Việc làm bất hoạt gen ở sinh vật, tế bào hoặc mô sẽ dẫn đến sự biểu hiện dưới mức của enzym mà có thể dẫn đến sự tích lũy của các tiền chất như axit hexanoic (sáu cacbon), axit octanoic (tám cacbon), axit nonanoic (chín cacbon), axit valeric (năm cacbon), axit heptanoic (bảy cacbon) hoặc các axit carboxylic khác, và/hoặc sự giảm canabinoit như THCA (tiền chất của THC) hoặc CBDA (tiền chất của CBD).

Sáng chế bao gồm quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô bằng cách biểu hiện hoặc biểu hiện quá mức enzym ngoại sinh theo sáng chế ở sinh vật, tế bào hoặc mô này, so với giống sinh vật, tế bào hoặc mô tương tự sinh trưởng trong các điều kiện tương tự nhưng không có enzym ngoại sinh theo sáng chế được biểu hiện hoặc được biểu hiện quá mức.

Sự biểu hiện hoặc biểu hiện quá mức của phân tử axit nucleic theo sáng chế có thể được thực hiện kết hợp với sự biểu hiện hoặc biểu hiện quá của một hoặc nhiều axit nucleic khác mã hoá một hoặc nhiều enzym trong con đường sinh tổng hợp cannabinoid. Một số ví dụ về các axit nucleic khác bao gồm các axit nucleic mã hóa: polyketit syntaza loại III, polyketit xyclaza, prenyltransferaza thơm và oxidoxylaza tạo thành cannabinoid. Ví dụ cụ thể về các enzym này bao gồm “olivetol syntaza”/polyketit syntaza, olivetolic axit syntaza, geranylpyrophosphat:olivetolat geranyltransferaza, Δ^9 -tetrahydrocannabinolic axit syntaza, canabidiolic axit syntaza hoặc canabichromenic axit syntaza. Sự tổng hợp alkanoyl-CoA trong điều kiện có mặt enzym polypeptit theo sáng chế có thể được thực hiện *in vivo* hoặc *in vitro*. Như được đề cập trước đây, các tổng hợp như vậy *in vivo* có thể được thực hiện bằng cách biểu hiện hoặc biểu hiện quá mức phân tử axit nucleic theo sáng chế ở sinh vật, tế bào hoặc mô.

Sự tổng hợp alkanoyl-CoA *in vitro* có thể xảy ra trong hệ không có tế bào. Để làm một phần của hệ không có tế bào *in vitro*, axit carboxylic và enzym theo sáng chế có thể được trộn cùng nhau trong bình phản ứng thích hợp để thực hiện phản ứng.

In vitro, các polypeptit theo sáng chế có thể được sử dụng kết hợp với các enzym khác để thực hiện sự tổng hợp hoàn toàn hợp chất cannabinoid từ tiền chất. Ví dụ, các enzym khác như vậy có thể có liên quan trong con đường sinh tổng hợp cannabinoid như được nêu trên Fig.1 (như “olivetol syntaza”/PKS, olivetolic axit syntaza, prenyltransferaza thơm, THCA syntaza, CBDA syntaza, CBCA syntaza).

Các polypeptit theo sáng chế có thể được sử dụng, *in vivo* hoặc *in vitro*, để tổng hợp chất tương tự của hợp chất cannabinoid không có trong tự nhiên ở loài chủ. Các chất tương tự như vậy có thể được sản xuất bằng cách sử dụng hợp chất axit carboxylic khác với các hợp chất được sử dụng để sản xuất hợp chất cannabinoid tự nhiên ở cây. Ví dụ, axit axetic, axit butyric, axit octanoic,

axit decanoic, axit lauric, axit myristic, axit palmitic; axit mạch nhánh như axit isovaleric; và axit hydroxyxinnamic như axit xinnamic.

Các thuật ngữ:

Để thuận tiện cho việc xem xét các phương án của phần bộc lộ này, các thuật ngữ cụ thể được giải thích sau đây:

Alkanoyl-CoA: alkanoyl-CoA là hợp chất cacbonyl béo có gốc coenzym A được liên kết với nguyên tử cacbon của nhóm cacbonyl thông qua cầu sulfua. Hợp chất alkanoyl-CoA được ưu tiên chứa từ 2 đến 10 nguyên tử cacbon trong phần cacbonyl béo của hợp chất này. Tốt hơn nữa nếu, alkanoyl-CoA là $\text{CoA-S-C(O)-(CH}_2\text{)}_n\text{-CH}_3$, trong đó n là số nguyên nằm trong khoảng từ 0 đến 8. Một số ví dụ về hợp chất alkanoyl-CoA là axetyl-CoA, butyryl-CoA, hexanoyl-CoA và octanoyl-CoA. Việc sử dụng axetyl-CoA tạo ra mạch bên methyl cho polyketit thơm thu được; việc sử dụng butyryl-CoA tạo ra mạch bên propyl; và việc sử dụng hexanoyl-CoA tạo ra mạch bên pentyl. Hexanoyl-CoA được đặc biệt ưu tiên. Canabinoit với mạch bên ngắn hơn tồn tại trong cây gai dầu (ví dụ, axit tetrahydrocannabinolic có mạch bên propyl thay vì mạch bên pentyl của THCA).

Sự thoái hóa codon: Cần phải hiểu rằng phần bộc lộ này bao gồm sự thoái hóa sử dụng codon như sẽ được hiểu bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật này và như được minh họa trong Bảng 1.

Bảng 1 Sự thoái hóa codon

Axit amin	Codon
Ala/A	GCT, GCC, GCA, GCG
Arg/R	CGT, CGC, CGA, CGG, AGA, AGG
Asn/N	AAT, AAC
Asp/D	GAT, GAC
Cys/C	TGT, UGC
Gln/Q	CAA, CAG
Glu/E	GAA, GAG

Gly/G	GGT, GGC, GGA, GGG
His/H	CAT, CAC
Ile/I	ATT, ATC, ATA
Leu/L	TTA, TTG, CTT, CTC, CTA, CTG
Lys/K	AAA, AAG
Met/M	ATG
Phe/F	TTT, TTC
Pro/P	CCT, CCC, CCA, CCG
Ser/S	TCT, TCC, TCA, TCG, AGT, AGC
Thr/T	ACT, ACC, ACA, ACG
Trp/W	TGG
Tyr/Y	TAT, TAC
Val/V	GTT, GTC, GTA, GTG
BẮT ĐẦU	ATG
KẾT THÚC	TAG, TGA, TAA

Trình tự nucleotit bổ sung: “Trình tự nucleotit bổ sung” của một trình tự được hiểu nghĩa là phân tử axit nucleic bất kỳ mà các nucleotit của nó là bổ sung với nucleotit của trình tự được nêu trong bản mô tả này, và hướng của nó được đảo ngược (trình tự đối song song).

Sự thay thế bảo toàn: Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng sự thay thế bảo toàn có thể được thực hiện trong trình tự axit amin của polypeptit mà không phá vỡ cấu trúc ba chiều hoặc chức năng của polypeptit này. Như vậy, sáng chế bao gồm các polypeptit chứa CsHCS1 và CsHCS2 được thế bảo toàn. Sự thay thế bảo toàn được thực hiện bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật này bằng cách thay thế axit amin có tính kỵ nước, độ phân cực, và chiều dài mạch R tương tự cho axit amin khác. Ngoài ra, bằng cách so sánh trình tự được sắp hàng của các protein tương đồng từ các loài khác nhau, sự thay thế bảo toàn có thể được nhận diện bằng cách đặt các gốc axit amin đã được làm đột biến giữa các loài mà không làm biến đổi

chức năng cơ bản của các protein đã được mã hóa. Bảng 2 đưa ra danh sách nêu làm ví dụ về sự thay thế bảo toàn.

Bảng 2 Sự thay thế bảo toàn

<u>Loại Axit amin</u>	<u>Axit amin có thể thay thế</u>
Ưa nước	Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser, Thr
Sulphydryl	Cys
Béo	Val, Ile, Leu, Met
Bazơ	Lys, Arg, His
Thơm	Phe, Tyr, Trp

Mức độ hoặc tỷ lệ phần trăm của độ tương đồng trình tự: Thuật ngữ "mức độ hoặc tỷ lệ phần trăm của độ tương đồng trình tự" chỉ mức độ hoặc tỷ lệ phần trăm của độ tương đồng trình tự giữa hai trình tự sau khi sắp hàng tối ưu.

Trình tự được phân lập và/hoặc được tinh chế tương đồng: “Trình tự được phân lập và/hoặc được tinh chế tương đồng” được hiểu là dùng để chỉ trình tự được phân lập và/hoặc được tinh chế có độ đồng nhất tính theo phần trăm với các bazơ của trình tự nucleotit, hoặc axit amin của trình tự polypeptit, ít nhất khoảng 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,6%, hoặc 99,7%. Tỷ lệ phần trăm này là hoàn toàn có tính chất thống kê, và có thể xếp loại độ khác biệt giữa hai nucleotit hoặc trình tự axit amin ngẫu nhiên và trên toàn bộ chiều dài của chúng. Độ đồng nhất trình tự có thể được xác định, ví dụ, bởi các chương trình máy tính được thiết kế để thực hiện một hoặc nhiều lần sắp hàng trình tự.

Việc làm tăng, giảm, điều biến, biến đổi hoặc các thuật ngữ tương tự: Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng các thuật ngữ như vậy chỉ sự so sánh với giống hoặc chủng tương tự sinh trưởng trong các điều kiện tương tự nhưng không có sự cải biến dẫn đến sự tăng, giảm, điều biến hoặc biến đổi. Trong một số trường hợp, giống hoặc chủng này có thể là

đối chứng chưa được biến nạp, đối chứng được biến nạp giả, hoặc đối chứng được biến nạp vector.

Được phân lập: Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng “được phân lập” chỉ các polypeptit hoặc axit nucleic mà đã “được phân lập” từ môi trường tự nhiên của chúng.

Trình tự nucleotit, polynucleotit, hoặc axit nucleic: “trình tự nucleotit, polynucleotit, hoặc axit nucleic” sẽ được hiểu là chỉ cả dạng sợi kép hoặc sợi đơn ở dạng monome và dime (đôi) và sản phẩm phiên mã của chúng.

Độ đồng nhất trình tự: Hai trình tự axit amin hoặc hai trình tự nucleotit được gọi là "đồng nhất" nếu trình tự của axit amin hoặc nucleotit trong hai trình tự này là giống nhau khi được sắp hàng tương ứng tối đa như được nêu dưới đây. Tỷ lệ phần trăm của độ đồng nhất trình tự (hoặc mức độ đồng nhất) được xác định bằng cách so sánh hai trình tự đã được sắp hàng tối ưu trên cửa sổ so sánh, trong đó một phần của trình tự peptit hoặc polynucleotit trong cửa sổ so sánh có thể bao gồm phần bổ sung hoặc phần khuyết (tức là, các khoảng trống) so với trình tự tham chiếu (không bao gồm phần bổ sung hoặc phần khuyết) để sắp hàng tối ưu hai trình tự. Tỷ lệ phần trăm được tính toán bằng cách xác định số lượng của các vị trí tại đó gốc axit amin hoặc bazơ axit nucleic đồng nhất xuất hiện trong cả hai trình tự để tạo ra số lượng các vị trí được làm phù hợp, chia số lượng các vị trí đã được làm phù hợp cho tổng số lượng các vị trí trong cửa sổ so sánh và nhân kết quả này với 100 để thu được tỷ lệ phần trăm của độ đồng nhất trình tự.

Sự sắp hàng tối ưu của các trình tự để so sánh có thể được tiến hành nhờ thuật toán độ tương đồng cục bộ của Smith and Waterman, *Ad. App. Math* 2: 482 (1981), thuật toán sắp hàng tương đồng của Needleman and Wunsch, *J. Mol. Biol.* 48: 443 (1970), phương pháp tìm kiếm độ tương tự của Pearson and Lipman, *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 2444 (1988), sự thực hiện bằng máy tính của các thuật toán này (GAP, BESTFIT, FASTA, và TFASTA trong Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis.), hoặc bởi sự kiểm tra bằng mắt thường.

Định nghĩa về độ đồng nhất trình tự được đưa ra ở trên là định nghĩa được người có hiểu biết trung bình về lĩnh vực kỹ thuật này sử dụng. Định nghĩa này tự nó không cần sự hỗ trợ của thuật toán bất kỳ, các thuật toán này chỉ giúp ích để đạt được sự sắp hàng tối ưu của các trình tự, đúng hơn là tính toán độ đồng nhất trình tự.

Từ định nghĩa đưa ra trên đây, có một giá trị đã được xác định rõ và chỉ một giá trị đối với độ đồng nhất trình tự giữa hai trình tự được so sánh, giá trị này tương ứng với giá trị thu được đối với sự sắp hàng tốt nhất hoặc tối ưu.

Lai nghiêm ngặt: Lai trong các điều kiện nghiêm ngặt với trình tự nucleotit được hiểu là chỉ việc lai trong điều kiện nhiệt độ và nồng độ ion được chọn theo cách sao cho chúng cho phép duy trì sự lai giữa hai mảnh của các phân tử axit nucleic bổ sung.

Các thể tương đồng của các gen mới nêu trong bản mô tả này thu được từ các sinh vật khác, ví dụ, cây, có thể thu được bằng cách sàng lọc các thư viện thích hợp bao gồm các thể tương đồng đó, trong đó việc sàng lọc được tiến hành với trình tự nucleotit của các gen cụ thể theo sáng chế, hoặc các phần hoặc mẫu dò của chúng, hoặc được nhận diện bởi tra cứu độ tương đồng trình tự sử dụng các chương trình tra cứu sắp hàng trình tự như BLAST hoặc FASTA.

Việc phân lập axit nucleic và tách dòng là đã được biết rõ. Tương tự, gen đã được phân lập có thể được cài xen vào vectơ và được biến nạp vào tế bào bởi các kỹ thuật thông thường đã biết rõ với người có hiểu biết trung bình về lĩnh vực kỹ thuật này. Phân tử axit nucleic có thể được biến nạp vào sinh vật. Như đã biết trong lĩnh vực kỹ thuật này, có nhiều cách mà nhờ đó gen, vectơ, cấu trúc và hệ biểu hiện có thể được đưa vào sinh vật, và sự kết hợp thành công của các kỹ thuật nuôi cấy mô và biến nạp đem lại các phương pháp hữu hiệu để tạo ra các sinh vật chuyển gen. Các phương pháp này có thể được sử dụng trong sáng chế và đã được mô tả trong một số tài liệu khác (Potrykus I (1991) Gene transfer to plants: Assessment of published approaches and results. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42: 205-225; Vasil I K (1994) Molecular

improvement of cereals. Plant Mol. Biol. 25: 925-937. Walden R, Wingender R (1995) Gene-transfer and plant regeneration techniques. Trends in Biotechnology 13: 324-331; Songstad DD, Somers DA, Griesbach RJ (1995) Advances in alternative DNA delivery techniques. Plant Cell Tissue Organ Cult. 40:1-15), và đã biết rõ với người có hiểu biết trung bình về lĩnh vực kỹ thuật này.

Các vector thích hợp là đã biết rõ với người có hiểu biết trung bình về lĩnh vực và được mô tả trong các tài liệu tham khảo về kỹ thuật chung như Pouwels et al., Cloning Vectors. A Laboratory Manual, Elsevier, Amsterdam (1986). Các vector đặc biệt thích hợp bao gồm vector Ti plasmit. Ví dụ, người có hiểu biết trung bình về lĩnh vực kỹ thuật này chắc chắn sẽ biết rằng, ngoài sự biến nạp qua trung gian *Agrobacterium* của *Arabidopsis* bởi sự ngấm lọc chân không (Bechtold N, Ellis J, Pelletier G (1993) In planta *Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C R Acad Sci Paris, Sciences de la vie/Life sciences 316: 1194-1199.) hoặc sự cấy nhiễm vết thương (Katavic V, Haughn GW, Reed D, Martin M, Kunst L (1994) In planta transformation of *Arabidopsis thaliana*. Mol. Gen. Genet. 245: 363-370.), cũng có thể biến nạp các loài cây khác, sử dụng sự biến nạp qua trung gian Ti-plasmit *Agrobacterium* (ví dụ, trụ dưới lá mầm (DeBlock M, DeBrouwer D, Tenning P (1989) Transformation of *Brassica napus* and *Brassica oleracea* using *Agrobacterium tumefaciens* and the expression of the bar and neo genes in transgenic plants. Plant Physiol. 91: 694-701) hoặc cuống lá mầm (Moloney MM, Walker JM, Sharma KK (1989) High efficiency transformation of *Brassica napus* using *Agrobacterium* vectors. Plant Cell Rep. 8: 238-242.) phương pháp lây nhiễm vết thương, kỹ thuật bắn phá bằng hạt/bắn gen (Sanford JC, Klein TM, Wolf ED, Allen N (1987) Delivery of substances into cells and tissues using a particle bombardment process. J. Part. Sci. Technol. 5: 27-37.) hoặc phương pháp biến nạp thể nguyên sinh được hỗ trợ bằng polyetylen glycol (Rhodes CA, Pierce DA, Mettler IJ, Mascarenhas D,

Detmer JJ (1988) Genetically transformed maize plants from protoplasts. *Science* 240: 204-207).

Như cũng sẽ rõ ràng với người có hiểu biết trung bình về lĩnh vực kỹ thuật này, và như được nêu trong các tài liệu khác (Meyer P (1995) Understanding and controlling transgene expression. *Trends in Biotechnology* 13: 332-337; Datla R, Anderson JW, Selvaraj G (1997) Plant promoters for transgene expression. *Biotechnology Annual Review* 3: 269-296.), có thể sử dụng gen khởi đầu được liên kết theo cách có điều khiển với phân tử axit nucleic để điều khiển sự điều hòa tăng hoặc giảm bất kỳ như đã được dự định sử dụng các gen khởi đầu không được điều hòa (tức là, gen khởi đầu cơ định) (ví dụ, các gen khởi đầu dựa trên CaMV35S), hoặc bằng cách sử dụng các gen khởi đầu mà có thể nhắm đích sự biểu hiện gen ở tế bào cụ thể, mô (ví dụ, gen khởi đầu napin để biểu hiện gen chuyển để phát triển các lá mầm của hạt), cơ quan (ví dụ, rễ), ở giai đoạn phát triển cụ thể, hoặc để đáp ứng với kích thích cụ thể từ bên ngoài (ví dụ, sốc nhiệt).

Gen khởi đầu để sử dụng theo sáng chế có thể là gen khởi đầu cảm ứng, gen khởi đầu cơ định, hoặc gen khởi đầu đặc hiệu mô hoặc nhiều tổ hợp khác nhau của các đặc tính này. Gen khởi đầu hữu dụng bao gồm, nhưng không giới hạn ở gen khởi đầu cơ định như gen khởi đầu virus vòng ăn mòn ở cây cảm chướng (CERV), gen khởi đầu 35S virus khảm ở súp lơ (CaMV), hoặc cụ thể hơn, gen khởi đầu virus khảm ở súp lơ được tăng cường kép, bao gồm hai gen khởi đầu CaMV 35S nối tiếp (được gọi là gen khởi đầu "35S kép"). Có thể mong muốn sử dụng gen khởi đầu đặc hiệu mô hoặc được điều hòa theo cách phát triển thay cho gen khởi đầu cơ định trong một số trường hợp nhất định. Gen khởi đầu đặc hiệu mô cho phép sự biểu hiện quá mức ở một số mô nhất định mà không ảnh hưởng đến sự biểu hiện ở các mô khác.

Gen khởi đầu và vùng điều hòa kết thúc sẽ có chức năng trong tế bào chủ và có thể là khác loại (tức là, không có trong tự nhiên) hoặc tương đồng (thu được từ loài chủ) với tế bào và gen.

Vùng điều hòa kết thúc có thể thu được từ vùng 3' của gen từ đó gen khởi đầu thu được hoặc từ gen khác. Vùng kết thúc thích hợp có thể được sử dụng đã được biết rõ trong lĩnh vực kỹ thuật này và bao gồm gen kết thúc *Agrobacterium tumefaciens* nopalín syntaza (Tnos), gen kết thúc *A. tumefaciens* mannopín syntaza (Tmas) và gen kết thúc CaMV 35S (T35S). Vùng kết thúc được đặc biệt ưu tiên để sử dụng theo sáng chế bao gồm vùng kết thúc cấu trúc siêu phân tử nhỏ ribuloza biphosphat carboxylaza ở cây đậu Hà Lan (TrbcS) hoặc vùng kết thúc Tnos. Các cấu trúc gen để sử dụng theo sáng chế có thể được sàng lọc một cách thích hợp về hoạt tính bởi, ví dụ, sự biến nạp vào cây chủ thông qua *Agrobacterium* và sàng lọc mức cannabinoid đã được biến đổi.

Phân tử axit nucleic theo sáng chế, hoặc mảnh của chúng, có thể được sử dụng để phong bế quá trình sinh tổng hợp cannabinoid ở sinh vật mà sản xuất một cách tự nhiên hợp chất cannabinoid. Việc làm bất hoạt sử dụng phân tử axit nucleic theo sáng chế có thể được thực hiện theo nhiều cách nhìn chung đã biết trong lĩnh vực kỹ thuật này, ví dụ, kỹ thuật ARN can thiệp (ARNi), kỹ thuật microARN nhân tạo, kỹ thuật bất hoạt gen được cảm ứng bởi virus (VIGS), kỹ thuật đối nghĩa, kỹ thuật đồng ức chế có nghĩa và kỹ thuật phát sinh đột biến có đích.

Kỹ thuật ARNi bao gồm sự biến nạp ổn định bằng cách sử dụng cấu trúc plasmit ARN can thiệp (ARNi) (Helliwell CA, Waterhouse PM (2005) Constructs and methods for hairpin RNA-mediated gene silencing in plants. *Methods Enzymology* 392:24-35). Các plasmit như vậy được tạo thành từ mảnh gen đích được làm bất hoạt trong cấu trúc lặp đảo ngược. Các đoạn lặp đảo ngược được phân cách bởi đệm, thường là intron. Cấu trúc ARNi được điều khiển bởi gen khởi đầu thích hợp, ví dụ, gen khởi đầu 35S virus khảm ở súp lơ (CaMV), được kết hợp vào hệ gen của cây và tiếp theo là sự phiên mã của gen chuyển dẫn đến phân tử ARN gấp nếp ngược trên chính nó để tạo thành ARN kẹp tóc sợi kép. Cấu trúc ARN sợi kép này được nhận ra bởi cây và được cắt thành các ARN nhỏ (chiều dài khoảng 21 nucleotit) được gọi là ARN

can thiệp nhỏ (siARN). siARN kết hợp với phức hợp protein (RISC) mà tiếp tục điều khiển quá trình thoái hóa của mRNA đối với gen đích.

Kỹ thuật microARN nhân tạo (amiARN) lợi dụng con đường microARN (miARN) mà thực hiện chức năng để làm bất hoạt các gen nội sinh ở cây và các sinh vật có nhân điển hình khác (Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* 18:1121-33; Alvarez JP, Pekker I, Goldshmidt A, Blum E, Amsellem Z, Eshed Y (2006) Endogenous and synthetic microRNAs stimulate simultaneous, efficient, and localized regulation of multiple targets in diverse species. *Plant Cell* 18:1134-51). Trong phương pháp này, các mảnh có chiều dài 21 nucleotit của gen cần được làm bất hoạt được đưa vào gen pre-miARN để tạo thành cấu trúc pre-amiARN. Cấu trúc pre-amiARN được chuyển vào hệ gen sinh vật sử dụng các phương pháp biến nạp đã được các người có hiểu biết trung bình về lĩnh vực kỹ thuật này biết rõ. Sau khi phiên mã pre-amiARN, xử lý tạo ra amiARN mà nhắm đích các gen có chung độ đồng nhất nucleotit với trình tự amiARN 21 nucleotit.

Trong kỹ thuật bất hoạt ARNi, hai yếu tố có thể ảnh hưởng đến việc lựa chọn chiều dài của mảnh. Mảnh càng ngắn, sự bất hoạt hiệu quả đạt được thường ít hơn, nhưng các kẹp tóc rất dài làm tăng cơ hội tái tổ hợp ở các chủng vi khuẩn chủ. Tính hiệu quả của việc làm bất hoạt dường như cũng phụ thuộc vào gen và có thể phản ánh khả năng tiếp cận của mRNA đích hoặc độ dồi dào tương đối của mRNA đích và ARN hình kẹp tóc trong các tế bào trong đó gen hoạt động. Chiều dài mảnh nằm trong khoảng từ 100 đến 800 bp, tốt hơn nếu nằm trong khoảng từ 300 đến 600 bp, nhìn chung là thích hợp để tối đa hóa hiệu quả bất hoạt thu được. Mỗi quan tâm khác là phần gen được nhắm đích. 5' UTR, vùng ghi mã, và mảnh 3' UTR có thể được sử dụng với các kết quả tốt như nhau. Vì cơ chế bất hoạt tùy thuộc vào độ tương đồng trình tự, có khả năng có sự bất hoạt chéo các trình tự mRNA có liên quan. Khi điều này không được mong muốn, vùng với độ tương tự trình tự thấp với các trình tự khác, như 5' hoặc 3' UTR, phải được chọn. Nguyên tắc để tránh sự bất hoạt tương đồng chéo

đường như là sử dụng các trình tự không có các khối đồng nhất trình tự lớn hơn 20 bazơ giữa cấu trúc và các trình tự gen không phải đích. Nhiều nguyên tắc giống như vậy được áp dụng để chọn lọc các vùng đích để thiết kế các amiARN.

Kỹ thuật bất hoạt gen được cảm ứng bởi virus (VIGS) là sự biến đổi kỹ thuật ARNi mà lợi dụng hàng rào kháng virus nội sinh của cây. Sự lây nhiễm của cây với các virus VIGS tái tổ hợp chứa mảnh ADN chủ dẫn đến sự bất hoạt gen sau phiên mã đối với gen đích. Theo một phương án, hệ VIGS dựa trên virus hình khúc chuông ở cây thuốc lá (TRV) có thể được sử dụng với trình tự nucleotit theo sáng chế.

Kỹ thuật đối nghĩa bao gồm đưa vào cây oligonucleotit đối nghĩa mà sẽ liên kết với ARN thông tin (mARN) được sản xuất bởi gen quan tâm. Oligonucleotit "đối nghĩa" có trình tự bazơ bổ sung với ARN thông tin (mARN) của gen, mà được gọi là trình tự "có nghĩa". Hoạt tính của mảnh có nghĩa của mARN bị phong bế bởi mảnh mARN đối nghĩa, bằng cách đó bất hoạt một cách hiệu quả sự biểu hiện gen. Việc áp dụng đối nghĩa để làm bất hoạt gen ở cây được mô tả chi tiết hơn bởi Stam M, de Bruin R, van Blokland R, van der Hoorn RA, Mol JN, Kooter JM (2000) Distinct features of post-transcriptional gene silencing by antisense transgenes in single copy and inverted T-DNA repeat loci. *Plant J.* 21:27-42.

Kỹ thuật đồng ức chế có nghĩa bao gồm việc đưa gen chuyển có nghĩa được biểu hiện cao vào cây dẫn đến sự biểu hiện giảm của cả gen chuyển và gen nội sinh (Depicker A, Montagu MV (1997) Post-transcriptional gene silencing in plants. *Curr Opin Cell Biol.* 9: 373-82). Hiệu quả thu được tùy thuộc vào độ đồng nhất trình tự giữa gen chuyển và gen nội sinh.

Kỹ thuật phát sinh đột biến có đích, ví dụ, TILLING (Targeting Induced Local Lesions IN Genomes) và "khuyết gen" sử dụng sự bắn phá neutron nhanh, có thể được sử dụng để làm bất hoạt chức năng gen ở sinh vật (Henikoff S, Till BJ, Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol* 135:630-6; Li X, Lassner M, Zhang Y. (2002)

Deleteagene: fast neutron deletion mutagenesis-based gene knockout system for plants. *Comp Funct Genomics*. 3: 158-60). TILLING bao gồm việc xử lý chất mầm hoặc các tế bào riêng rẽ bằng chất gây đột biến để gây ra các đột biến điểm mà sau đó được phát hiện trong các gen quan tâm bằng cách sử dụng phương pháp nhảy để phát hiện đột biến đơn nucleotit. Việc phát hiện các đột biến mong muốn (ví dụ, các đột biến dẫn đến sự bất hoạt sản phẩm gen quan tâm) có thể được thực hiện, ví dụ, bởi phương pháp PCR. Ví dụ, có thể điều chế đoạn mồi oligonucleotit thu được từ gen quan tâm và có thể sử dụng PCR để khuếch đại các vùng gen quan tâm từ sinh vật trong quần thể đã được làm đột biến. Các gen đột biến đã được khuếch đại có thể được ủ với các gen kiểu dại để phát hiện các bắt cặp sai giữa các gen đột biến và gen kiểu dại. Các sai lệch đã được phát hiện có thể được truy ngược đến sinh vật mà có gen đột biến nhờ đó phát hiện sinh vật đã được làm đột biến nào sẽ có sự biểu hiện mong muốn (ví dụ, bất hoạt gen quan tâm). Sau đó, các sinh vật này có thể được nhân giống một cách chọn lọc để tạo ra quần thể có sự biểu hiện mong muốn. TILLING có thể tạo ra dãy alen bao gồm các đột biến sai nghĩa và các đột biến làm bất hoạt, mà thể hiện sự biểu hiện giảm của gen đã được nhắm đích. TILLING được xem là phương pháp thích hợp để bất hoạt gen mà không liên quan đến việc đưa vào gen chuyển, và do đó có thể dễ được chấp nhận hơn đối với người tiêu dùng. Sự bắn phá neutron nhanh sẽ kích thích các đột biến, tức là, khuyết đoạn, trong hệ gen sinh vật, các khuyết đoạn này cũng có thể được phát hiện bằng cách sử dụng PCR theo cách tương tự với TILLING.

Người có hiểu biết trung bình về lĩnh vực kỹ thuật này sẽ hiểu rằng các quy trình theo sáng chế cũng có thể được tiến hành trong môi trường không có tế bào trong điều kiện có mặt một hoặc nhiều axit carboxylic.

Các phương án theo sáng chế có thể có các cải biến và các dạng thay thế khác ngoài các ví dụ cụ thể được bao gồm trong bản mô tả này. Như vậy, các phương án theo sáng chế không bị giới hạn ở các dạng cụ thể đã được bộc lộ.

Ví dụ thực hiện sáng chế

Ví dụ 1: Khuếch đại và tách dòng CsHCS1 và CsHCS2

CsHCS1 và CsHCS2 được khuếch đại bằng PCR từ các dòng cADN plasmit sử dụng các đoạn môi được liệt kê trong Bảng 3 và Phusion™ polymeraza (Finnzymes). Các sản phẩm PCR được tinh chế và được tách dòng vào vector tiếp nhận pCR8/GW/TOPO (Invitrogen). Sau khi biến nạp vào tế bào *E. coli* TOP10 (Invitrogen), các dòng riêng rẽ được kiểm tra bằng cách giải trình tự. Gen *CsHCS1* và *CsHCS2* được tái tổ hợp vào vector đích pHIS8/GW sử dụng LR recombinaza (Invitrogen). Sản phẩm phản ứng LR được biến nạp vào các tế bào TOP10 và được kiểm tra bằng cách giải trình tự.

Bảng 3: Oligonucleotit

Tên	Trình tự (5'-3')
CsHCS1 xuôi (SEQ ID NO: 5)	ATGGGTAAGAATTACAAGTCCCT
CsHCS1 ngược (SEQ ID NO: 6)	GAGCTCTCATTCAAAGTGAGAAAAT TGCTG
CsHCS2 xuôi (SEQ ID NO: 7)	ATGGAGAAATCTGGGTATGGAAG
CsHCS2 ngược (SEQ ID NO: 8)	TCACATGTTGGAGCGTACTTTC
MCS xuôi (SEQ ID NO: 9)	ATGAGCAACCATCTTTTCGACG
MCS ngược (SEQ ID NO: 10)	TTACGTCCTGGTATAAAGATCGGC

pHIS8/GW-CsHCS1 và pHIS8/GW-CsHCS2 được biến nạp vào các tế bào *E. coli* Rosetta 2 (Merck). Các khuẩn lạc riêng rẽ được sử dụng để cấy các giống cấy quy mô nhỏ của môi trường LB lỏng chứa cloramphenicol và kanamycin, các giống cấy này được sử dụng để cấy 500 mL môi trường LB lỏng không có chất kháng sinh. Sau khi sinh trưởng đến OD₆₀₀ bằng 0,6, sự biểu hiện được cảm ứng bằng cách bổ sung IPTG đến nồng độ 0,2 µM. Sau đó, các giống cấy biểu hiện CsHCS1 được sinh trưởng ở nhiệt độ 12°C có lắc trong 24 giờ, trong khi đó các giống cấy CsHCS2 được sinh trưởng ở nhiệt độ 37°C trong 16 giờ. Các nhiệt độ khác nhau được sử dụng vì quan sát thấy rằng CsHCS1 không sản xuất protein hòa tan ở nhiệt độ cao hơn.

Các tế bào được thu hoạch bằng cách ly tâm và được tái tạo huyền phù trong 10 mL đệm tan His-tag (Tris-HCl 50 mM pH 7, NaCl 500 mM, imidazol 2,5 mM, glyxerol 10% theo thể tích, β-mercaptoetanol 10 mM, Tween™ 20 1%

theo thể tích, và lysozym 750 $\mu\text{g/mL}$. Các tế bào đã được tái tạo huyền phù được ủ trên đá trong 1 giờ sau đó được làm tan bằng cách khuấy âm. Sau khi ly tâm trong 20 phút ở tốc độ 12.000 g ở nhiệt độ 4°C, phân đoạn protein hòa tan được bổ sung vào 160 μL nhựa Talon™ (Clontech) đã được rửa trước bằng đệm rửa His-tag (HWB; Tris-HCl 50 mM pH 7, NaCl 500 mM, imidazol 2,5 mM, glyxerol 10%, β -mercaptoetanol 10 mM). Các mẫu được ủ có lắc nhẹ ở nhiệt độ 4°C, sau đó nhựa được tách riêng bằng cách ly tâm (700 g trong 5 phút). Nhựa được tái tạo huyền phù trong đệm HWB và được rửa có lắc nhẹ ở nhiệt độ 4°C sau đó được ly tâm. Sau đó, bước rửa được lặp lại hai lần và nhựa đã được tái tạo huyền phù được nạp lên trên cột sắc ký và được để ráo nước. Sau khi rửa nhựa bằng 10 mL đệm HWB, protein đã được gắn thẻ His được rửa giải bằng cách bổ sung 2,5 mL đệm rửa giải His-tag (Tris-HCl 50 mM pH 7, NaCl 500 mM, imidazol 250 mM, glyxerol 10% theo thể tích, β -mercaptoetanol 10 mM). Nước giải hấp được trao đổi đệm vào đệm bảo quản (HEPES 50 mM pH 9, glyxerol 10% theo thể tích, MgCl_2 2 mM, và dithiothreitol 2 mM) sử dụng cột PD10 (Amersham Biosciences). Độ tinh khiết của các protein đã được tách riêng được kiểm tra bởi SDS-PAGE, và nồng độ protein được xác định bằng thử nghiệm Bradford.

Ví dụ 2: Phân tích hoạt tính hexanoyl-CoA synthetaza

Hoạt tính hexanoyl-CoA synthetaza được xác định bằng cách ủ 0,1 μg enzym trong 20 μL hỗn hợp phản ứng chứa HEPES 50 mM pH 9, ATP 8 mM, MgCl_2 10 mM, CoA 0,5 mM, và natri hexanoat 4 mM. Các phản ứng được ủ trong 10 phút ở nhiệt độ 40°C, được kết thúc bằng 2 μL HCl 1 N và được bảo quản trên đá cho đến khi phân tích.

Hỗn hợp phản ứng được pha loãng 1:100 với nước và sau đó được tách riêng bằng cách sử dụng hệ Waters Acquity UPLC được lắp cột Acquity UPLC BEH C18 (cỡ hạt 1,7 μm , cột 2,1 x 50 mm), và được phân tích bởi MS/MS sử dụng khối phổ kế ba-bốn cực Micromass Quattro Ultima. Hệ dung môi được sử dụng là đệm A: TEA 5 mM và axit axetic 3 mM trong nước, và đệm B: TEA 5 mM và axit axetic 3 mM trong 95:5 metanol:nước. Chương trình dòng chảy

được thể hiện trong Bảng 4. Các thiết lập của khối phổ kế là: Kiểu dương ESI, năng lượng va chạm 27 V, điện áp hình nón 135 V, quét trong 866>359 lần chuyển tiếp.

Bảng 4: Chương trình dòng chảy đối với sắc ký chất lỏng

Thời gian (phút)	Tốc độ dòng chảy (mL/phút)	%A	%B
0	0,2	94,7	5,3
1	0,2	94,7	5,3
6	0,2	85,0	15,0
10	0,2	0,1	50,0
21	0,2	0,1	99,9
22	0,8	0,1	99,9
24	0,8	0,1	99,9
24,1	0,4	94,7	5,3
27	0,4	94,7	5,3
27,1	0,2	94,7	5,3

Như được thể hiện trên Fig.2, CsHCS1 và CsHCS2 xúc tác sự tạo thành hexanoyl-CoA từ hexanoat và CoA. Fig. 2A và 2B thể hiện sự rửa giải của chuẩn hexanoyl-CoA tin cậy. Fig.2C thể hiện mẫu thử nghiệm đầy đủ bao gồm CsHCS1, HEPES 50 mM pH 9, ATP 8 mM, MgCl₂ 10 mM, CoA 0,5 mM, và natri hexanoat 4 mM. Hợp chất có cùng sự chuyển tiếp khối lượng và thời gian rửa giải như chuẩn hexanoyl-CoA tin cậy có thể được quan sát thấy tại thời điểm 9,25 phút. Fig.2D thể hiện mẫu thử nghiệm đầy đủ bao gồm CsHCS2, HEPES 50 mM pH 9, 8 mM ATP, MgCl₂ 10 mM, CoA 0,5 mM, và natri hexanoat 4 mM. Hợp chất có cùng sự chuyển tiếp khối lượng và thời gian rửa giải như chuẩn hexanoyl-CoA tin cậy có thể được quan sát thấy tại thời điểm 9,25 phút. Fig. 2E và 2F thể hiện các đối chứng âm với các enzym CsHCS1 và CsHCS2 được làm bất hoạt (được đun sôi). Như có thể được thấy trên các Fig. 2E và 2F, các thử nghiệm này cho thấy không có sự tổng hợp hexanoyl-CoA.

Cả CsHCS1 và CsHCS2 đều thể hiện sự tối ưu của nhiệt độ và độ pH tương ứng là 40°C và pH 9. Trong thử nghiệm khoảng cation hóa trị hai, CsHCS1 sử dụng một cách tối ưu Mg²⁺ và sử dụng Mn²⁺ và Co²⁺ ở mức thấp hơn. Hoạt tính CsHCS2 là cao nhất khi sử dụng Co²⁺, nhưng cũng được quan sát thấy là cao đối với Mg²⁺, Mn²⁺, và ở mức thấp hơn đối với Ca²⁺. Mỗi liên

quan về mặt sinh học của hoạt tính cao đối với Co^{2+} là không rõ ràng và Mg^{2+} được sử dụng cho tất cả các thử nghiệm tiếp theo.

Với hexanoat, CsHCS1 có K_m bằng $6,1 \pm 1,0$ mM, $V_{\max}$ bằng $15,6 \pm 1,7$ pKat và k_{cat} bằng $4,5$ giây⁻¹. CsHCS2 có K_m bằng 320 nM, $V_{\max}$ bằng 1,7 pKat, và k_{cat} bằng $57,6$ giây⁻¹.

Ví dụ 3: Thử nghiệm với các axit carboxylic khác nhau

Để thử nghiệm khoảng axit carboxylic mà CsHCS1 và CsHCS2 có thể hoạt hóa, các thử nghiệm enzym được tiến hành với khoảng rộng của axit carboxylic và ATP hạn chế. Các điều kiện thử nghiệm được sử dụng là tương tự với Schneider K et al. (2005). A new type of peroxisomal acyl-coenzyme A synthetase from *Arabidopsis thaliana* has the catalytic capacity to activate biosynthetic precursors of jasmonic acid. The Journal of Biological Chemistry, 280:13962-72. Nói một cách ngắn gọn, enzym HCS đã được tinh chế (1 μg) được ủ với cơ chất axit carboxylic 500 μM và CoA 100 μM trong mẫu thử nghiệm chứa HEPES 100 mM pH 9, MgCl_2 250 μM , ATP 50 μM và DTT 1 mM. Tất cả các cơ chất axit carboxylic đều được hòa tan trong 2% theo thể tích Triton™ X-100, dẫn đến nồng độ cuối cùng là 0,05% Triton X-100 trong mẫu thử nghiệm. Sau khi cho phản ứng trong 3 giờ ở nhiệt độ 29°C, các lượng nhỏ 10 μL của các phản ứng được chuyển sang đĩa có 96 lỗ để xác định ATP chưa được tiêu thụ dựa trên luxiferin/luxiferaza. Các đĩa được phân tích bằng máy đếm 1420 Multilabel (PerkinElmer). 90 μL dung dịch chứa Tris 100 mM pH 7,8, MgCl_2 1 mM, 2,3 μg luxiferin, và 0,5 μg luxiferaza được đưa vào mỗi lỗ, và sau khi lắc trong 2 giây, sự phát quang được xác định trong 15 giây mà không có bộ lọc tại chỗ. Các chỉ số thấp hơn, so với phản ứng không có cơ chất axit carboxylic, thể hiện lượng hoạt tính enzym cao hơn và do đó việc sử dụng cơ chất cao hơn. Kết quả được thể hiện trên Fig.3, trong đó thanh báo lỗi biểu thị % mức độ sai số, $n=3$.

Như được thể hiện trên Fig.3A, CsHCS1 được quan sát thấy là sử dụng hexanoat (sáu cacbon), octanoat (tám cacbon), và nonanoat (chín cacbon), và

sử dụng ở mức ít hơn valerat (năm cacbon) và heptanoat (bảy cacbon), làm cơ chất. Trái lại, như được thể hiện trên Fig.3B, CsHCS2 thể hiện sự hỗn tạp lớn hơn, và có thể sử dụng khoảng cơ chất rộng, từ propanoat (C3) đến arachidoat (C20), và nhiều phenylpropanoit (xinnamat, ferulat, và ở mức ít hơn *p*-coumarat).

Trong thử nghiệm riêng rẽ, các tính chất động lực học của CsHCS1 và CsHCS3 được xác định một cách chính xác hơn đối với CoA và các axit béo mạch ngắn (butanoat), trung bình (hexanoat và decanoat), và mạch dài (palmitat) đại diện (Bảng 5). Nồng độ CoA cao ức chế CsHCS1. Sử dụng mô hình thử nghiệm sự ức chế cơ chất hồi quy phi tuyến tính, K_i của CsHCS1 được ước tính là $5,101 \pm 1,8$ mM. CoA không ức chế CsHCS2 ở các nồng độ được thử nghiệm. Axit decanoic ức chế CsHCS2 ($K_i = 120,8 \pm 47,9$ μ M) và ức chế không đáng kể CsHCS1 ở các nồng độ lớn hơn 4 mM (K_i không được xác định). Các số liệu này cho thấy xu hướng giống như số liệu động lực học đã được đưa ra trên đây.

Bảng 5: Các tính chất động lực học của CsHCS1 và CsHCS2.

Cơ chất	CsHCS1		CsHCS2		
	K_m	V_{max} (pKat)	K_m	V_{max} (pKat)	k_{cat} (giây ⁻¹)
CoA	$0,26 \pm 0,05$ μ M	-	$0,16 \pm 0,01$ μ M	-	-
butanoat	>10 mM	-	>10 mM	-	-
hexanoat	$3,7 \pm 0,7$ mM	$6,8 \pm 0,7$	261 ± 37 μ M	$1,8 \pm 0,05$	57,6
decanoat	$1,7 \pm 0,2$ mM	$1,8 \pm 0,7$	$16,1 \pm 5,8$ μ M	$1,6 \pm 0,1$	10,0
palmitat	n.d. ^a	-	$1,3 \pm 0,5$ μ M	$0,4 \pm 0,01$	2,4

^a không được xác định do thiếu hoạt tính xúc tác

Ví dụ 4: Tổng hợp axit olivetolic sử dụng CsHCS2

CsHCS2 được sử dụng để tổng hợp hóa enzym axit polyketit olivetolic thơm. Axit olivetolic là tiền chất thích hợp đầu tiên cho quá trình sinh tổng hợp canabinoit. Quy trình tổng hợp *in vitro* này được tiến hành bằng cách sử dụng bốn enzym tái tổ hợp: CsHCS2, malonyl-CoA synthetaza (MCS) từ *Rhizobium*

leguminosarum, “olivetol syntaza”/polyketit syntaza từ cây gai dầu, và olivetolic axit syntaza từ cây gai dầu.

Malonyl-CoA synthetaza (MCS) được khuếch đại từ ADN hệ gen của *Rhizobium leguminosarum* với đoạn mã MCS xuôi và MCS ngược (xem Bảng 3). Sản phẩm PCR được tách dòng vào vector pCR8/GW/TOPO (Invitrogen) được tái tổ hợp vào vector pHIS8/GW. Sau khi kiểm tra bằng cách giải trình tự, plasmit được biến nạp vào *E. coli* Rosetta II (DE3). MCS tái tổ hợp được biểu hiện và được tinh chế như được nêu đối với enzym CsHCS.

Việc tách dòng, biểu hiện và tinh chế “olivetol syntaza”/polyketit syntaza và olivetolic axit syntaza được tiến hành như sau. Để biểu hiện trong các tế bào *E. coli*, khung đọc mở của polyketit syntaza/olivetol syntaza và olivetolic axit syntaza được khuếch đại bởi PCR, được tách dòng vào pHIS8/GW đối với polyketit syntaza/olivetol syntaza hoặc pET100 (Invitrogen) đối với olivetolic axit syntaza và được biến nạp vào *E. coli* BL21 (DE3) (Invitrogen). Việc tách dòng được kiểm tra bằng cách giải trình tự.

Olivetolic axit syntaza được biểu hiện trong môi trường nuôi cấy canh lớn dung tích 200 mL trong khi polyketit syntaza/olivetol syntaza được sinh trưởng trong môi trường nuôi cấy dung tích 1 L. Cả hai môi trường nuôi cấy đều được ủ ở nhiệt độ 30°C/lắc 150 vòng/phút, được cảm ứng bằng IPTG 0,5 μ M và được sinh trưởng qua đêm. Các môi trường nuôi cấy được ly tâm ở tốc độ 16.000 g trong 20 phút, và các viên tròn được làm tan bằng cách xử lý bằng lysozym và khuấy âm. Các dịch tan đã được làm trong được trộn với nhựa Talon (200 μ L đối với olivetolic axit syntaza, 1 mL đối với polyketit syntaza/olivetol syntaza; Clontech), rửa bằng 5 mL đệm rửa His-tag (Tris-HCl 50 mM (pH 7), NaCl 150 mM, imidazol 20 mM, β -mercaptoetanol 10 mM) và các protein tái tổ hợp được rửa giải bằng cách sử dụng đệm rửa giải His-tag (Tris HCl 20 mM (pH 7), NaCl 150 mM, imidazol 100 mM, β -mercaptoetanol 10 mM). Nước giải hấp được cô bằng cách sử dụng thiết bị cô YM10 và được trao đổi đệm với đệm bảo quản (HEPES 20 mM (pH 7,5), NaCl 25 mM, glyxerol 10%, DTT 5 mM). Dung dịch protein cuối cùng được định lượng bằng

cách sử dụng bộ kit thử nghiệm protein RC/DC (Bio-Rad) mà phát hiện thấy nồng độ protein bằng 0,5 mg/mL (olivetolic axit syntaza) và 5,6 mg/mL (polyketit syntaza/olivetol syntaza). Phân tích gel SDS-PAGE xác nhận độ tinh khiết của cả hai protein.

Khả năng kết hợp sự tổng hợp hexanoyl-CoA với sự tổng hợp polyketit thơm sử dụng các chất phản ứng không đắt tiền được thử nghiệm bằng cách thực hiện các mẫu thử nghiệm enzym gồm hexanoat 4 mM, malonat 8 mM, CoA 0,4 mM, ATP 0,4 mM, MgCl₂ 5 mM, DTT 2 mM, HEPES 20 mM pH 7,5, 0,3 µg CsHCS2, 12 µg MCS, 8 µg “olivetol syntaza”/PKS và 10 µg OAS. Phản ứng được ủ ở nhiệt độ trong phòng trong 16 giờ, được axit hóa và được chiết trong etyl axetat. Phân đoạn phân cực được thu hồi, được làm bay hơi cho đến khô và được tái tạo huyền phù trong 50 µL metanol. Lượng nhỏ (5 µL) được phân tích bằng LCMS, và các sản phẩm được xác định bởi thời gian lưu và khối lượng của chúng (xem Fig.4).

Ví dụ 5: Tổng hợp axit olivetolic trong nấm men sử dụng CsHCS1 và CsHCS2

Nấm men (*Saccharomyces cerevisiae*) được biến đổi di truyền để sản xuất axit olivetolic bằng cách sử dụng CsHCS1 và CsHCS2 để tổng hợp hexanoyl-CoA, và thể dung hợp “olivetol syntaza”/polyketit syntaza (PKS) và olivetolic axit syntaza (OAS) của cây gai dầu để tạo thành axit olivetolic.

OAS được tách dòng trong khung với “olivetol syntaza”/PKS sử dụng trình tự liên kết tổng hợp mã hóa cho axit amin AATSGTGSTGSTGSGRSTGSTGSTGSGRSHMV (SEQ ID NO: 11) trong vector biểu hiện của nấm men pESC-Trp (Stratagene) dưới sự kiểm soát của gen khởi đầu GAL10. Khung đọc mở của CsHCS1 được tách dòng vào vector biểu hiện của nấm men pYESDEST52-Ura (Invitrogen) sử dụng công nghệ Gateway. Khung đọc mở của CsHCS2 được tách dòng vào vector biểu hiện của nấm men pESC-Trp (Stratagene) dưới sự kiểm soát của gen khởi đầu GAL1.

Các tế bào nấm men (InvSc I, Invitrogen) được biến nạp với các cấu trúc nêu trên (thể dung hợp OAS:“olivetol syntaza”/PKS một mình, thể dung hợp OAS:“olivetol syntaza”/PKS và CsHCS1; thể dung hợp OAS:“olivetol syntaza”/PKS và CsHCS2) và các thể biến nạp được sinh trưởng trên đĩa SD-Trp trong 3 ngày ở nhiệt độ 28°C. Đối với mỗi thể biến nạp, khuẩn lạc đơn được cấy vào 3 mL môi trường SD-Trp glucoza và được ủ có lắc ở nhiệt độ 28°C trong 2 ngày. Lượng nhỏ 0,5 mL của giống cấy mỗi được sử dụng để cấy 10 mL môi trường SD-Trp galactoza chứa natri hexanoat 1 mM và được ủ ở nhiệt độ 20°C trong 4 ngày. Giống cấy hoàn toàn được chiết với etyl axetat, được làm khô và phần cặn được tái tạo huyền phù trong 100 µL của 30% axetonitril/70% nước/0,05% axit formic. Các sản phẩm được phân tích bằng cách sử dụng LCMS (xem Fig.5).

Ví dụ 6: Vai trò của CsHSC1 và CsHSC2 trong quá trình sinh tổng hợp cannabinoid ở cây

Thông qua phân tích dựa trên trình tự của tập dữ liệu EST tằm lông và thử nghiệm hóa sinh của năm AAE, hai AAE có hoạt tính hexanoyl-CoA synthetaza được xác định (CsHSC1 và CsHSC2). Để xác định AAE nào trong số hai AAE này có khả năng có liên quan đến con đường sinh tổng hợp cannabinoid, tiến hành qRT-PCR và thử nghiệm định vị dưới tế bào.

qRT-PCR của sự biểu hiện CsHSC1 và CsHSC2

Các cây ‘Finola’ được phát triển từ hạt cho đến giai đoạn giữa thời kỳ ra hoa. Rễ, thân, lá, hoa cái (với các tằm lông và sau khi tách riêng các tằm lông sử dụng Beadbeater), các tế bào tằm lông và hoa đực được lấy mẫu từ ba cây. ARN tổng cộng được tách riêng như nêu trên. ARN có $A_{260}:A_{280} > 1,9$ và thể hiện các băng ribosom riêng biệt trên gel biến tính. cADN sợi thứ nhất được tổng hợp bằng cách sử dụng 0,5 µg ARN với bộ kit tổng hợp cADN QuantiTect (Qiagen). Mỗi 20 µL mẫu cADN được pha loãng 1:4 với nước, và 1 µL được sử dụng làm khuôn PCR. Các đoạn mồi đặc hiệu gen được thiết kế để sản xuất các đơn vị siêu sao chép có kích thước 90-200 bp. Các phản ứng PCR (20 µL) được tiến hành trong đĩa có 96 lỗ sử dụng thử nghiệm dựa trên SYBR màu lục

(QuantiFast SYBR Green kit, Qiagen) với thiết bị StepOne Plus (Applied Biosystems). Các thông số của chu kỳ được sử dụng là nhiệt độ 95°C trong 5 phút sau đó là 40 chu kỳ ở nhiệt độ 95°C trong 10 giây, 60°C trong 30 giây, và protocol phân ly chuẩn (95°C 15 giây, 60°C trong 1 phút, 60-95°C với số gia 0,3°C trong 15 giây). Các thử nghiệm được tiến hành bằng cách sử dụng cADN từ ba cây với hai bản sao kỹ thuật. *Actin*, mà được thấy là có sự biểu hiện ổn định trong tất cả mô được thử nghiệm, được sử dụng làm gen tham chiếu. Hiệu quả đối với tất cả các cặp đoạn mồi là 90-110% như được tính toán bằng cách sử dụng phương pháp đường cong chuẩn. Các giá trị C_t được tính toán bằng cách sử dụng phần mềm StepOne Software (Applied Biosystems). Phương pháp $2^{-\Delta\Delta C_t}$ được sử dụng để phân tích sự biểu hiện gen tương đối.

Sự định vị dưới tế bào của CsHSC1 và CsHSC2

Các thể dung hợp YFP:CsHSC1 và YFP:CsHSC2 được tạo cấu trúc bằng cách tái tổ hợp vào pEARLYGATE104 (Earley, K.W., Haag, J.R., Pontes, O., Oppen, K., Juehne, T., Song, K. and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J.* 45, 616-629.) sử dụng LR recombinaza (Invitrogen). Để tạo ra cấu trúc OLS:CFP, OLS thiếu codon kết thúc được tách dòng vào pCR8/GW/TOPO trước khi tái tổ hợp vào pEARLYGATE102 sử dụng LR recombinaza. Gen đánh dấu peroxisom PX-CK (Nelson, B.K., Cai, X. and Nebenführ, A. (2007). Multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *Plant J.* 51, 1126-1136.) là từ ABRC (www.arabidopsis.org). Các plasmit được biến nạp vào *Agrobacterium tumefaciens* GV3101 bằng phương pháp đục lỗ điện và được lựa chọn trên đĩa LB chứa 10 µg/mL rifampaxin và 50 µg/mL kanamycin. Lá của các cây *Nicotiana benthamiana* hai tuần tuổi được ngâm lọc dung dịch *Agrobacterium* ở OD₆₀₀ bằng 0,02 (Sparkes, I.A., Runions, J., Kearns, A. and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protocol.* 1, 2019-2025.). Hai ngày sau khi ngâm lọc, các tế bào biểu bì lá được hiển thị hóa bằng

cách sử dụng kính hiển vi đồng tụ Zeiss LSM510. CFP được hiển thị hóa với sự kích thích ở 458 nm và thu thập hình ảnh với bộ lọc lấy dải 475-525 nm; YFP ở 514 nm với bộ lọc lấy dải 530-600 nm. Các hình ảnh được thu gom và được phân tích bằng cách sử dụng gói phần mềm Zeiss LSM.

Số liệu chứng tỏ rằng CsHSC1 là enzyme có liên quan đến quá trình sinh tổng hợp cannabinoids. CsHSC1 là thể phiên mã dồi dào nhất trong danh mục EST, và số liệu qRT-PCR cho thấy rằng sự biểu hiện của nó lớn hơn 100 lần trong các tế bào tủy lông so với các mô khác (Fig.6). Ngoài ra, CsHSC1 được định vị ở phần bào tan như được chứng tỏ bởi thử nghiệm định vị dưới tế bào, trong cùng ngăn mà enzyme cannabinoids giả định OLS được định vị. Sự ưu tiên cơ chất của CsHSC1 cung cấp bằng chứng bổ sung về vai trò của nó trong quá trình sinh tổng hợp cannabinoids vì nó thể hiện tính đặc hiệu đối với hexanoat và các axyl CoA béo mạch ngắn khác lớn hơn so với CsHSC2 (Fig.3).

Mặc dù CsHSC1 là enzyme có khả năng có liên quan đến quá trình sinh tổng hợp cannabinoids ở cây, CsHSC2 là hiệu quả hơn so với CsHSC1 đối với việc tổng hợp hexanoyl-CoA. Tuy nhiên, CsHSC2 được định vị ở peroxisom và không rõ cách hexanoyl-CoA được tạo thành trong ngăn này có thể được xuất đến bào lập ở đó pha tổng hợp polyketit của con đường cannabinoids được định vị. CsHSC2 chấp nhận khoảng cơ chất rất rộng, cho thấy rằng nó là axyl-CoA synthetaza phổ biến hơn mà có thể thực hiện chức năng trong quá trình β -oxy hóa peroxisom.

Cả CsHSC1 và CsHSC2 đều là công cụ công nghiệp có giá trị. Việc làm bất hoạt CsHSC1 ở cây gai dầu có thể dẫn đến sự giảm đáng kể mức cannabinoids ở cây, điều này là rất mong muốn đối với các nhà nhân giống cây gai dầu. Sự biểu hiện quá mức của CsHSC1 ở cây gai dầu có thể dẫn đến mức cannabinoids cao, điều này là hữu ích cho mục đích dược phẩm. Mặt khác, CsHSC2 sẽ đặc biệt hữu ích để thiết lập lại sự tạo thành cannabinoids ở các vi sinh vật hoặc trong hệ *in vitro*.

Ví dụ 7: Tạo ra các thể đột biến trong các gen CsHSC1 và CsHSC2 sử dụng các tổn thương cục bộ cảm ứng đích trong hệ gen (TILLING)

Việc nhận diện cây gai dầu có các thể đột biến trong các gen CsHSC1 hoặc CsHSC2 có thể được thực hiện bằng cách sử dụng TILLING. Quần thể cây gai dầu đã được làm đột biến được sàng lọc bằng cách sử dụng đoạn mồi oligonucleotit và PCR để khuếch đại các gen quan tâm. Các gen đột biến đã được khuếch đại được ủ với các gen kiểu dại để phát hiện các bắt cặp sai giữa các gen đột biến và các gen kiểu dại. Các sai khác đã được phát hiện được sử dụng để nhận diện cây chứa đột biến trong một trong hai gen CsHSC1 hoặc CsHSC2. Cây chứa đột biến mà dẫn đến các axit amin được biến đổi ở các vị trí cần thiết cho độ ổn định hoặc hoạt tính alkanoyl-CoA synthetaza của các protein CsHSC1 hoặc CsHSC2 không có khả năng sản xuất tiền chất alkanoyl-CoA cho quá trình sinh tổng hợp cannabinoids. Các cây thu được chứa mức giảm hoặc mức được làm biến đổi của sản phẩm cannabinoids.

Sáng chế đề xuất gen mã hóa hai enzym alkanoyl CoA synthetaza từ cây gai dầu. Các gen này có thể được sử dụng để tạo ra, thông qua việc nhân giống, phát sinh đột biến có đích hoặc biến đổi di truyền, cây gai dầu với sự sản xuất cannabinoids được tăng cường. Ngoài ra, việc làm bất hoạt hoặc gây câm gen theo sáng chế trong cây gai dầu có thể được sử dụng để phong bế quá trình sinh tổng hợp cannabinoids và bằng cách đó làm giảm sự sản xuất các cannabinoids như THCA, tiền chất của THC, trong cây gai dầu (ví dụ, cây gai dầu công nghiệp). Các gen theo sáng chế có thể được sử dụng, kết hợp với gen mã hóa cho các enzym khác trong con đường cannabinoids, để biến đổi di truyền quá trình sinh tổng hợp cannabinoids ở các cây khác hoặc ở các vi sinh vật hoặc trong hệ thống không có tế bào, hoặc để sản xuất chất tương tự của hợp chất cannabinoids hoặc chất tương tự của tiền chất cannabinoids.

Trong toàn bộ bản mô tả này, các tài liệu công bố được đưa vào bản mô tả này bằng cách viện dẫn.

Các ưu điểm khác vốn có của sáng chế là rõ ràng với người có hiểu biết trung bình về lĩnh vực kỹ thuật này. Các phương án được mô tả trong bản mô tả này nhằm mục đích minh họa và không nhằm mục đích giới hạn phạm vi của sáng chế như được yêu cầu bảo hộ. Các biến đổi của các phương án nêu trên sẽ

rõ ràng với người có hiểu biết trung bình về lĩnh vực kỹ thuật này và được dự định để bao gồm trong các điểm yêu cầu bảo hộ sau đây.

YÊU CẦU BẢO HỘ

1. Phân tử axit nucleic được phân lập hoặc được tinh chế chứa trình tự nucleotit có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 1, và mã hóa polypeptit có hoạt tính alkanoyl-CoA synthetaza.
2. Phân tử axit nucleic theo điểm 1, trong đó trình tự nucleotit là trình tự nucleotit thoái hóa codon của SEQ ID NO: 1.
3. Phân tử axit nucleic được phân lập hoặc được tinh chế, trong đó trình tự nucleotit là như được nêu trong SEQ ID NO: 1 hoặc trình tự nucleotit thoái hóa codon của nó.
4. Phân tử axit nucleic được phân lập hoặc được tinh chế chứa trình tự nucleotit có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 3, và mã hóa polypeptit có hoạt tính alkanoyl-CoA synthetaza.
5. Phân tử axit nucleic theo điểm 4, trong đó trình tự nucleotit là trình tự nucleotit thoái hóa codon của SEQ ID NO: 3.
6. Phân tử axit nucleic được phân lập hoặc được tinh chế, trong đó trình tự nucleotit là như được nêu trong SEQ ID NO: 3 hoặc trình tự nucleotit thoái hóa codon của nó.
7. Polypeptit được phân lập hoặc được tinh chế chứa trình tự axit amin có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 2 và có hoạt tính alkanoyl-CoA synthetaza.
8. Polypeptit theo điểm 7, trong đó trình tự axit amin là trình tự axit amin được thể bảo toàn của SEQ ID NO: 2.
9. Polypeptit được phân lập hoặc được tinh chế chứa trình tự axit amin có ít nhất 85% độ đồng nhất trình tự với SEQ ID NO: 4 và có hoạt tính alkanoyl-CoA synthetaza.
10. Polypeptit theo điểm 9, trong đó trình tự axit amin là trình tự axit amin được thể bảo toàn của SEQ ID NO: 4.

11. Hệ biểu hiện chứa phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6.
12. Tế bào chủ được biến nạp với phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6.
13. Quy trình tổng hợp alkanoyl-CoA bao gồm bước cho axit béo phản ứng với CoA trong sự có mặt của polypeptit theo điểm bất kỳ trong số các điểm từ 7 đến 10.
14. Quy trình theo điểm 13, trong đó axit béo có từ 2 đến 10 nguyên tử cacbon.
15. Quy trình theo điểm 13, trong đó axit béo là axit hexanoic và alkanoyl-CoA bao gồm hexanoyl-CoA.
16. Quy trình biến đổi hàm lượng của hợp chất canabinoit ở cây, tế bào hoặc mô sản xuất canabinoit bao gồm bước sử dụng phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6, hoặc một phần của nó, để làm bất hoạt ở cây, tế bào hoặc mô sản xuất canabinoit này gen mã hóa enzym xúc tác sự tổng hợp alkanoyl-CoA và là CsHCS1 hoặc CsHCS2, so với giống cây, tế bào hoặc mô sản xuất canabinoit tương tự sinh trưởng trong các điều kiện tương tự nhưng không sử dụng phân tử axit nucleic này để làm bất hoạt.
17. Quy trình biến đổi hàm lượng của hợp chất canabinoit ở cây, tế bào hoặc mô sản xuất canabinoit bao gồm bước làm đột biến gen ở cây, tế bào hoặc mô sản xuất canabinoit này, và sử dụng phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6 để lựa chọn cây, tế bào hoặc mô sản xuất canabinoit chứa các thể đột biến hoặc các biến thể của gen mã hóa enzym xúc tác sự tổng hợp alkanoyl-CoA và là CsHCS1 hoặc CsHCS2.
18. Quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô khác với cây, tế bào hoặc mô sản xuất canabinoit, trong đó sinh vật này không phải là người, bao gồm bước biểu hiện hoặc biểu hiện quá mức phân tử axit nucleic ngoại sinh theo điểm bất kỳ trong số các điểm từ 1 đến 6 ở sinh vật, tế bào hoặc mô này trong sự có mặt của axit hexanoic và CoA so với giống sinh vật, tế bào hoặc mô tương tự sinh trưởng trong các điều kiện tương tự

nhưng không có sự biểu hiện hoặc biểu hiện quá mức của phân tử axit nucleic ngoại sinh này,

trong đó phân tử axit nucleic ngoại sinh được biểu hiện hoặc biểu hiện quá mức kết hợp với sự biểu hiện hoặc biểu hiện quá mức của các axit nucleic khác mã hóa “olivetol syntaza”/PKS, olivetolic axit syntaza, prenyltransferaza thơm, và enzym được chọn từ THCA syntaza, CBDA syntaza, hoặc CBCA syntaza.

19. Quy trình biến đổi hàm lượng của hợp chất canabinoit ở sinh vật, tế bào hoặc mô khác với cây, tế bào hoặc mô sản xuất canabinoit, trong đó sinh vật này không phải là người, bao gồm bước biểu hiện hoặc biểu hiện quá mức phân tử axit nucleic ngoại sinh mã hóa polypeptit theo điểm bất kỳ trong số các điểm từ 7 đến 10 ở sinh vật, tế bào hoặc mô này trong sự có mặt của axit hexanoic và CoA so với giống sinh vật, tế bào hoặc mô tương tự sinh trưởng trong các điều kiện tương tự nhưng không có sự biểu hiện hoặc biểu hiện quá mức của phân tử axit nucleic ngoại sinh này,

trong đó phân tử axit nucleic ngoại sinh được biểu hiện hoặc được biểu hiện quá mức kết hợp với sự biểu hiện hoặc biểu hiện quá mức của các axit nucleic khác mã hoá “olivetol syntaza”/PKS, olivetolic axit syntaza, prenyltransferaza thơm và enzym được chọn từ THCA syntaza, CBDA syntaza, hoặc CBCA syntaza.

20. Quy trình theo điểm 18 hoặc 19, trong đó hàm lượng của hợp chất canabinoit được làm tăng.

21. Quy trình theo điểm 16 hoặc 17, trong đó hàm lượng của hợp chất canabinoit được làm giảm.

22. Quy trình tổng hợp hợp chất canabinoit ở sinh vật, tế bào hoặc mô khác với cây, tế bào hoặc mô sản xuất canabinoit, trong đó sinh vật này không phải là người, bao gồm bước biểu hiện phân tử nucleic ngoại sinh theo điểm bất kỳ trong số các điểm từ 1 đến 6 ở sinh vật, tế bào hoặc mô này trong sự có mặt của axit hexanoic và CoA,

trong đó phân tử axit nucleic ngoại sinh được biểu hiện hoặc biểu hiện quá mức kết hợp với sự biểu hiện hoặc biểu hiện quá mức của các axit nucleic khác mã hóa “olivetol syntaza”/PKS, olivetolic axit syntaza, prenyltransferaza thơm và enzym được chọn từ THCA syntaza, CBDA syntaza, hoặc CBCA syntaza.

23. Quy trình làm tăng sự sản xuất hợp chất canabinoit ở cây, tế bào hoặc mô sản xuất canabinoit bao gồm bước biểu hiện phân tử nucleic ngoại sinh theo điểm bất kỳ trong số các điểm từ 1 đến 6 ở cây, tế bào hoặc mô này trong sự có mặt của axit béo và CoA.

24. Quy trình theo điểm bất kỳ trong số các điểm từ 18 đến 20 và 22, trong đó sinh vật, tế bào hoặc mô là từ vi sinh vật.

25. Quy trình theo điểm 24, trong đó vi sinh vật là nấm men *Saccharomyces cerevisiae* hoặc *E. coli*.

26. Quy trình theo điểm bất kỳ trong số các điểm từ 16 đến 25, trong đó hợp chất canabinoit là một hoặc nhiều trong số axit canabigerolic, axit Δ^9 -tetrahydrocannabinolic, axit canabidiolic, axit canabichromenic, Δ^9 -tetrahydrocannabinol, canabidiol, canabichromen hoặc chất tương tự của chúng chứa mạch bên có chiều dài từ 1 đến 9 nguyên tử cacbon.

27. Quy trình tổng hợp alkanoyl-CoA trong phản ứng không có tế bào *in vitro*, quy trình này bao gồm bước cho axit béo phản ứng với coenzym A thông qua tác động của polypeptit theo điểm bất kỳ trong số các điểm từ 7 đến 10.

28. Cấu trúc chứa phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6.

29. Vector chứa phân tử axit nucleic theo điểm bất kỳ trong số các điểm từ 1 đến 6.

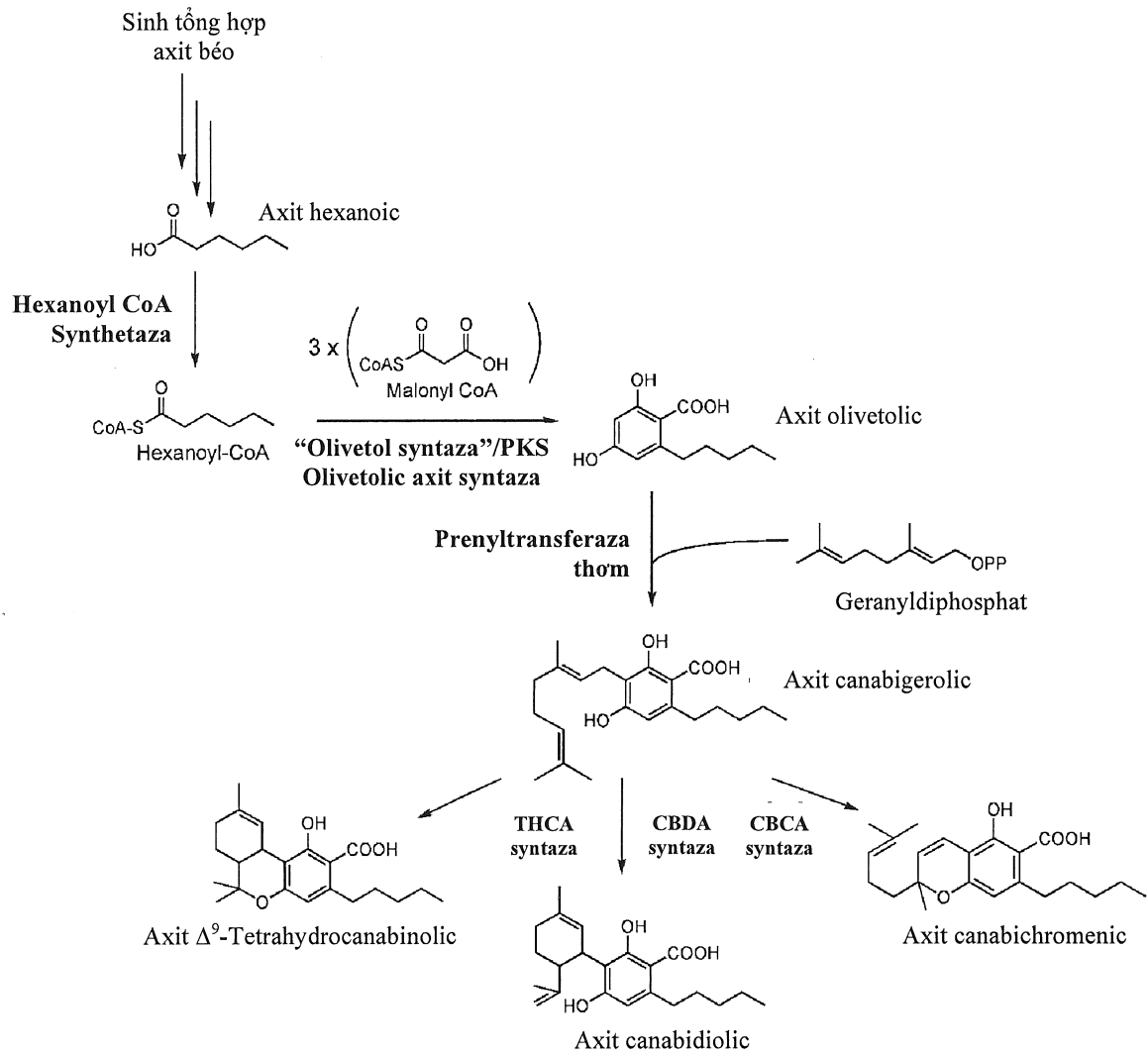


Fig. 1

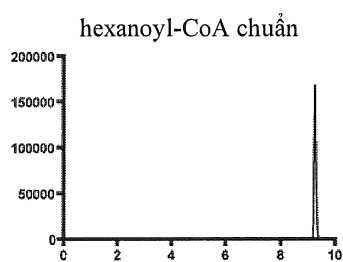


Fig. 2A

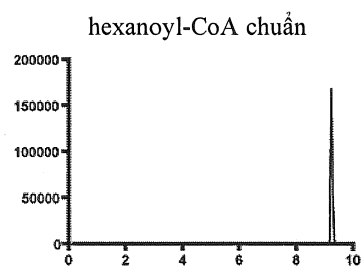


Fig. 2B

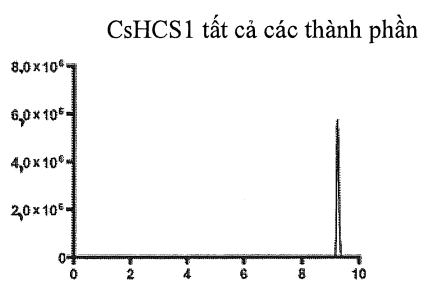


Fig. 2C

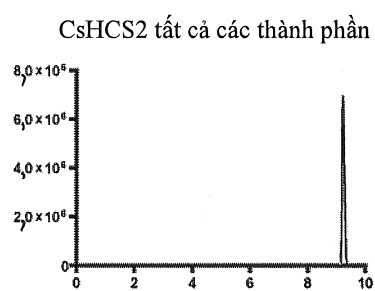


Fig. 2D

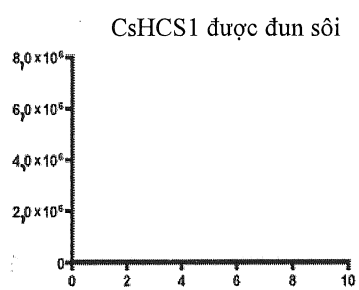


Fig. 2E

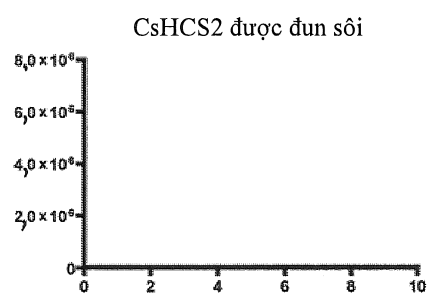


Fig. 2F

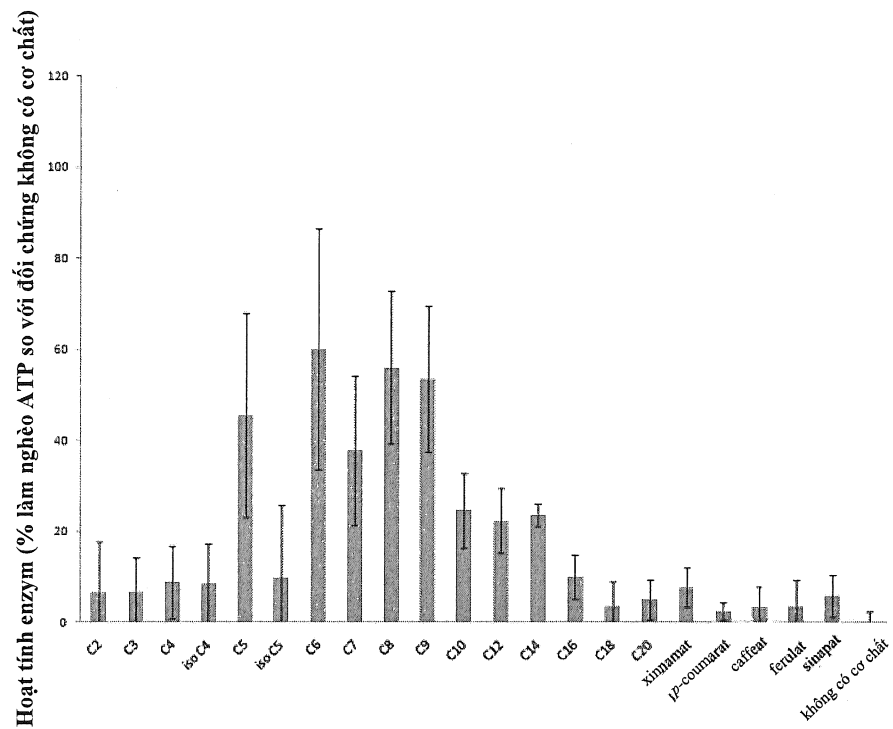


Fig. 3A

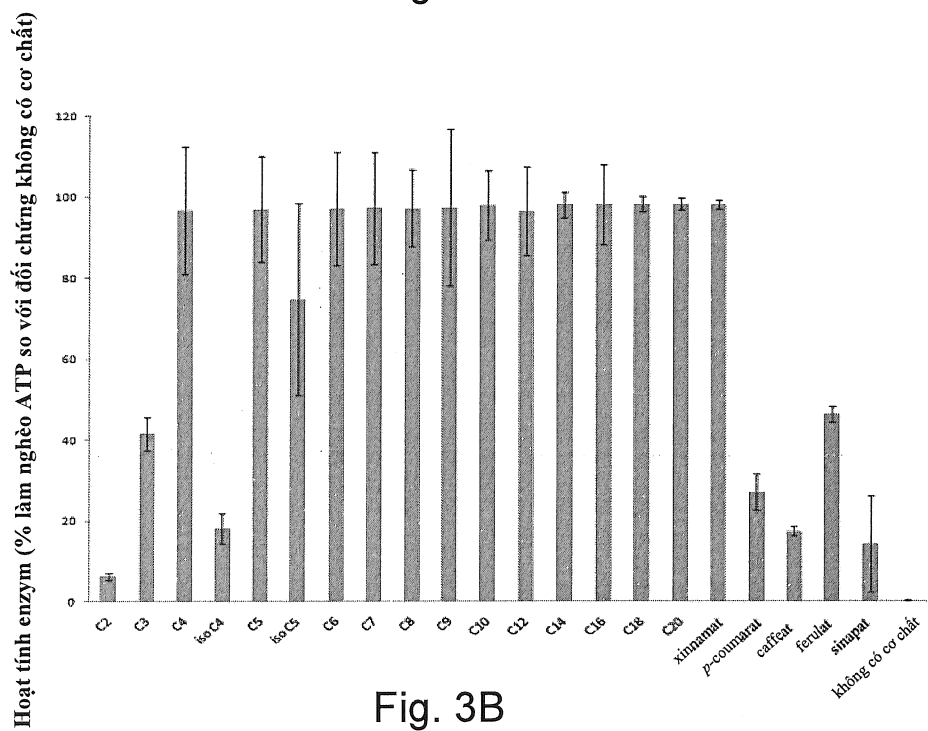


Fig. 3B

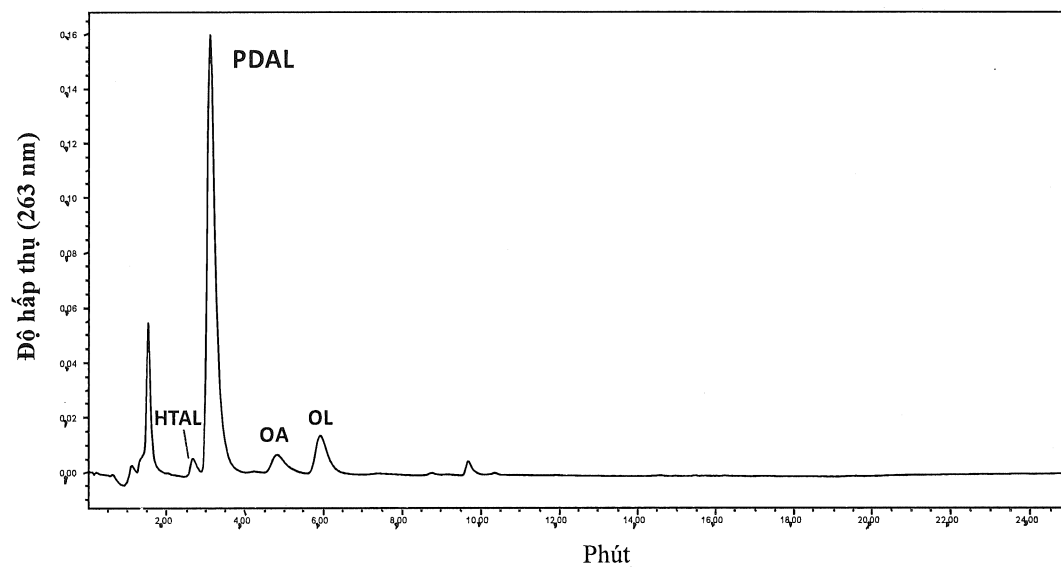


Fig. 4

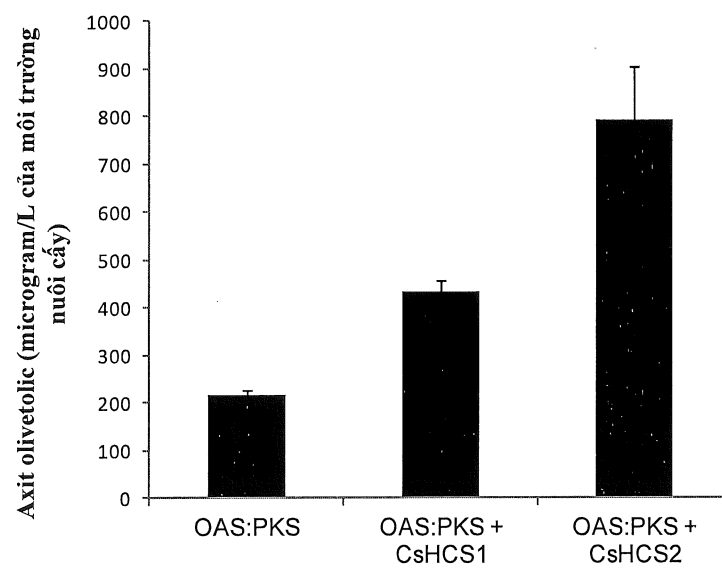
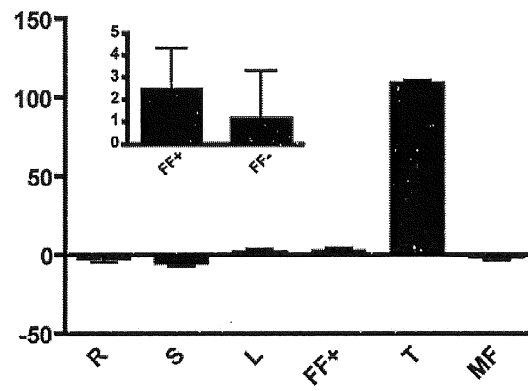


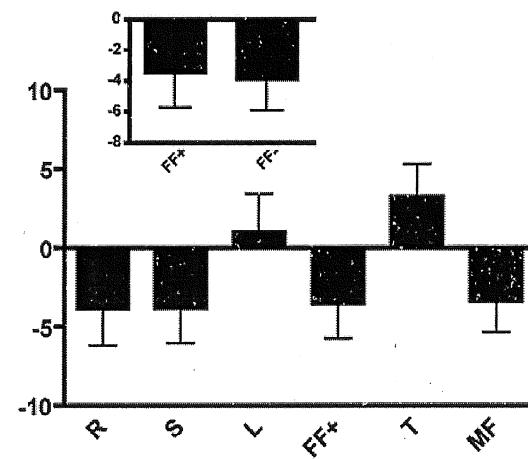
Fig. 5

Độ dời đảo phiên mã (số lần chênh lệch so với lá)

CsHCS1



CsHCS2



CBDA syntaza

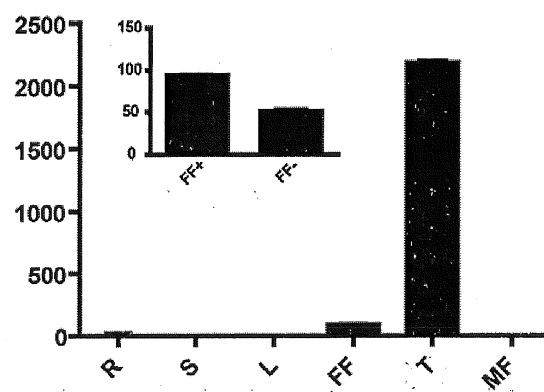


Fig. 6

DANH MỤC TRÌNH TỰ

<110> National Research Council of Canada
University of Saskatchewan
Page, Jonathan E
Stout, Jason M

<120> Phân tử axit nucleic, vectơ chứa phân tử axit nucleic này, quy trình tổng hợp alkanoyl-coA và quy trình biến đổi hàm lượng của hợp chất canabinoit

<130> 12416-98

<150> US 61/507,331

<151> 2011-07-13

<160> 11

<170> PatentIn version 3.5

<210> 1

<211> 2163

<212> ADN

<213> Cannabis sativa

<400> 1
atgggtaaga attacaagtc cctggactct gttgtggcct ctgacttcat agccctaggt
60

atcacctctg aagttgctga gacactccat ggtagactgg ccgagatcgt gtgtaattat
120

ggcgctgccca ctccccaac atggatcaat attgccaacc atattctgtc gcctgacctc
180

cccttctccc tgcaccagat gctcttctat ggttgctata aagacttttg accctgcccct
240

cctgcttgga taccgaccc ggagaaagta aagtccacca atctgggcgc acttttgga
300

aagcgaggaa aagagttttt gggagtcaag tataaggatc ccatttcaag cttttctcat
360

ttccaagaat tttctgtaag aaaccctgag gtgtattgga gaacagtact aatggatgag
420

atgaagataa gtttttcaaa ggatccagaa tgtatattgc gtagagatga tattaataat
480

ccagggggta gtgaatggct tccaggaggt tatcttaact cagcaaagaa ttgcttgaat
540

gtaaatagta acaagaaatt gaatgatata atgattgtat ggcgtgatga aggaaatgat
600

gatttgctc taaacaaatt gacacttgac caattgcgta aacgtgtttg gttagttggt
660

tatgcacttg aagaaatggg tttggagaag ggttggtgcaa ttgcaattga tatgccaatg
720

catgtggatg ctgtgggttat ctatctagct attgttcttg cgggatatgt agttgtttct
780

attgctgata gtttttctgc tcctgaaata tcaacaagac ttcgactatc aaaagcaaaa
840

gccattttta cacaggatca tattattcgt gggaagaagc gtattccctt atacagtaga
900

gttggtggaag ccaagtctcc catggccatt gttattcctt gtagtggctc taatattggt
960

gcagaattgc gtgatggcga tatttcttgg gattactttc tagaaagagc aaaagagttt
1020

aaaaattgtg aatttactgc tagagaacaa ccagttgatg cctatacaaa catcctcttc
1080

tcattctggaa caacagggga gccaaaggca attccatgga ctcaagcaac tcctttaaaa
1140

gcagctgcag atgggtggag ccatttggac attaggaaag gtgatgtcat tgtttggccc
1200

actaatcttg gttggatgat gggtccttgg ctgggtctatg cttcactcct taatggggct
1260

tctattgcct tgtataatgg atcaccactt gtttctggct ttgccaaatt tgtgcaggat
1320

gctaaagtaa caatgctagg tgtggtccct agtattgttc gatcatggaa aagtaccaat
1380

tgtgttagtg gctatgattg gtccaccatc cgttgctttt cctcttcttg tgaagcatct
1440

aatgtagatg aatacctatg gttgatgggg agagcaaact acaagcctgt tatcgaaatg
1500

tgtggtggca cagaaattgg tgggtgcattt tctgctggct ctttcttaca agctcaatca
1560

ttatcttcat ttagttcaca atgtatgggt tgcactttat acatacttga caagaatggt
1620

tatccaatgc ctaaaaacaa accaggaatt ggtgaattag cgcttggtcc agtcatgttt
1680

ggagcatcga agactctggt gaatggtaat caccatgatg tttattttaa gggaatgcct
1740

acattgaatg gagaggtttt aaggaggcat ggggacattt ttgagcttac atctaattggt
1800

tattatcatg cacatgggtcg tgcagatgat acaatgaata ttggaggcat caagattagt
1860

tccatagaga ttgaacgagt ttgtaatgaa gttgatgaca gagttttcga gacaactgct
1920

attggagtgc cacctttggg cggtggacct gagcaattag taattttctt tgtattaaaa
1980

gattcaaagt atacaactat tgacttaaata caattgaggt tatctttcaa cttgggttta
2040

cagaagaaac taaatcctct gttcaaggct actcgtgttg tgcctctttc atcacttccg
2100

agaacagcaa ccaacaagat catgagaagg gttttgcgcc agcaattttc tcactttgaa
2160

tga
2163

<210> 2
<211> 719
<212> PRT
<213> Cannabis sativa

<400> 2

Met	Gly	Lys	Asn	Tyr	Lys	Ser	Leu	Asp	Ser	Val	Val	Ala	Ser	Asp	Phe
1			5						10					15	

Ile	Ala	Leu	Gly	Ile	Thr	Ser	Glu	Val	Ala	Glu	Thr	Leu	His	Gly	Arg
		20						25					30		

Leu	Ala	Glu	Ile	Val	Cys	Asn	Tyr	Gly	Ala	Ala	Thr	Pro	Gln	Thr	Trp
		35					40					45			

Ile	Asn	Ile	Ala	Asn	His	Ile	Leu	Ser	Pro	Asp	Leu	Pro	Phe	Ser	Leu
	50					55					60				

His	Gln	Met	Leu	Phe	Tyr	Gly	Cys	Tyr	Lys	Asp	Phe	Gly	Pro	Ala	Pro
65					70					75				80	

Pro	Ala	Trp	Ile	Pro	Asp	Pro	Glu	Lys	Val	Lys	Ser	Thr	Asn	Leu	Gly
			85						90					95	

Ala	Leu	Leu	Glu	Lys	Arg	Gly	Lys	Glu	Phe	Leu	Gly	Val	Lys	Tyr	Lys
			100					105					110		

Asp Pro Ile Ser Ser Phe Ser His Phe Gln Glu Phe Ser Val Arg Asn

115								120								125
Pro	Glu	Val	Tyr	Trp	Arg	Thr	Val	Leu	Met	Asp	Glu	Met	Lys	Ile	Ser	
130						135					140					
Phe	Ser	Lys	Asp	Pro	Glu	Cys	Ile	Leu	Arg	Arg	Asp	Asp	Ile	Asn	Asn	
145					150					155					160	
Pro	Gly	Gly	Ser	Glu	Trp	Leu	Pro	Gly	Gly	Tyr	Leu	Asn	Ser	Ala	Lys	
				165					170						175	
Asn	Cys	Leu	Asn	Val	Asn	Ser	Asn	Lys	Lys	Leu	Asn	Asp	Thr	Met	Ile	
			180					185					190			
Val	Trp	Arg	Asp	Glu	Gly	Asn	Asp	Asp	Leu	Pro	Leu	Asn	Lys	Leu	Thr	
		195					200					205				
Leu	Asp	Gln	Leu	Arg	Lys	Arg	Val	Trp	Leu	Val	Gly	Tyr	Ala	Leu	Glu	
	210					215					220					
Glu	Met	Gly	Leu	Glu	Lys	Gly	Cys	Ala	Ile	Ala	Ile	Asp	Met	Pro	Met	
225					230					235					240	
His	Val	Asp	Ala	Val	Val	Ile	Tyr	Leu	Ala	Ile	Val	Leu	Ala	Gly	Tyr	
			245					250						255		
Val	Val	Val	Ser	Ile	Ala	Asp	Ser	Phe	Ser	Ala	Pro	Glu	Ile	Ser	Thr	
			260					265					270			
Arg	Leu	Arg	Leu	Ser	Lys	Ala	Lys	Ala	Ile	Phe	Thr	Gln	Asp	His	Ile	
		275					280					285				
Ile	Arg	Gly	Lys	Lys	Arg	Ile	Pro	Leu	Tyr	Ser	Arg	Val	Val	Glu	Ala	
	290					295					300					
Lys	Ser	Pro	Met	Ala	Ile	Val	Ile	Pro	Cys	Ser	Gly	Ser	Asn	Ile	Gly	
305					310					315					320	
Ala	Glu	Leu	Arg	Asp	Gly	Asp	Ile	Ser	Trp	Asp	Tyr	Phe	Leu	Glu	Arg	
				325					330					335		
Ala	Lys	Glu	Phe	Lys	Asn	Cys	Glu	Phe	Thr	Ala	Arg	Glu	Gln	Pro	Val	
			340					345					350			

Asp Ala Tyr Thr Asn Ile Leu Phe Ser Ser Gly Thr Thr Gly Glu Pro
 355 360 365

Lys Ala Ile Pro Trp Thr Gln Ala Thr Pro Leu Lys Ala Ala Ala Asp
 370 375 380

Gly Trp Ser His Leu Asp Ile Arg Lys Gly Asp Val Ile Val Trp Pro
 385 390 395 400

Thr Asn Leu Gly Trp Met Met Gly Pro Trp Leu Val Tyr Ala Ser Leu
 405 410 415

Leu Asn Gly Ala Ser Ile Ala Leu Tyr Asn Gly Ser Pro Leu Val Ser
 420 425 430

Gly Phe Ala Lys Phe Val Gln Asp Ala Lys Val Thr Met Leu Gly Val
 435 440 445

Val Pro Ser Ile Val Arg Ser Trp Lys Ser Thr Asn Cys Val Ser Gly
 450 455 460

Tyr Asp Trp Ser Thr Ile Arg Cys Phe Ser Ser Ser Gly Glu Ala Ser
 465 470 475 480

Asn Val Asp Glu Tyr Leu Trp Leu Met Gly Arg Ala Asn Tyr Lys Pro
 485 490 495

Val Ile Glu Met Cys Gly Gly Thr Glu Ile Gly Gly Ala Phe Ser Ala
 500 505 510

Gly Ser Phe Leu Gln Ala Gln Ser Leu Ser Ser Phe Ser Ser Gln Cys
 515 520 525

Met Gly Cys Thr Leu Tyr Ile Leu Asp Lys Asn Gly Tyr Pro Met Pro
 530 535 540

Lys Asn Lys Pro Gly Ile Gly Glu Leu Ala Leu Gly Pro Val Met Phe
 545 550 555 560

Gly Ala Ser Lys Thr Leu Leu Asn Gly Asn His His Asp Val Tyr Phe
 565 570 575

Lys Gly Met Pro Thr Leu Asn Gly Glu Val Leu Arg Arg His Gly Asp

580 585 590
 Ile Phe Glu Leu Thr Ser Asn Gly Tyr Tyr His Ala His Gly Arg Ala
 595 600 605
 Asp Asp Thr Met Asn Ile Gly Gly Ile Lys Ile Ser Ser Ile Glu Ile
 610 615 620
 Glu Arg Val Cys Asn Glu Val Asp Asp Arg Val Phe Glu Thr Thr Ala
 625 630 635 640
 Ile Gly Val Pro Pro Leu Gly Gly Gly Pro Glu Gln Leu Val Ile Phe
 645 650 655
 Phe Val Leu Lys Asp Ser Asn Asp Thr Thr Ile Asp Leu Asn Gln Leu
 660 665 670
 Arg Leu Ser Phe Asn Leu Gly Leu Gln Lys Lys Leu Asn Pro Leu Phe
 675 680 685
 Lys Val Thr Arg Val Val Pro Leu Ser Ser Leu Pro Arg Thr Ala Thr
 690 695 700
 Asn Lys Ile Met Arg Arg Val Leu Arg Gln Phe Ser His Phe Glu
 705 710 715
 <210> 3
 <211> 1547
 <212> ADN
 <213> Cannabis sativa
 <400> 3
 atggagaaat ctttttcaga aactcatctt catacccaca aaagccagct ctcattgatt
 60
 ccgaaaccaa ccaaatactc tccttttccc acttcaaate tacgggttatc aagggtctccc
 120
 atggcctttct caatctgggt atcaagaaaa acgacgtcgt tctcatctac gccctaatt
 180
 ctatccactt ccctgtttgt ttcttgggaa ttatagcctc tggagccatt gccactacct
 240
 caaatcctct ctacacagtt tccgagcttt ccaaacaggt caaggattcc aatcccaaac
 300
 tcattatcac cgttcctcaa ctcttggaaa aagtaaaggg tttcaatctc cccacgattc
 360

taattggtcc tgattctgaa caagaatctt ctagtgataa agtaatgacc tttaacgatt
420

tgggtcaactt aggtgggtcg tctggctcag aatttccaat tgttgatgat tttaaagcaga
480

gtgacactgc tgcgctattg tactcatctg gcacaacggg aatgagtaaa ggtgtggttt
540

tgactcaciaa aaacttcatt gcctcttctt taatggtgac aatggagcaa gacctagttg
600

gagagatgga taatgtgttt ctatgctttt tgccaatggt tcatgtattt ggtttggcta
660

tcatcaccta tgctcagttg cagagaggaa aactgttat ttcaatggcg agatttgacc
720

ttgagaagat gttaaaagat gtggaaaagt ataaagttac ccatttgtgg gttgtgcctc
780

ctgtgatact ggctctgagt aagaacagta tgggtgaagaa gtttaatctt tcttctataa
840

agtatatattg ctccggtgca gctcctttgg gcaaagattt aatggaggag tgctctaagg
900

ttgttcctta tgggtattgtt gctcagggat atgggtatgac agaaacttgt gggattgtat
960

ccatggagga tataagagga ggtaaacgaa atagtgggtc agctggaatg ctggcatctg
1020

gagtagaagc ccagatagtt agtgtagata cactgaagcc cttacctctt aatcaattgg
1080

gggagatatg ggtgaagggg cctaatatga tgcaagggtta cttcaataac ccacaggcaa
1140

ccaagttgac tatagataag aaagggttggg tacatactgg tgatcttgga tattttgatg
1200

aagatggaca tctttatggt gttgaccgta taaaagagct catcaaatat aaaggatttc
1260

aggttgctcc tgctgagctt gaaggattgc ttgtttctca ccctgaaata ctgatgctg
1320

ttgtgattcc atttctgac gctgaagcgg gtgaagtccc agttgcttat gttgtgcgct
1380

ctcccaacag ttcattaacc gaaaatgatg tgaagaaatt tatcgcgggc caggttgcat
1440

ctttcaaaag attgagaaaa gtaacattta taaacagtgt cccgaaatct gcttcgggga
1500

aaatcctcag aagagaactc attcagaaag tacgctccaa catgtga
1547

<210> 4
<211> 543
<212> PRT
<213> Cannabis sativa

<400> 4

Met Glu Lys Ser Gly Tyr Gly Arg Asp Gly Ile Tyr Arg Ser Leu Arg
1 5 10 15

Pro Pro Leu His Leu Pro Asn Asn Asn Asn Leu Ser Met Val Ser Phe
20 25 30

Leu Phe Arg Asn Ser Ser Ser Tyr Pro Gln Lys Pro Ala Leu Ile Asp
35 40 45

Ser Glu Thr Asn Gln Ile Leu Ser Phe Ser His Phe Lys Ser Thr Val
50 55 60

Ile Lys Val Ser His Gly Phe Leu Asn Leu Gly Ile Lys Lys Asn Asp
65 70 75 80

Val Val Leu Ile Tyr Ala Pro Asn Ser Ile His Phe Pro Val Cys Phe
85 90 95

Leu Gly Ile Ile Ala Ser Gly Ala Ile Ala Thr Thr Ser Asn Pro Leu
100 105 110

Tyr Thr Val Ser Glu Leu Ser Lys Gln Val Lys Asp Ser Asn Pro Lys
115 120 125

Leu Ile Ile Thr Val Pro Gln Leu Leu Glu Lys Val Lys Gly Phe Asn
130 135 140

Leu Pro Thr Ile Leu Ile Gly Pro Asp Ser Glu Gln Glu Ser Ser Ser
145 150 155 160

Asp Lys Val Met Thr Phe Asn Asp Leu Val Asn Leu Gly Gly Ser Ser
165 170 175

Gly Ser Glu Phe Pro Ile Val Asp Asp Phe Lys Gln Ser Asp Thr Ala
180 185 190

Ala Leu Leu Tyr Ser Ser Gly Thr Thr Gly Met Ser Lys Gly Val Val
 195 200 205

Leu Thr His Lys Asn Phe Ile Ala Ser Ser Leu Met Val Thr Met Glu
 210 215 220

Gln Asp Leu Val Gly Glu Met Asp Asn Val Phe Leu Cys Phe Leu Pro
 225 230 235 240

Met Phe His Val Phe Gly Leu Ala Ile Ile Thr Tyr Ala Gln Leu Gln
 245 250 255

Arg Gly Asn Thr Val Ile Ser Met Ala Arg Phe Asp Leu Glu Lys Met
 260 265 270

Leu Lys Asp Val Glu Lys Tyr Lys Val Thr His Leu Trp Val Val Pro
 275 280 285

Pro Val Ile Leu Ala Leu Ser Lys Asn Ser Met Val Lys Lys Phe Asn
 290 295 300

Leu Ser Ser Ile Lys Tyr Ile Gly Ser Gly Ala Ala Pro Leu Gly Lys
 305 310 315 320

Asp Leu Met Glu Glu Cys Ser Lys Val Val Pro Tyr Gly Ile Val Ala
 325 330 335

Gln Gly Tyr Gly Met Thr Glu Thr Cys Gly Ile Val Ser Met Glu Asp
 340 345 350

Ile Arg Gly Gly Lys Arg Asn Ser Gly Ser Ala Gly Met Leu Ala Ser
 355 360 365

Gly Val Glu Ala Gln Ile Val Ser Val Asp Thr Leu Lys Pro Leu Pro
 370 375 380

Pro Asn Gln Leu Gly Glu Ile Trp Val Lys Gly Pro Asn Met Met Gln
 385 390 395 400

Gly Tyr Phe Asn Asn Pro Gln Ala Thr Lys Leu Thr Ile Asp Lys Lys
 405 410 415

Gly Trp Val His Thr Gly Asp Leu Gly Tyr Phe Asp Glu Asp Gly His

420							425					430				
Leu	Tyr	Val	Val	Asp	Arg	Ile	Lys	Glu	Leu	Ile	Lys	Tyr	Lys	Gly	Phe	
435							440					445				
Gln	Val	Ala	Pro	Ala	Glu	Leu	Glu	Gly	Leu	Leu	Val	Ser	His	Pro	Glu	
450							455					460				
Ile	Leu	Asp	Ala	Val	Val	Ile	Pro	Phe	Pro	Asp	Ala	Glu	Ala	Gly	Glu	
465							470					475				
Val	Pro	Val	Ala	Tyr	Val	Val	Arg	Ser	Pro	Asn	Ser	Ser	Leu	Thr	Glu	
485							490					495				
Asn	Asp	Val	Lys	Lys	Phe	Ile	Ala	Gly	Gln	Val	Ala	Ser	Phe	Lys	Arg	
500							505					510				
Leu	Arg	Lys	Val	Thr	Phe	Ile	Asn	Ser	Val	Pro	Lys	Ser	Ala	Ser	Gly	
515							520					525				
Lys	Ile	Leu	Arg	Arg	Glu	Leu	Ile	Gln	Lys	Val	Arg	Ser	Asn	Met		
530							535					540				

```
<210>      5
<211>      23
<212>      ADN
<213>      Trình tự nhân tạo

<220>
<223>      CsHCS1 đoạn mỗi xuôi

<400>      5
atgggtaaga attacaagtc cct
23
```

```
<210>      6
<211>      30
<212>      ADN
<213>      Trình tự nhân tạo

<220>
<223>      CSHCS1 đoạn mỗi ngược

<400>      6
gagctctcat tcaaagtgag aaaattgctg
30
```

<210> 7

<211> 23
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> CsHCS2 đoạn mỗi xuôi

<400> 7
 atggagaaat ctgggtatgg aag
 23

<210> 8
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> CsHCS2 đoạn mỗi ngược

<400> 8
 tcacatgttg gagcgtactt tc
 22

<210> 9
 <211> 22
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> MCS đoạn mỗi xuôi

<400> 9
 atgagcaacc atcttttcga cg
 22

<210> 10
 <211> 24
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> MCS đoạn mỗi ngược

<400> 10
 ttacgtcctg gtataaagat cggc
 24

<210> 11
 <211> 33
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> Yếu tố liên kết tổng hợp

<400> 11

Ala	Ala	Thr	Ser	Gly	Ser	Thr	Gly	Ser	Thr	Gly	Ser	Thr	Gly	Ser	Gly
1				5				10					15		

Arg	Ser	Thr	Gly	Ser	Thr	Gly	Ser	Thr	Gly	Ser	Gly	Arg	Ser	His	Met
			20				25						30		

Val