



(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) 

1-0021745

(51)⁷ **C12P 7/04, 7/00, 5/00, C12N 15/52, C12P 5/02, 7/02**

(13) **B**

(21) 1-2011-02981

(22) 25.05.2007

(62) 1-2008-02948

(86) PCT/US2007/069807 25.05.2007

(87) WO2007/140339 06.12.2007

(30) 60/808,989 26.05.2006 US

60/870,592 18.12.2006 US

(45) 25.09.2019 378

(43) 27.08.2012 293

(73) AMYRIS BIOTECHNOLOGIES, INC. (US)

5980 Horton Street, Suite 450, Emeryville, CA 94608, United States of America

(72) RENNINGER, Neil, Stephen (US), NEWMAN, Jack (US), REILING, Keith, Kinkead (US), REGENTIN, Rika (US), PADDON, Christopher, John (US)

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

(54) **PHƯƠNG PHÁP SẢN XUẤT ISOPRENOIT**

(57) Sáng chế đề xuất phương pháp sản xuất ổn định isoprenoit thông qua một hoặc nhiều quá trình sinh tổng hợp. Sáng chế còn đề xuất axit nucleic, enzym, vectơ biểu hiện, và tế bào chủ đã được cải biến di truyền để thực hiện phương pháp này. Sáng chế còn đề xuất phương pháp lên men cho sản lượng isoprenoit cao từ tế bào chủ đã được cải biến di truyền này.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất ổn định isoprenoit thông qua một hoặc nhiều quá trình sinh tổng hợp. Sáng chế còn đề cập đến phương pháp lên men cho sản lượng isoprenoit cao từ tế bào chủ đã được cải biến di truyền.

Tình trạng kỹ thuật của sáng chế

Isoprenoit có ở khắp nơi trong tự nhiên. Chúng bao gồm một họ đa dạng gồm hơn 40.000 sản phẩm riêng rẽ, nhiều loại trong số này là rất quan trọng đối với các sinh vật sống. Isoprenoit giúp duy trì tính lưu động của tế bào, chức năng vận chuyển điện tử, và các chức năng chuyển hoá khác. Một lượng lớn các isoprenoit tự nhiên và tổng hợp là hữu ích làm thuốc, mỹ phẩm, nước hoa, chất nhuộm và chất tạo màu, thuốc diệt nấm, thuốc chống nhiễm khuẩn, chất dinh dưỡng, và sản phẩm trung gian hoá học mịn.

Sản phẩm isoprenoit thường chứa năm đơn vị cacbon isopentenyl diphosphat (isopentenyl diphosphate - IPP) lặp lại, mặc dù các isoprenoit và polyterpen bất quy tắc đã được báo cáo. Trong tự nhiên, isoprenoit được tổng hợp bằng cách ngưng tụ liên tục tiền chất IPP của chúng và dạng đồng phân dimethylallyl pyrophosphat (dimethylallyl pyrophosphate - DMAPP) của nó. Hai quá trình đối với các tiền chất này là đã biết. Các sinh vật nhân chuẩn, trừ thực vật, thường sử dụng quá trình phụ thuộc mevalonat (mevalonate-dependent - MEV) để chuyển hoá axetyl coenzym A (axetyl-CoA) thành IPP, chất này sau đó được đồng phân hoá thành DMAPP. Sinh vật nhân sơ, trừ một số ngoại lệ, thường chỉ sử dụng quá trình không phụ thuộc mevalonat hoặc quá trình deoxyxyluloza-5-phosphat (deoxyxylulose-5-phosphate - DXP) để sản xuất IPP và DMAPP. Thực vật sử dụng cả quá trình MEV và quá trình DXP. Xem ấn phẩm: Rohmer et al. (1993) *Biochem. J.* 295:517-524; Lange et al. (2000) *Proc. Natl. Acad. Sci. USA* 97(24):13172-13177; Rohdich et al. (2002) *Proc. Natl. Acad. Sci. USA* 99:1158-1163.

Theo truyền thống, isoprenoit đã được sản xuất bằng cách chiết từ các nguồn tự nhiên như thực vật, vi sinh vật, và động vật. Tuy nhiên, sản lượng thu được bằng cách chiết này thường rất thấp do có nhiều hạn chế đáng kể. Thứ nhất, phần lớn isoprenoit tích lũy trong tự nhiên chỉ với lượng nhỏ. Thứ hai, nói chung các nguồn sinh vật này không phù hợp để nuôi cấy ở quy mô lớn cần thiết để sản xuất lượng isoprenoit mong muốn khả thi về mặt thương mại. Thứ ba, yêu cầu về một số dung môi độc để chiết isoprenoit, đòi hỏi phải có các quy trình xử lý và tiêu hủy đặc biệt, và do đó làm phức tạp quy trình sản xuất isoprenoit ở quy mô thương mại.

Các quá trình chuyển hoá MEV và DXP được làm sáng tỏ khiến cho việc sản xuất sinh tổng hợp isoprenoit có thể thực hiện được. Ví dụ, vi sinh vật đã được thiết kế di truyền để biểu hiện quá mức một phần hoặc toàn bộ quá trình mevalonat để sản xuất isoprenoit có tên là amorpha-4,11-dien (Patent Mỹ số 7,172,886 và 7,192,751). Các nghiên cứu khác tập trung vào việc làm cân bằng nguồn glyxeraldehyt-3-phosphat và pyruvat, hoặc tập trung vào việc làm tăng mức biểu hiện của 1-deoxy-D-xyluloza-5-phosphat synthaza (1-deoxy-D-xylulose-5-phosphate synthase - dxs) và IPP isomeraza (IPP isomerase - idi). Xem ấn phẩm: Farmer et al. (2001) *Biotechnol. Prog.* **17**:57-61; Kajiwara et al. (1997) *Biochem. J.* **324**:421-426; và Kim et al. (2001) *Biotechnol. Bioeng.* **72**:408-415.

Tuy nhiên, do nhu cầu cần có lượng sản phẩm isoprenoit rất lớn cho nhiều ứng dụng thương mại, vẫn cần có các hệ biểu hiện và các quy trình lên men mà sản xuất được nhiều isoprenoit hơn so với các công nghệ hiện tại. Việc định hướng lại tối ưu sự chuyển hoá vi sinh vật theo hướng sản xuất isoprenoit đòi hỏi là quá trình sinh tổng hợp đã đưa vào được cải biến di truyền thích hợp để tập trung cacbon vào việc vừa sản xuất isoprenoit một cách hiệu quả và vừa ngăn ngừa sự tích lũy các mức gây độc của các sản phẩm chuyển hóa trung gian trong khoảng thời gian kéo dài. Sáng chế giải quyết nhu cầu này và còn có các ưu điểm.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất chế phẩm và phương pháp sản xuất mạnh isoprenoit bằng cách sử dụng các enzym của quá trình isopentenyl pyrophosphat được kiểm soát bởi ít nhất một yếu tố điều hòa khác loại hoặc các điều kiện lên men, hoặc riêng rẽ hoặc kết hợp. Ví dụ không giới hạn về các isoprenoit thích hợp bao gồm: hemiterpen (được tạo ra từ 1 đơn vị isopren) như isopren; monoterpen (được tạo ra từ 2 đơn vị isopren) như myrcen; sesquiterpen (được tạo ra từ 3 đơn vị isopren) như amorph-4,11-dien; diterpen (được tạo ra từ bốn đơn vị isopren) như taxadien; triterpen (được tạo ra từ 6 đơn vị isopren) như squalen; tetraterpen (được tạo ra từ 8 isoprenoit) như β -caroten; và polyterpen (được tạo ra từ nhiều hơn 8 đơn vị isopren) như polyisopren.

Theo một khía cạnh, phương pháp sản xuất isoprenoit bao gồm các bước (a) thu nhiều tế bào chủ chứa quá trình enzym để tạo ra isopentenyl pyrophosphat trong đó tất cả các enzym của quá trình này đều được kiểm soát bởi ít nhất một yếu tố điều hòa phiên mã khác loại; và (b) nuôi cấy tế bào chủ này trong môi trường ở các điều kiện dưới điểm cực thuận so với các điều kiện mà sẽ tạo ra tốc độ sinh trưởng riêng tối đa đối với tế bào chủ này. Theo một số phương án, quá trình này là quá trình mevalonat. Theo phương án khác, quá trình này là quá trình DXP. Theo các phương án khác, ít nhất một trình tự điều hòa phiên mã khác loại có thể cảm ứng được. Theo các phương án khác, các enzym của quá trình này được kiểm soát bởi một yếu tố điều hòa phiên mã duy nhất. Theo phương án khác, các enzym của quá trình này được kiểm soát bởi nhiều yếu tố điều hòa phiên mã khác loại.

Theo một số phương án, quá trình này bao gồm trình tự axit nucleic mã hóa enzym của quá trình mevalonat từ sinh vật nhân sơ có quá trình mevalonat nội sinh. Ví dụ về sinh vật nhân sơ có quá trình mevalonat nội sinh bao gồm, nhưng không bị giới hạn ở chi *Enterococcus*, chi *Pseudomonas*, và chi *Staphylococcus*. Theo một phương án, enzym của quá trình mevalonat này được chọn từ axetyl-CoA thiolaza, HMG-CoA synthaza, HMG-CoA reductaza,

và mevalonat kinaza. Theo phương án khác, trình tự axit nucleic khác loại mã hóa HMG-CoA reductaza lớp II.

Theo phương án khác, tế bào chủ được nuôi cấy trong môi trường trong đó mức dinh dưỡng và/hoặc nhiệt độ được duy trì ở mức thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ này. Theo phương án khác, tế bào chủ được nuôi cấy trong môi trường trong đó nguồn cacbon được duy trì ở mức để tạo ra thấp hơn khoảng 90%, 75%, 50%, 25%, 10%, hoặc ở tỷ lệ bất kỳ nằm trong khoảng từ 90% đến 10% tốc độ sinh trưởng riêng tối đa. Theo phương án khác, tế bào chủ được nuôi cấy trong môi trường trong đó nguồn nitơ được duy trì ở mức tạo ra thấp hơn khoảng 90%, 75%, 50%, 25%, 10%, hoặc ở tỷ lệ bất kỳ nằm trong khoảng từ 90% đến 10% tốc độ sinh trưởng riêng tối đa. Theo phương án khác, tế bào chủ được nuôi cấy trong môi trường trong đó nhiệt độ được duy trì ở mức để tạo ra thấp hơn khoảng 90%, 75%, 50%, 25%, 10% hoặc ở tỷ lệ bất kỳ nằm trong khoảng từ 90% đến 10% tốc độ sinh trưởng riêng tối đa. Theo phương án khác, nhiệt độ môi trường được duy trì thấp hơn ít nhất là khoảng 2°C, 4°C, 5°C, 6°C, 8°C, 10°C, 15°C, hoặc 20°C so với nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa.

Theo phương án khác nữa, phương pháp sản xuất isoprenoit hoặc tiền chất isoprenoit bao gồm các bước (i) tiến hành phản ứng lên men bao gồm môi trường lên men và nhiều tế bào chủ đã được cải biến di truyền mà sản xuất isoprenoit trong điều kiện sao cho (a) môi trường lên men được giữ ở nhiệt độ thấp hơn nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ này; (b) môi trường lên men chứa nguồn cacbon có mặt với lượng thấp hơn lượng sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ; và/hoặc (c) môi trường lên men chứa nguồn nitơ có mặt với lượng thấp hơn lượng sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ; (ii) thu hồi isoprenoit được tạo ra trong một hoặc nhiều điều kiện nêu trong mục từ (a) đến (c). Theo một khía cạnh, isoprenoit được sản xuất ra trong ít nhất hai điều kiện nêu trong mục từ (a) đến (c). Theo khía cạnh khác, isoprenoit được sản xuất ra trong tất cả các điều kiện nêu trong mục từ (a) đến (c).

Mô tả vắn tắt các hình vẽ

Fig.1A là sơ đồ biểu diễn quá trình mevalonat ("MEV") để sản xuất isopentenyl pyrophosphat ("IPP").

Fig.1B là sơ đồ biểu diễn quá trình 1-deoxy-D-xyluloza 5-diphosphat ("DXP") để sản xuất isopentenyl pyrophosphat ("IPP") và dimethylalyl pyrophosphat ("DMAPP"). Dxs là 1-deoxy-D-xyluloza-5-phosphat synthaza; Dxr là 1-deoxy-D-xyluloza-5-phosphat reductoisomeraza (còn được gọi là IspC); IspD là 4-diphosphoxytidyl-2C-metyl-D-erythritol synthaza; IspE là 4-diphosphoxytidyl-2C-metyl-D-erythritol synthaza; IspF là 2C-metyl-D-erythritol 2,4-xyclodiphosphat synthaza; IspG là 1-hydroxy-2-metyl-2-(E)-butenyl 4-diphosphat synthaza (IspG); và ispH là isopentenyl/dimethylalyl diphosphat synthaza.

Fig.2 là sơ đồ biểu diễn quá trình chuyển hoá isopentenyl pyrophosphat ("IPP") và dimethylalyl pyrophosphat ("DMAPP") thành geranyl pyrophosphat ("GPP"), farnesyl pyrophosphat ("FPP"), và geranylgeranyl pyrophosphat ("GGPP"), và quá trình tổng hợp các isoprenoit khác nhau.

Fig.3 thể hiện bản đồ plasmit biểu hiện pMBIS-gpps.

Fig.4 thể hiện bản đồ plasmit biểu hiện pAM408.

Fig.5 thể hiện bản đồ plasmit biểu hiện pAM424.

Fig.6 thể hiện bản đồ plasmit biểu hiện pTrc99A-ADS, pTrc99A-FSA, pTrc99A-LLS, pTrc99A-LMS, pTrc99A-GTS, pTrc99A-APS, pTrc99A-BPS, pTrc99A-PHS, pTrc99A-TS, pTrc99A-CS, pTrc99A-SS, và pAM373.

Fig.7A-C là sơ đồ cấu trúc của plasmit pAM489-pAM498 và pAM328.

Fig.8 thể hiện hoạt tính đặc hiệu cao hơn và độ ổn định gia tăng của HMGR-CoA reductaza (HMGR) ở *Enterococcus faecalis* so với HMG-CoA reductaza bị cắt cụt (tHMGR) ở *Saccharomyces cerevisiae*.

Fig.9 thể hiện mối tương quan giữa khối lượng tế bào khô (dry cell weight - "DCW") trên lít và OD₆₀₀.

Fig.10A-B thể hiện năng suất sản xuất amorph-4,11-dien riêng và theo thể tích gia tăng của các chủng chủ mang các gen HMGR và HMGS ở *Staphylococcus aureus* so với các chủng chủ mang gen tHMGR và HMGS ở *Saccharomyces cerevisiae*.

Fig.11A-B thể hiện ảnh hưởng của nhiệt độ thấp hơn đến năng suất sản xuất amorph-4,11-dien của chủng chủ *Escherichia coli*.

Fig.12A-D thể hiện ảnh hưởng của mức glucoza giảm đến năng suất sản xuất amorph-4,11-dien của chủng chủ *Escherichia coli*.

Fig.13A-B thể hiện ảnh hưởng kết hợp của nhiệt độ thấp hơn và mức glucoza giảm đến năng suất sản xuất amorph-4,11-dien của chủng chủ *Escherichia coli*.

Fig.14A-E và Fig.15A-E thể hiện ảnh hưởng kết hợp của nhiệt độ thấp hơn và mức glucoza và nitơ giảm đến năng suất sản xuất amorph-4,11-dien của chủng chủ *Escherichia coli*.

Fig.16 thể hiện quá trình sản xuất amorph-4,11-dien qua quá trình DXP bởi chủng chủ *Escherichia coli*.

Mô tả chi tiết sáng chế

Các định nghĩa

Trừ khi có quy định khác, tất cả các thuật ngữ khoa học và kỹ thuật được dùng ở đây có cùng nghĩa như thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Các thuật ngữ được dùng ở đây sẽ được định nghĩa như sau:

Thuật ngữ "tuỳ chọn" hoặc "tuỳ ý" được dùng để chỉ đặc điểm hoặc cấu trúc đã được mô tả có thể có mặt hoặc có thể không có mặt, hoặc sự kiện hoặc tình huống đã được mô tả có thể xuất hiện hoặc có thể không xuất hiện, và sự mô tả này bao gồm cả các trường hợp trong đó đặc điểm hoặc cấu trúc cụ thể có mặt và các trường hợp trong đó đặc điểm hoặc cấu trúc không có mặt, hoặc các trường hợp trong đó sự kiện hoặc tình huống xuất hiện và các ví dụ trong đó sự kiện hoặc tình huống không xuất hiện.

Thuật ngữ "quá trình chuyển hoá" được dùng ở đây để chỉ quá trình dị hoá hoặc quá trình đồng hoá. Quá trình đồng hoá liên quan đến quá trình tạo cấu trúc phân tử lớn hơn từ các phân tử nhỏ hơn, là quá trình đòi hỏi năng lượng. Quá trình dị hoá liên quan đến việc phá vỡ các phân tử lớn hơn, thường giải phóng năng lượng.

Thuật ngữ "quá trình mevalonat" hoặc "quá trình MEV" được dùng ở đây để chỉ quá trình sinh tổng hợp chuyển hoá axetyl-CoA thành IPP. Quá trình MEV được minh họa dưới dạng sơ đồ trên Fig.1A.

Thuật ngữ "quá trình deoxyxyluloza 5-phosphat" hoặc "quá trình DXP" được dùng ở đây để chỉ quá trình chuyển hoá glyxeraldehyt-3-phosphat và pyruvat thành IPP và DMAPP. Quá trình DXP được minh họa dưới dạng sơ đồ trên Fig.1B.

Thuật ngữ "pyrophosphat" có thể được dùng thay thế lẫn nhau với "diphosphat" trong bản mô tả này.

Thuật ngữ "vector biểu hiện" hoặc "vector" được dùng để chỉ axit nucleic mà được tải nạp, biến nạp, hoặc gây nhiễm tế bào chủ, nhờ đó làm cho tế bào này sản xuất ra các axit nucleic và/hoặc protein ngoài các axit nucleic và/hoặc protein tự nhiên đối với tế bào này, hoặc làm cho tế bào này biểu hiện các axit nucleic và/hoặc protein theo cách không tự nhiên đối với tế bào này.

Thuật ngữ "nội sinh" được dùng để chỉ chất hoặc quy trình xảy ra một cách tự nhiên, ví dụ, trong tế bào chủ không tái tổ hợp.

Thuật ngữ "quá trình enzym để sản xuất isopentenyl pyrophosphat" được dùng để chỉ quá trình bất kỳ có khả năng sản xuất isopentyl pyrophosphat, bao gồm, nhưng không bị giới hạn ở, quá trình mevalonat hoặc quá trình DXP.

Thuật ngữ "axit nucleic" được dùng để chỉ dạng polyme của các nucleotit có chiều dài bất kỳ, hoặc là ribonucleotit hoặc là deoxynucleotit. Do đó, thuật ngữ này bao gồm, nhưng không bị giới hạn ở, ADN hoặc ARN sợi đơn, sợi kép, hoặc đa sợi, ADN của hệ gen, cADN, thể lai ADN-ARN, hoặc

polyme chứa các bazơ purin và pyrimidin hoặc các bazơ nucleotit tự nhiên, bazơ nucleotit đã được cải biến về mặt hoá học, hoặc hoá sinh, các bazơ nucleotit không có trong tự nhiên, hoặc đã được tạo dẫn xuất.

Thuật ngữ "operon" được dùng để chỉ hai hoặc nhiều trình tự nucleotit liên kế, mỗi trình tự mã hóa sản phẩm gen như ARN hoặc protein, và sự biểu hiện của chúng được điều hòa một cách đồng thời bởi một hoặc nhiều yếu tố kiểm soát (ví dụ, yếu tố khởi đầu).

Thuật ngữ "sản phẩm gen" được dùng để chỉ ARN được mã hoá bởi ADN (hoặc ngược lại) hoặc protein được mã hoá bởi ARN hoặc ADN, trong đó gen thường sẽ chứa một hoặc nhiều trình tự nucleotit mã hóa protein, và cũng có thể chứa các intron và các trình tự nucleotit không mã hóa khác.

Thuật ngữ "protein" được dùng để chỉ dạng polyme của axit amin có chiều dài bất kỳ, mà có thể chứa các axit amin được mã hóa và không được mã hóa, các axit amin đã được cải biến về mặt hoá học hoặc hoá sinh hoặc các axit amin đã được tạo dẫn xuất, và các polypeptit có khung peptit đã được cải biến.

Thuật ngữ "axit nucleic khác loại" ở đây được dùng để chỉ axit nucleic trong đó ít nhất một trong số các điều kiện sau được đáp ứng: (a) axit nucleic là ngoại lai ("ngoại sinh") đối với (tức là, không tìm thấy trong tự nhiên ở) tế bào chủ đã cho; (b) axit nucleic chứa các trình tự nucleotit mà tìm thấy trong tự nhiên ở (tức là, là "nội sinh với") tế bào chủ đã cho, nhưng trình tự nucleotit này được sản xuất ra với lượng không bình thường (ví dụ, lớn hơn lượng mong đợi hoặc lớn hơn lượng được tìm thấy trong tự nhiên) ở tế bào này; (c) axit nucleic chứa các trình tự nucleotit khác biệt về trình tự với trình tự nucleotit nội sinh, nhưng trình tự nucleotit này mã hóa cùng một protein (có trình tự axit amin giống nhau hoặc gần như giống nhau) và được sản xuất ra với lượng không bình thường (ví dụ, lớn hơn lượng mong đợi hoặc lớn hơn lượng tìm thấy trong tự nhiên) ở tế bào này; hoặc (d) axit nucleic chứa hai hoặc nhiều trình tự nucleotit không được tìm thấy trong cùng mối tương quan với nhau trong tự nhiên (ví dụ, axit nucleic là tái tổ hợp).

Thuật ngữ "gen chuyển" được dùng để chỉ gen được đưa theo cách ngoại sinh vào tế bào chủ. Nó có thể chứa axit nucleic nội sinh hoặc ngoại sinh, hoặc axit nucleic khác loại.

Thuật ngữ "vật chủ tái tổ hợp" (còn được gọi là "tế bào chủ đã được cải biến di truyền" hoặc "vi sinh vật chủ đã được cải biến di truyền") được dùng để chỉ tế bào chủ chứa axit nucleic khác loại theo sáng chế.

Thuật ngữ "axit nucleic ngoại sinh" được dùng để chỉ axit nucleic được đưa theo cách ngoại sinh vào tế bào chủ, và do vậy không được tìm thấy một cách thông thường hoặc trong tự nhiên trong và/hoặc được sản xuất bởi tế bào đã cho trong tự nhiên.

Thuật ngữ "yếu tố điều hòa" được dùng để chỉ các trình tự kiểm soát quá trình phiên mã và dịch mã, như yếu tố khởi đầu, yếu tố tăng cường, tín hiệu polyadenyl hoá, yếu tố kết thúc, tín hiệu phân huỷ protein, và các dạng tương tự, mà tạo ra và/hoặc điều hòa sự biểu hiện của trình tự mã hóa và/hoặc việc sản xuất polypeptit được mã hoá ở tế bào chủ.

Thuật ngữ "biến nạp" được dùng để chỉ sự thay đổi di truyền vĩnh viễn hoặc tạm thời được gây ra ở tế bào sau khi đưa axit nucleic mới vào. Sự thay đổi di truyền ("sự cải biến") có thể được thực hiện bằng cách đưa ADN mới vào hệ gen của tế bào chủ, hoặc bằng cách duy trì tạm thời hoặc ổn định ADN mới để làm yếu tố episom. Ở tế bào nhân chuẩn, sự thay đổi di truyền vĩnh viễn thường đạt được bằng cách đưa ADN vào hệ gen của tế bào. Ở tế bào nhân sơ, sự thay đổi di truyền vĩnh viễn có thể được đưa vào trong nhiễm sắc thể hoặc qua các yếu tố ngoài nhiễm sắc thể như các plasmit và vectơ biểu hiện, mà có thể chứa một hoặc nhiều gen chỉ thị có khả năng chọn lọc để trợ giúp việc duy trì chúng ở tế bào chủ tái tổ hợp.

Thuật ngữ "được liên kết chức năng" được dùng để chỉ sự sắp xếp kề nhau trong đó các thành phần đã được mô tả có mối quan hệ cho phép chúng thực hiện chức năng theo cách mong muốn. Ví dụ, yếu tố khởi đầu được liên kết chức năng với trình tự nucleotit nếu yếu tố khởi đầu này tác động đến quá trình phiên mã hoặc biểu hiện của trình tự nucleotit này.

Thuật ngữ “tế bào chủ” và “vi sinh vật chủ” được dùng thay thế lẫn nhau ở đây để chỉ tế bào sống cổ khuẩn, tế bào sống vi khuẩn, hoặc tế bào sống nhân chuẩn mà axit nucleic khác loại có thể được đưa vào hoặc được chèn vào đó. Thuật ngữ này còn đề cập đến thế hệ con của tế bào ban đầu, mà có thể không nhất thiết phải hoàn toàn giống về hình thái học hoặc về hệ gen hoặc toàn bộ ADN bổ sung so với tế bào mẹ ban đầu, do sự đột biến tự nhiên, đột biến ngẫu nhiên, hoặc đột biến có chủ ý.

Thuật ngữ "tổng hợp" được dùng liên quan tới axit nucleic có nghĩa là sự gắn nối của các khối cấu trúc oligonucleotit được tổng hợp hoá học để tạo ra các phân đoạn gen, sau đó, các phân đoạn gen này được lắp ráp nhờ enzym để tạo ra toàn bộ gen. Quá trình tổng hợp axit nucleic bằng "phương pháp hoá học" có nghĩa là các nucleotit thành phần được lắp ráp *in vitro*.

Thuật ngữ "tự nhiên" như được áp dụng đối với axit nucleic, tế bào, hoặc sinh vật, được dùng để chỉ axit nucleic, tế bào, hoặc sinh vật có trong tự nhiên. Ví dụ, trình tự polypeptit hoặc polynucleotit có mặt ở sinh vật không gây bệnh (không bị bệnh) có thể phân lập được từ nguồn tự nhiên và chưa được con người cải biến có chủ ý trong phòng thí nghiệm là tự nhiên.

Thuật ngữ "có trong tự nhiên" khi được áp dụng đối với axit nucleic, enzym, tế bào, hoặc sinh vật, được dùng để chỉ axit nucleic, enzym, tế bào, hoặc sinh vật được tìm thấy trong tự nhiên. Ví dụ, trình tự polypeptit hoặc polynucleotit có mặt trong sinh vật có thể được phân lập từ nguồn tự nhiên và chưa được con người cải biến có chủ ý trong phòng thí nghiệm là có trong tự nhiên.

Thuật ngữ "đoạn có hoạt tính sinh học" như được áp dụng đối với protein, polypeptit hoặc enzym được dùng để chỉ (các) phần chức năng của protein hoặc polypeptit hoặc enzym này. Các dạng tương đương về mặt chức năng có thể có trình tự axit amin biến đổi có thể xuất hiện, ví dụ, là kết quả của sự dư thừa codon và sự tương đương về mặt chức năng được biết là xuất hiện tự nhiên trong trình tự axit nucleic và do đó các protein được mã hoá. Protein hoặc peptit tương đương về chức năng có thể được tạo ra theo cách khác bằng

cách áp dụng công nghệ ADN tái tổ hợp, trong đó các thay đổi về cấu trúc protein có thể được thiết kế di truyền, dựa trên sự xem xét về đặc tính của axit amin đang được trao đổi.

Thuật ngữ "isoprenoit", "hợp chất isoprenoit", "sản phẩm isoprenoit", "terpen", "hợp chất terpen", "terpenoit", và "hợp chất terpenoit" được dùng thay thế lẫn nhau ở đây. Các thuật ngữ này được dùng để chỉ hợp chất có khả năng được tạo ra từ IPP.

Trừ khi có quy định khác, các dạng mạo từ số ít bao gồm cả đề cập đến số nhiều. Như vậy, ví dụ, việc đề cập đến "vector biểu hiện" bao gồm một vector biểu hiện cũng như nhiều vector biểu hiện, và việc đề cập đến "tế bào chủ" bao gồm việc đề cập đến một hoặc nhiều tế bào chủ, và các đề cập tương tự. Cũng lưu ý rằng các điểm yêu cầu bảo hộ có thể được soạn thảo để loại trừ dấu hiệu tùy chọn bất kỳ. Như vậy, sự trình bày này được dùng để làm cơ sở tiền đề cho việc dùng thuật ngữ mang tính chất loại trừ như "chỉ có" "chỉ" và các thuật ngữ tương tự đối với các dấu hiệu của yêu cầu bảo hộ hoặc sử dụng giới hạn "phủ định".

Trừ khi có quy định khác, sáng chế không bị giới hạn ở các trình tự, vector biểu hiện, enzym, vi sinh vật chủ, hoặc quy trình cụ thể, vì chúng có thể thay đổi theo hiểu biết của người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Thuật ngữ được dùng ở đây chỉ nhằm mục đích mô tả sáng chế theo các phương án thực hiện cụ thể và không nhằm mục đích giới hạn nó.

Tế bào chủ

Có thể sử dụng tế bào chủ thích hợp bất kỳ để thực hiện sáng chế. Theo một phương án, tế bào chủ này là vi sinh vật chủ đã được cải biến di truyền trong đó các phân tử axit nucleic đã được chèn, được làm mất hoặc được cải biến, (tức là được gây đột biến, ví dụ, bằng cách chèn, làm mất, thay thế, và/hoặc đảo các nucleotit), để sản xuất ra hợp chất isoprenoit hoặc dẫn xuất isoprenoit mong muốn, hoặc làm tăng hiệu suất của hợp chất isoprenoit hoặc dẫn xuất isoprenoit mong muốn. Theo một phương án khác, tế bào chủ có khả năng sinh trưởng trong môi trường nuôi cấy lỏng. Trái lại, "tế bào đối chứng" là

đối tượng hoặc mẫu thay thế được sử dụng trong thử nghiệm nhằm mục đích so sánh, và thường là tế bào mẹ không chứa (các) cải biến được tạo ra cho tế bào chủ tương ứng.

Ví dụ minh họa về tế bào chủ thích hợp bao gồm tế bào cổ khuẩn, tế bào nhân sơ, hoặc tế bào nhân chuẩn bất kỳ. Ví dụ về tế bào cổ khuẩn bao gồm, nhưng không bị giới hạn ở các tế bào thuộc loài: *Aeropyrum*, *Archaeoglobus*, *Halobacterium*, *Methanococcus*, *Methanobacterium*, *Pyrococcus*, *Sulfolobus*, và *Thermoplasma*. Ví dụ minh họa về các chủng cổ khuẩn bao gồm, nhưng không bị giới hạn ở: *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Pyrococcus abyssi*, *Pyrococcus horikoshii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium*.

Ví dụ về tế bào nhân sơ bao gồm, nhưng không bị giới hạn ở các tế bào thuộc chi: *Agrobacterium*, *Alicyclobacillus*, *Anabaena*, *Anacystis*, *Arthrobacter*, *Azobacter*, *Bacillus*, *Brevibacterium*, *Chromatium*, *Clostridium*, *Corynebacterium*, *Enterobacter*, *Erwinia*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Mesorhizobium*, *Methylobacterium*, *Microbacterium*, *Phomnidium*, *Pseudomonas*, *Rhodobacter*, *Rhodopseudomonas*, *Rhodospirillum*, *Rhodococcus*, *Salmonella*, *Scenedesmus*, *Serratia*, *Shigella*, *Staphylococcus*, *Streptomyces*, *Synnecoccus*, và *Zymomonas*.

Ví dụ minh họa về các chủng vi khuẩn chưa có nhân điển hình bao gồm, nhưng không bị giới hạn ở: *Bacillus subtilis*, *Bacillus amyloliquefacines*, *Brevibacterium ammoniagenes*, *Brevibacterium immariophilum*, *Clostridium beigerinckii*, *Enterobacter sakazakii*, *Escherichia coli*, *Lactococcus lactis*, *Mesorhizobium loti*, *Pseudomonas aeruginosa*, *Pseudomonas mevalonii*, *Pseudomonas pudica*, *Rhodobacter capsulatus*, *Rhodobacter sphaeroides*, *Rhodospirillum rubrum*, *Salmonella enterica*, *Salmonella typhi*, *Salmonella typhimurium*, *Shigella dysenteriae*, *Shigella flexneri*, *Shigella sonnei*, *Staphylococcus aureus*, và các chủng vi khuẩn tương tự.

Nói chung, nếu tế bào chủ vi khuẩn được sử dụng, thì chúng không gây bệnh được ưu tiên. Ví dụ minh họa về các chủng không gây bệnh bao gồm, nhưng không bị giới hạn ở *Bacillus subtilis*, *Escherichia coli*, *Lactobacillus acidophilus*, *Lactobacillus helveticus*, *Pseudomonas aeruginosa*, *Pseudomonas mevalonii*, *Pseudomonas putida*, *Rhodobacter sphaeroides*, *Rhodobacter capsulatus*, *Rhodospirillum rubrum*, và các chủng vi khuẩn tương tự.

Ví dụ về tế bào nhân chuẩn bao gồm, nhưng không bị giới hạn ở tế bào nấm. Ví dụ về tế bào nấm bao gồm, nhưng không bị giới hạn ở các tế bào thuộc chi: *Aspergillus*, *Candida*, *Chrysosporium*, *Cryptococcus*, *Fusarium*, *Kluyveromyces*, *Neotryphodium*, *Neurospora*, *Penicillium*, *Pichia*, *Saccharomyces*, *Trichoderma* và *Xanthophyllomyces* (trước đây là *Phaffia*).

Ví dụ minh họa về các chủng sinh vật nhân chuẩn bao gồm, nhưng không bị giới hạn ở: *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Candida albicans*, *Chrysosporium lucknowense*, *Fusarium graminearum*, *Fusarium venenatum*, *Kluyveromyces lactis*, *Neurospora crassa*, *Pichia angusta*, *Pichia finlandica*, *Pichia kodamae*, *Pichia membranaefaciens*, *Pichia methanolica*, *Pichia opuntiae*, *Pichia pastoris*, *Pichia pijperi*, *Pichia quercuum*, *Pichia salictaria*, *Pichia thermotolerans*, *Pichia trehalophila*, *Pichia stipitis*, *Streptomyces ambofaciens*, *Streptomyces aureofaciens*, *Streptomyces aureus*, *Saccaromyces bayanus*, *Saccaromyces boulardi*, *Saccharomyces cerevisiae*, *Streptomyces fungicidicus*, *Streptomyces griseochromogenes*, *Streptomyces griseus*, *Streptomyces lividans*, *Streptomyces olivogriseus*, *Streptomyces rameus*, *Streptomyces tanashiensis*, *Streptomyces vinaceus*, *Trichoderma reesei* và *Xanthophyllomyces dendrorhous* (trước đây là *Phaffia rhodozyma*).

Nói chung, nếu tế bào nhân chuẩn được sử dụng, thì chúng không gây bệnh được ưu tiên. Ví dụ minh họa về các chủng không gây bệnh bao gồm, nhưng không bị giới hạn ở: *Fusarium graminearum*, *Fusarium venenatum*, *Pichia pastoris*, *Saccaromyces boulardi*, và *Saccaromyces cerevisiae*.

Ngoài ra, các chủng nhất định đã được Cục quản lý Thực phẩm và Dược phẩm Hoa kỳ chỉ định là GRAS hoặc được đánh giá chung là an toàn. Các chủng này bao gồm: *Bacillus subtilis*, *Lactibacillus acidophilus*, *Lactobacillus helveticus*, và *Saccharomyces cerevisiae*.

Các quá trình IPP

Tế bào chủ theo sáng chế chứa hoặc sử dụng quá trình MEV, quá trình DXP hoặc cả hai quá trình này để tổng hợp IPP và dạng đồng phân của nó, DMAPP. Nói chung, các sinh vật có nhân điển hình không phải thực vật chỉ sử dụng quá trình MEV isoprenoit để chuyển hoá axetyl-CoA thành IPP, chất này sau đó được đồng phân hoá thành DMAPP. Sinh vật nhân sơ, trừ một số ngoại lệ, sử dụng quá trình không phụ thuộc mevalonat hoặc quá trình DXP để sản xuất IPP và DMAPP một cách riêng biệt thông qua điểm nhánh. Nói chung, thực vật sử dụng cả quá trình MEV và DXP để tổng hợp IPP.

Quá trình MEV

Fig.1A là sơ đồ biểu diễn quá trình MEV. Nói chung, quá trình này bao gồm sáu bước.

Ở bước thứ nhất, hai phân tử axetyl-coenzym A được kết hợp nhờ enzym để tạo thành axetoaxetyl-CoA. Enzym đã biết để xúc tác cho bước này là, ví dụ, axetyl-CoA thiolaza (còn được gọi là axetyl-CoA axetyltransferaza). Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở các số đăng ký ngân hàng gen và sinh vật mà từ đó các trình tự này thu được: (NC_000913 REGION: 2324131..2325315; *Escherichia coli*), (D49362; *Paracoccus denitrificans*), và (L20428; *Saccharomyces cerevisiae*).

Ở bước thứ hai của quá trình MEV, axetoaxetyl-CoA được ngưng tụ bằng enzym với phân tử axetyl-CoA khác để tạo ra 3-hydroxy-3-metylglutaryl-CoA (3-hydroxy-3-methylglutaryl-CoA – (HMG-CoA)). Enzym đã biết để xúc tác cho bước này là, ví dụ, HMG-CoA synthaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (NC_001145. bộ 19061..20536; *Saccharomyces cerevisiae*), (X96617; *Saccharomyces cerevisiae*), (X83882;

Arabidopsis thaliana), (AB037907; *Kitasatospora griseola*), (BT007302; *Homo sapiens*), và (NC_002758, Locus tag SAV2546, Gene ID 1122571; *Staphylococcus aureus*).

Ở bước thứ ba, HMG-CoA được chuyển hoá bằng enzym thành mevalonat. Enzym đã biết để xúc tác cho bước này là, ví dụ, HMG-CoA reductaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (NM_206548; *Drosophila melanogaster*), (NC_002758, Locus tag SAV2545, Gene ID 1122570; *Staphylococcus aureus*), (NM_204485; *Gallus gallus*), (AB015627; *Streptomyces sp.* KO 3988), (AF542543; *Nicotiana attenuata*), (AB037907; *Kitasatospora griseola*), (AX128213, cung cấp trình tự mã hóa HMGR bị cắt cụt; *Saccharomyces cerevisiae*), và (NC_001145: bổ sung (115734..118898; *Saccharomyces cerevisiae*).

Ở bước thứ tư, mevalonat được phosphoryl hoá bằng enzym để tạo thành mevalonat 5-phosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, mevalonat kinaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (L77688; *Arabidopsis thaliana*), và (X55875; *Saccharomyces cerevisiae*).

Ở bước thứ năm, nhóm phosphat thứ hai được bổ sung bằng enzym vào mevalonat 5-phosphat để tạo thành mevalonat 5-pyrophosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, phosphomevalonat kinaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AF429385; *Hevea brasiliensis*), (NM_006556; *Homo sapiens*), và (NC_001145. bộ 712315..713670; *Saccharomyces cerevisiae*).

Ở bước thứ sáu, mevalonat 5-pyrophosphat được chuyển hoá bằng enzym thành IPP. Enzym đã biết để xúc tác cho bước này là, ví dụ, mevalonat pyrophosphat decarboxylaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (X97557; *Saccharomyces cerevisiae*), (AF290095; *Enterococcus faecium*), và (U49260; *Homo sapiens*).

Nếu IPP được chuyển hoá thành DMAPP, thì cần phải có bước thứ bảy. Enzym đã biết để xúc tác cho bước này là, ví dụ, IPP isomeraza. Ví dụ minh

họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (NC_000913, 3031087..3031635; *Escherichia coli*), và (AF082326; *Haematococcus pluvialis*). Nếu cần phải có sự chuyển hoá thành DMAPP, thì sự biểu hiện IPP isomeraza được tăng lên đảm bảo rằng việc chuyển hoá IPP thành DMAPP không biểu thị bước giới hạn tốc độ trong toàn bộ quá trình này.

Quá trình DXP

Quá trình DXP được thể hiện bằng sơ đồ trên Fig.1B. Nói chung, quá trình DXP bao gồm bảy bước. Ở bước thứ nhất, pyruvat được ngưng tụ với D-glyxeraldehyt 3-phosphat để tạo ra 1-deoxy-D-xyluloza-5-phosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, 1-deoxy-D-xyluloza-5-phosphat synthaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AF035440; *Escherichia coli*), (NC_002947, locus tag PP0527; *Pseudomonas putida* KT2440), (CP000026, locus tag SPA2301; *Salmonella enterica* Paratyphi, xem ATCC 9150), (NC_007493, locus tag RSP_0254; *Rhodobacter sphaeroides* 2.4.1), (NC_005296, locus tag RPA0952; *Rhodopseudomonas palustris* CGA009), (NC_004556, locus tag PD1293; *Xylella fastidiosa* Temecula1), và (NC_003076, locus tag AT5G11380; *Arabidopsis thaliana*).

Ở bước thứ hai, 1-deoxy-D-xyluloza-5-phosphat được chuyển hoá thành 2C-metyl-D-erythritol-4-phosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, 1-deoxy-D-xyluloza-5-phosphat reductoisomeraza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AB013300; *Escherichia coli*), (AF148852; *Arabidopsis thaliana*), (NC_002947, locus tag PP1597; *Pseudomonas putida* KT2440), (AL939124, locus tag SCO5694; *Streptomyces coelicolor* A3(2)), (NC_007493, locus tag RSP_2709; *Rhodobacter sphaeroides* 2.4.1), và (NC_007492, locus tag Pfl_1107; *Pseudomonas fluorescens* PfO-1).

Ở bước thứ ba, 2C-metyl-D-erythritol-4-phosphat được chuyển hoá thành 4-diphosphoxytidyl-2C-metyl-D-erythritol. Enzym đã biết để xúc tác cho bước này là, ví dụ, 4-diphosphoxytidyl-2C-metyl-D-erythritol synthaza. Ví dụ

minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AF230736; *Escherichia coli*), (NC_007493, locus_tag RSP_2835; *Rhodobacter sphaeroides* 2.4.1), (NC_003071, locus_tag AT2G02500; *Arabidopsis thaliana*), và (NC_002947, locus_tag PP1614; *Pseudomonas putida* KT2440).

Ở bước thứ tư, 4-diphosphoxytidyl-2C-metyl-D-erythritol được chuyển hoá thành 4-diphosphoxytidyl-2C-metyl-D-erythritol-2-phosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, 4-diphosphoxytidyl-2C-metyl-D-erythritol kinaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AF216300; *Escherichia coli*) và (NC_007493, locus_tag RSP_1779; *Rhodobacter sphaeroides* 2.4.1).

Ở bước thứ năm, 4-diphosphoxytidyl-2C-metyl-D-erythritol-2-phosphat được chuyển hoá thành 2C-metyl-D-erythritol 2,4-xyclodiphosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, 2C-metyl-D-erythritol 2,4-xyclodiphosphat synthaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AF230738, *Escherichia coli*), (NC_007493, locus_tag RSP_6071; *Rhodobacter sphaeroides* 2.4.1), và (NC_002947, locus_tag PP1618; *Pseudomonas putida* KT2440).

Ở bước thứ sáu, 2C-metyl-D-erythritol 2,4-xyclodiphosphat được chuyển hoá thành 1-hydroxy-2-metyl-2-(E)-butenyl-4-diphosphat. Enzym đã biết để xúc tác cho bước này là, ví dụ, 1-hydroxy-2-metyl-2-(E)-butenyl-4-diphosphat synthaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở: (AY033515, *Escherichia coli*), (NC_002947, locus_tag PP0853; *Pseudomonas putida* KT2440), và (NC_007493, locus_tag RSP_2982; *Rhodobacter sphaeroides* 2.4.1).

Ở bước thứ bảy, 1-hydroxy-2-metyl-2-(E)-butenyl-4-diphosphat được chuyển hoá thành IPP hoặc dạng đồng phân của nó, DMAPP. Enzym đã biết để xúc tác cho bước này là, ví dụ, isopentyl/dimetylalyl diphosphat synthaza. Ví dụ minh họa về các trình tự nucleotit bao gồm, nhưng không bị giới hạn ở:

(AY062212; *Escherichia coli*) và (NC_002947, locus_tag PP0606; *Pseudomonas putida* KT2440).

Theo một số phương án, "sự giao tiếp chéo" (hoặc sự gây nhiễu) giữa chính các quá trình trao đổi chất của tế bào chủ và các quy trình khác liên quan đến việc sản xuất IPP như được đề xuất theo sáng chế được giảm đến mức tối thiểu hoặc được loại bỏ hoàn toàn. Ví dụ, sự giao tiếp chéo được giảm đến mức tối thiểu hoặc được loại bỏ hoàn toàn khi vi sinh vật chủ chỉ dựa trên quá trình DXP để tổng hợp IPP, và quá trình MEV được đưa vào để tạo ra IPP bổ sung. Các sinh vật chủ này sẽ không được trang bị để thay đổi sự biểu hiện của các enzym của quá trình MEV hoặc xử lý các sản phẩm trung gian liên quan đến quá trình MEV. Sinh vật chỉ dựa hoặc chủ yếu dựa vào quá trình DXP là, ví dụ, *Escherichia coli*.

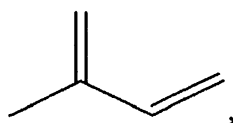
Theo một số phương án, tế bào chủ sản xuất ra IPP thông qua quá trình MEV, hoặc một cách duy nhất hoặc kết hợp với quá trình DXP. Theo phương án khác, quá trình DXP của vật chủ bị bất hoạt về mặt chức năng sao cho tế bào chủ sản xuất ra IPP một cách duy nhất thông qua quá trình MEV được đưa vào theo cách khác loại. Quá trình DXP có thể bị bất hoạt về mặt chức năng bằng cách làm bất hoạt sự biểu hiện gen hoặc bất hoạt chức năng của một hoặc nhiều enzym của quá trình DXP.

Các hợp chất C₅

Các hợp chất C₅ theo sáng chế thường được tạo ra từ IPP hoặc DMAPP. Các hợp chất này còn được gọi là hemiterpen vì chúng được tạo ra từ một đơn vị isopren đơn (IPP hoặc DMAPP).

Isopren

Isopren, có công thức cấu tạo:



được tìm thấy trong nhiều thực vật. Isopren được sản xuất từ IPP nhờ isopren synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (AB198190; *Populus alba*) và (AJ294819; *Populus alba* x *Populus tremula*).

Các hợp chất C₁₀

Các hợp chất C₁₀ theo sáng chế thường được tạo ra từ geranyl pyrophosphat (GPP) mà được tạo thành bằng việc ngưng tụ IPP với DMAPP. Enzym đã biết để xúc tác cho bước này là, ví dụ, geranyl pyrophosphat synthaza. Các hợp chất C₁₀ này còn được gọi là monoterpen vì chúng được tạo ra từ hai đơn vị isopren.

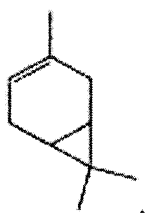
Fig.2 là sơ đồ thể hiện cách IPP và DMAPP có thể sản xuất ra GPP, chất này có thể được xử lý tiếp thành monoterpen.

Ví dụ minh họa về các trình tự nucleotit mã hóa geranyl pyrophosphat synthaza bao gồm, nhưng không bị giới hạn ở: (AF513111; *Abies grandis*), (AF513112; *Abies grandis*), (AF513113; *Abies grandis*), (AY534686; *Antirrhinum majus*), (AY534687; *Antirrhinum majus*), (Y17376; *Arabidopsis thaliana*), (AE016877, Locus AP11092; *Bacillus cereus*; ATCC 14579), (AJ243739; *Citrus sinensis*), (AY534745; *Clarkia breweri*), (AY953508; *Ips pini*), (DQ286930; *Lycopersicon esculentum*), (AF182828; *Mentha x piperita*), (AF182827; *Mentha x piperita*), (MPI249453; *Mentha x piperita*), (PZE431697, Locus CAD24425; *Paracoccus zeaxanthinifaciens*), (AY866498; *Picrorhiza kurrooa*), (AY351862; *Vitis vinifera*), và (AF203881, Locus AAF12843; *Zymomonas mobilis*).

Sau đó, GPP được chuyển hoá tiếp thành nhiều hợp chất C₁₀ khác nhau. Ví dụ minh họa về các hợp chất C₁₀ bao gồm, nhưng không bị giới hạn ở:

Caren

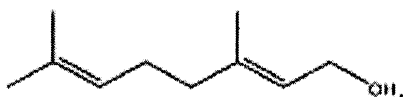
Caren, có công thức cấu tạo:



được tìm thấy trong nhựa của nhiều loại cây, đặc biệt là cây thông. Caren được sản xuất từ GPP nhờ caren synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (AF461460, REGION 43..1926; *Picea abies*) và (AF527416, REGION: 78..1871; *Salvia stenophylla*).

Geraniol

Geraniol (còn được gọi là rhodnol), có công thức cấu tạo:



là thành phần chính của dầu hoa hồng và dầu sả hoa hồng. Nó cũng có mặt trong cây phong lữ, cây chanh, và cây sả. Geraniol được sản xuất từ GPP nhờ geraniol synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (AJ457070; *Cinnamomum tenuipilum*), (AY362553; *Ocimum basilicum*), (DQ234300; chủng *Perilla frutescens* 1864), (DQ234299; chủng *Perilla citriodora* 1861), (DQ234298; chủng *Perilla citriodora* 4935), và (DQ088667; *Perilla citriodora*).

Linalool

Linalool, có công thức cấu tạo:

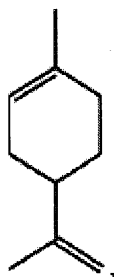


được tìm thấy trong nhiều loài hoa và thực vật làm gia vị như hạt cây rau mùi. Linalool được sản xuất từ GPP nhờ linalool synthaza. Ví dụ minh họa về các

trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (AF497485; *Arabidopsis thaliana*), (AC002294, Locus AAB71482; *Arabidopsis thaliana*), (AY059757; *Arabidopsis thaliana*), (NM_104793; *Arabidopsis thaliana*), (AF154124; *Artemisia annua*), (AF067603; *Clarkia breweri*), (AF067602; *Clarkia concinna*), (AF067601; *Clarkia breweri*), (U58314; *Clarkia breweri*), (AY840091; *Lycopersicon esculentum*), (DQ263741; *Lavandula angustifolia*), (AY083653; *Mentha citrate*), (AY693647; *Ocimum basilicum*), (XM_463918; *Oryza sativa*), (AP004078, Locus BAD07605; *Oryza sativa*), (XM_463918, Locus_XP_463918; *Oryza sativa*), (AY917193; *Perilla citriodora*), (AF271259; *Perilla frutescens*), (AY473623; *Picea abies*), (DQ195274; *Picea sitchensis*), và (AF444798; giống *Perilla frutescens* var. *crispa* số 79).

Limonen

Limonen, có công thức cấu tạo:



được tìm thấy trong vỏ quả cây có mùi và cây bạc hà. Limonen được sản xuất từ GPP nhờ limonen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (+)-limonen synthaza (AF514287, REGION: 47..1867; *Citrus limon*) và (AY055214, REGION: 48..1889; *Agastache rugosa*) và (-)-limonen synthaza (DQ195275, REGION: 1..1905; *Picea sitchensis*), (AF006193, REGION: 73..1986; *Abies grandis*), và (MHC4SLSP, REGION: 29..1828; *Mentha spicata*).

Myrxen

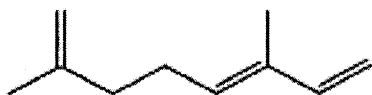
Myrxen, có công thức cấu tạo:



được tìm thấy trong tinh dầu ở nhiều loại thực vật bao gồm cây nguyệt quế, cỏ roi ngựa, và cây thanh mai mà từ loại cây này nó có tên như vậy. Myrcen được sản xuất từ GPP nhờ myrcen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (U87908; *Abies grandis*), (AY195609; *Antirrhinum majus*), (AY195608; *Antirrhinum majus*), (NM_127982; *Arabidopsis thaliana* TPS1O), (NM_113485; *Arabidopsis thaliana* ATTPS-CIN), (NM_113483; *Arabidopsis thaliana* ATTPS-CIN), (AF271259; *Perilla frutescens*), (AY473626; *Picea abies*), (AF369919; *Picea abies*), và (AJ304839; *Quercus ilex*).

Ocimen

α - và β -Ocimen, lần lượt có công thức cấu tạo:



và



được tìm thấy ở nhiều loại cây và quả bao gồm *Ocimum basilicum* và được sản xuất từ GPP nhờ ocimen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (AY195607; *Antirrhinum majus*), (AY195609; *Antirrhinum majus*), (AY195608; *Antirrhinum majus*), (AK221024; *Arabidopsis thaliana*), (NM_113485; *Arabidopsis thaliana* ATTPS-CIN), (NM_113483; *Arabidopsis thaliana* ATTPS-CIN), (NM_117775; *Arabidopsis thaliana* ATTPS03), (NM_001036574; *Arabidopsis thaliana* ATTPS03), (NM_127982; *Arabidopsis thaliana* TPS10), (AB110642; *Citrus unshiu* CitMTSL4), và (AY575970; *Lotus corniculatus* var. *japonicus*).

α -Pinen

α -Pinen, có công thức cấu tạo:



được tìm thấy trong cây thông và bạch đàn. α -Pinen được sản xuất từ GPP nhờ α -pinen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (+) α -pinen synthaza (AF543530, REGION 1...1887, *Pinus taeda*), (-) α -pinen synthaza (AF543527, REGION 32...1921, *Pinus taeda*), và (+)/(-) α -pinen synthaza (AGU87909, REGION 6111892, *Abies grandis*).

α -Pinen

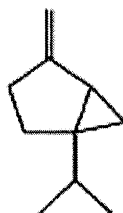
β -Pinen, có công thức cấu tạo:



được tìm thấy trong cây thông, cây hương thảo, cây mùi tây, cây thì là, cây húng quế, và cây hoa hồng. β -Pinen được sản xuất từ GPP nhờ β -pinen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở (-) β -pinen synthaza (AF276072, REGION: 1..1749, *Artemisia annua*) và (AF514288, REGION: 26..1834, *Citrus limon*).

Sabinen

Sabinen, có công thức cấu tạo:



được tìm thấy trong hạt tiêu đen, hạt cà rốt, cây xô thơm, và cây chè. Sabinen được sản xuất từ GPP bởi sabinen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở AF051901, REGION 26..1798 từ *Salvia officinalis*.

γ -Terpinen

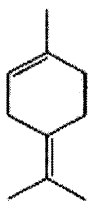
γ -Terpinen, có công thức cấu tạo:



là thành phần của tinh dầu từ quả cây có múi. Bằng quá trình hoá sinh, γ -terpinen được sản xuất từ GPP nhờ γ -terpinen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm: (AF514286, REGION: 30..1832 từ *Citrus limon*) và (AB110640, REGION: 1..1803 từ *Citrus unshiu*)

Terpinolen

Terpinolen, có công thức cấu tạo:



được tìm thấy trong cây lý chua, cây bách, cây ôi, cây vải, cây đu đủ, cây thông, và cây chè. Terpinolen được sản xuất từ GPP nhờ terpinolen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở AY906866, REGION: 10..1887 từ *Pseudotsuga menziesii*.

Các hợp chất C₁₅

Các hợp chất C₁₅ theo sáng chế thường được tạo ra từ farnesyl pyrophosphat (FPP), chất này được sản xuất bằng cách ngưng tụ hai phân tử IPP với một phân tử DMAPP. Enzym đã biết để xúc tác cho bước này là, ví dụ, farnesyl pyrophosphat synthaza. Các hợp chất C₁₅ này còn được gọi là sesquiterpen vì chúng được tạo ra từ ba đơn vị isopren.

Fig.2 là đồ thể hiện cách IPP và DMAPP có thể được kết hợp để sản xuất FPP, chất này có thể được xử lý tiếp thành sesquiterpen.

Ví dụ minh họa về các trình tự nucleotit mã hóa farnesyl pyrophosphat synthaza bao gồm, nhưng không bị giới hạn ở: (ATU80605; *Arabidopsis thaliana*), (ATHFPS2R, *Arabidopsis thaliana*), (AAU36376, *Artemisia annua*), (AF461050; *Bos taurus*), (D00694; *Escherichia coli* K-12), (AE009951, Locus AAL95523; *Fusobacterium nucleatum subsp. nucleatum* ATCC 25586), (GFFPPSGEN; *Gibberella fujikuroi*), (CP000009, Locus AAW60034, *Gluconobacter oxydans* 621H), (AF019892, *Helianthus annuus*), (HUMFAPS; *Homo sapiens*), (KLFPFSQCR, *Kluyveromyces lactis*), (LAU15777, *Lupinus albus*), (LAU20771, *Lupinus albus*), (AF309508, *Mus musculus*), (NCFPPSGEN; *Neurospora crassa*), (PAFPS1, *Parthenium argentatum*), (PAFPS2, *Parthenium argentatum*), (RATFAPS; *Rattus norvegicus*), (YSCFPP; *Saccharomyces cerevisiae*), (D89104; *Schizosaccharomyces pombe*), (CP000003, Locus AAT87386; *Streptococcus pyogenes*), (CP000017, Locus AAZ51849, *Streptococcus pyogenes*), (NC_008022, Locus YP_598856, *Streptococcus pyogenes* MGAS10270), (NC_008023, Locus YP_600845, *Streptococcus pyogenes* MGAS2096), (NC_008024, Locus YP_602832, *Streptococcus pyogenes* MGAS 10750), và (MZEFPS, *Zea mays*).

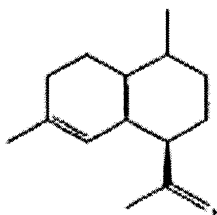
Theo cách khác, FPP còn có thể được sản xuất bằng cách cộng IPP vào GPP. Ví dụ minh họa về các trình tự nucleotit mã hóa enzym có khả năng tham gia vào phản ứng này bao gồm, nhưng không bị giới hạn ở: (AE000657, Locus AAC06913, *Aquifex aeolicus* VF5), (NM_202836, *Arabidopsis thaliana*), (D84432, Locus BAA12575; *Bacillus subtilis*), (U12678, Locus AAC28894; *Bradyrhizobium japonicum* USDA 110), (BACFDPS; *Geobacillus*

stearothermophilus), (NC_002940, Locus NP_873754; *Haemophilus ducreyi* 35000HP), (L42023, Locus AAC23087, *Haemophilus influenzae* Rd KW20), (J05262, *Homo sapiens*), (YP_395294, *Lactobacillus sakei* subsp. *sakei* 23K), (NC_005823, Locus YP_000273; *Leptospira interrogans* serovar *Copenhageni* str *Fiocruz* L1-130), (AB003187; *Micrococcus luteus*), (NC_002946, Locus YP_208768; *Neisseria gonorrhoeae* FA 1090), (U00090, Locus AAB91752, *Rhizobium* sp. NGR234), (J05091; *Saccharomyces cerevisiae*), (CP000031, Locus AAV93568; *Silicibacter pomeroyi* DSS-3), (AE008481, Locus AAK99890; *Streptococcus pneumoniae* R6), và (NC_004556, Locus NP_779706; *Xylella fastidiosa* Temecula1).

Sau đó FPP được chuyển hoá tiếp thành nhiều hợp chất C₁₅ khác nhau. Ví dụ minh họa về các hợp chất C₁₅ bao gồm, nhưng không bị giới hạn ở:

Amorphadien

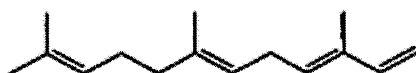
Amorphadien, có công thức cấu tạo:



là tiền chất đối với artemisinin được sản xuất bởi *Artemisia anna*. Amorphadien được sản xuất từ FPP nhờ amorphadien synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp là SEQ ID NO. 37 của Patent Mỹ số 7,192,751.

α-Farnesen

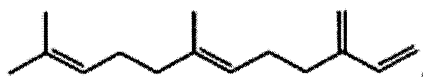
α-Farnesen, có công thức cấu tạo:



được tìm thấy trong nhiều nguồn sinh học khác nhau bao gồm, nhưng không bị giới hạn ở tuyến Dufour của kiến và trong vỏ táo và vỏ lê. α -Farnesen được sản xuất từ FPP nhờ α -farnesen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở DQ309034 từ *Pyrus communis* cultivar d'Anjou (lê; tên gen AFS1) và AY182241 từ *Malus domestica* (táo; gen AFS1). Xem ấn phẩm: Pechouus *et al.*, *Planta* **219(1)**:84-94 (2004).

β -Farnesen

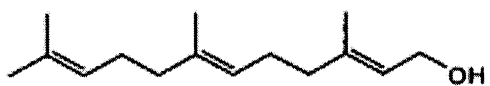
β -Farnesen, có công thức cấu tạo:



được tìm thấy trong nhiều nguồn sinh học khác nhau bao gồm, nhưng không bị giới hạn ở rễ vùng và tinh dầu như từ tinh dầu bạc hà. Trong một số thực vật như khoai tây dại, β -farnesen được tổng hợp để làm chất diệt côn trùng tự nhiên. β -Farnesen được sản xuất từ FPP nhờ β -farnesen synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở số đăng ký ngân hàng gen AF024615 từ *Mentha x piperita* (bạc hà; gen Tsps 11), và AY835398 từ *Artemisia annua*. Xem ấn phẩm: Picaud *et al.*, *Phytochemistry* **66(9)**: 961-967 (2005).

Farnesol

Farnesol, có công thức cấu tạo:

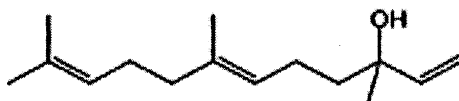


được tìm thấy trong nhiều nguồn sinh học khác nhau bao gồm côn trùng và tinh dầu như từ cây sả chanh, dầu hoa cam, cây hoa anh thảo, cây sả, cây hoa huệ, và cây hoa hồng. Farnesol được sản xuất từ FPP nhờ hydroxylaza như farnesol synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở số đăng ký ngân hàng gen AF529266 từ *Zea mays* và

YDR481C từ *Saccharomyces cerevisiae* (gen Pho8). Xem ấn phẩm: Song, L., *Applied Biochemistry and Biotechnology* **128**:149-158 (2006).

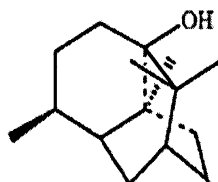
Nerolidol

Nerolidol, có công thức cấu tạo:



còn được gọi là peruvicol, và được tìm thấy ở nhiều nguồn sinh học khác nhau bao gồm là tinh dầu như từ dầu hoa cam, cây gừng, cây hoa nhài, cây oải hương, cây chè, và cây sả. Nerolidol được sản xuất từ FPP nhờ hydroxylaza như nerolidol synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở AF529266 từ *Zea mays* (ngô, gen tps1).

Patchoulol

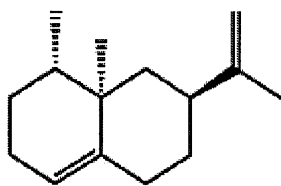


Patchoulol, có công thức cấu tạo:

còn được biết đến là rượu cây hoắc hương và là thành phần của tinh dầu của *Pogostemon patchouli*. Patchoulol được sản xuất từ FPP nhờ patchoulol synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở AY508730 REGION: 1..1659 từ *Pogostemon cablin*.

Valencen

Valencen, có công thức cấu tạo:



là một trong số các thành phần hoá học chính của mùi và vị của cây cam và được tìm thấy ở vỏ quả cam. Valencen được sản xuất từ FPP nhờ nootkaton synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở AF441124 REGION: 1..1647 từ *Citrus sinensis* và AY917195 REGION: 1..1653 từ *Perilla frutescens*.

Các hợp chất C₂₀

Các hợp chất C₂₀ theo sáng chế thường được tạo ra từ geranylgeraniol pyrophosphat (GGPP) mà được sản xuất bằng việc ngưng tụ ba phân tử IPP với một phân tử DMAPP. Enzym đã biết để xúc tác cho bước này là, ví dụ, geranylgeranyl pyrophosphat synthaza. Các hợp chất C₂₀ này còn được gọi là diterpen vì chúng được tạo ra từ bốn đơn vị isopren.

Fig.2 là sơ đồ thể hiện cách IPP và DMAPP có thể được kết hợp để sản xuất GGPP, chất này có thể được xử lý tiếp thành diterpen, hoặc có thể được xử lý tiếp để sản xuất carotenoid.

Ví dụ minh họa về các trình tự nucleotit mã hóa geranylgeranyl pyrophosphat synthaza bao gồm, nhưng không bị giới hạn ở: (ATHGERPYRS, *Arabidopsis thaliana*), (BT005328, *Arabidopsis thaliana*), (NM_119845, *Arabidopsis thaliana*), (NZ_AAJM01000380, Locus ZP_00743052, *Bacillus thuringiensis serovar israelensis*, ATCC 35646 sql563), (CRGGPPS, *Catharanthus roseus*), (NZ_AABF02000074, Locus ZP_00144509, *Fusobacterium nucleatum subsp. vincentii*, ATCC 49256), (GFGGPPSGN, *Gibberella fujikuroi*), (AY 371321; *Ginkgo biloba*), (AB055496; *Hevea brasiliensis*), (AB017971; *Homo sapiens*), (MCI276129; *Mucor circinelloides f. lusitanicus*), (AB016044; *Mus musculus*), (AABX01000298, Locus NCU01427; *Neurospora crassa*), (NCU20940; *Neurospora crassa*), (NZ_AAKL01000008, Locus ZP_00943566; *Ralstonia solanacearum*

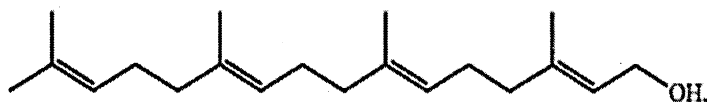
UW551), (AB118238; *Rattus norvegicus*), (SCU31632; *Saccharomyces cerevisiae*), (AB016095; *Synechococcus elongates*), (SAGGPS; *Sinapis alba*), (SSOGDS; *Sulfolobus acidocaldarius*), (NC_007759, Locus_YP 461832; *Syntrophus aciditrophicus* SB), và (NC_006840, Locus_YP 204095; *Vibrio fischeri* ES114).

Theo cách khác, GGPP còn có thể được sản xuất bằng cách cộng IPP vào FPP. Ví dụ minh họa về các trình tự nucleotit mã hóa enzym có khả năng tham gia phản ứng này bao gồm, nhưng không bị giới hạn ở: (NM_112315; *Arabidopsis thaliana*), (ERWCRTE; *Pantoea agglomerans*), (D90087, Locus BAA14124; *Pantoea ananatis*), (X52291, Locus CAA36538; *Rhodobacter capsulatus*), (AF195122, Locus AAF24294; *Rhodobacter sphaeroides*), và (NC_004350, Locus NP_721015; *Streptococcus mutans* UA159).

Sau đó, GGPP được chuyển hoá tiếp thành nhiều C₂₀ isoprenoit khác nhau. Ví dụ minh họa về các hợp chất C₂₀ bao gồm, nhưng không bị giới hạn ở:

Geranylgeraniol

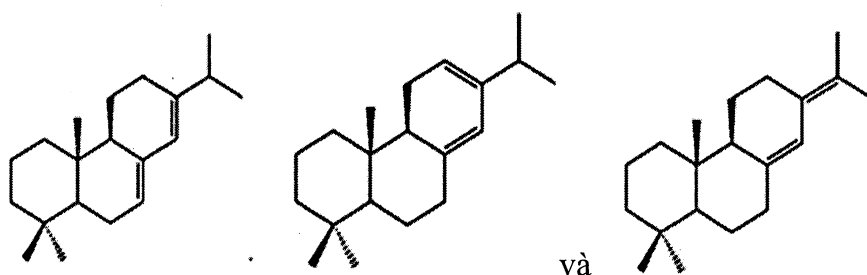
Geranylgeraniol, có công thức cấu tạo:



là thành phần của dầu gỗ từ *Cedrela toona* và của dầu lạnh. Geranylgeraniol có thể được sản xuất, ví dụ, bằng cách bổ sung gen mã hóa phosphatase vào các cấu trúc biểu hiện sau gen mã hóa GGPP synthase.

Abietadien

Abietadien bao gồm các chất đồng phân sau:



và được tìm thấy ở cây như *Abies grandis*. Abietadien được sản xuất bởi abietadien synthaza. Ví dụ minh họa về các trình tự nucleotit thích hợp bao gồm, nhưng không bị giới hạn ở: (U50768; *Abies grandis*) và (AY473621; *Picea abies*).

Các hợp chất C₂₀₊

Các hợp chất C₂₀₊ cũng bao gồm trong phạm vi của sáng chế. Ví dụ minh họa về các hợp chất này bao gồm sesterterpen (hợp chất C₂₅ được tạo ra từ năm đơn vị isopren), triterpen (hợp chất C₃₀ được tạo ra từ sáu đơn vị isopren), và tetraterpen (hợp chất C₄₀ được tạo ra từ tám đơn vị isopren). Các hợp chất này được tạo thành bằng cách sử dụng các phương pháp tương tự được nêu ở đây và thay thế hoặc bổ sung các trình tự nucleotit mã hóa (các synthaza thích hợp).

Thao tác di truyền các quá trình

Sáng chế sử dụng quá trình MEV và/hoặc DXP được thao tác di truyền để tác động đến việc sản xuất các isoprenoit ở mức cao trong tế bào chủ. Quá trình này thường được thao tác di truyền thông qua công nghệ ADN tái tổ hợp bằng cách biểu hiện các trình tự khác loại mã hóa các enzyme trong ít nhất một trong số các quá trình này.

Axit nucleotit của đối tượng có thể được biểu hiện bằng vector đơn hoặc đa. Ví dụ, vector biểu hiện đơn có thể bao gồm ít nhất hai, ba, bốn, năm, hoặc toàn bộ các trình tự khác loại mã hóa toàn bộ enzyme của quá trình DXP hoặc MEV. Mặc dù việc chọn lựa vector đơn hoặc đa có thể tùy thuộc vào kích thước của các trình tự khác loại và khả năng của vector, nhưng điều này sẽ phụ thuộc phần lớn vào tổng sản lượng của isoprenoit thu được mà vector này có khả năng tạo ra khi được biểu hiện trong tế bào chủ đã chọn. Vector của đối tượng có thể giữ được khả năng sao chép theo cách episom hoặc là phần không thể thiếu của hệ gen của tế bào chủ. Thông thường, đặc tính thứ hai này được ưu tiên cho sự nhân giống được duy trì liên tục của tế bào chủ.

Trong các tế bào chủ nhất định, một hoặc nhiều trình tự khác loại mã hóa enzym của quá trình MEV hoặc DXP có thể được kiểm soát bởi một hoặc nhiều operon. Trong một vài trường hợp, hai hoặc ba hệ operon tạo ra sản lượng isoprenoit cao hơn so với một hệ operon.

Khi cần, các trình tự axit nucleic của đối tượng có thể được cải biến để mang lại sự ưa thích codon của tế bào chủ đã được chọn để tác động đến sự biểu hiện cao hơn của các trình tự này trong tế bào chủ. Ví dụ, theo một số phương án, trình tự nucleotit của đối tượng sẽ được cải biến sự ưa thích codon của nấm men. Xem, ví dụ, ấn phẩm: Bennetzen and Hall (1982) *J. Biol. Chem* 257(6): 3026-3031. Ví dụ không giới hạn khác là, theo các phương án khác, các trình tự nucleotit sẽ được cải biến sự ưa thích codon của *E. coli*. Xem, ví dụ, ấn phẩm: Gouy and Gautier (1982) *Nucleic Acids Res.* 10(22) 7055-7074, Eyre-Walker (1996) *Mol. Biol. Evol.* 13(6):864-872. Xem ấn phẩm: Nakamura et al. (2000) *Nucleic Acids Res.* 28(1) 292. Bảng sử dụng codon dùng cho nhiều sinh vật là có sẵn, bảng này có thể được dùng để tham khảo trong việc thiết kế các trình tự theo sáng chế. Việc sử dụng các codon phổ biến của các vi sinh vật chủ đã cho thường làm tăng khả năng xảy ra dịch mã, và do đó làm tăng mức biểu hiện của các trình tự mong muốn.

Việc sản xuất các axit nucleic của đối tượng có thể được tiến hành bằng nhiều kỹ thuật tái tổ hợp và quy trình tổng hợp thông thường khác nhau. Tóm lại, các axit nucleic của đối tượng có thể được sản xuất từ các đoạn ADN, cADN, và ARN của hệ gen, tất cả các loại này có thể được tách chiết trực tiếp từ tế bào hoặc được sản xuất bằng cách tái tổ hợp bởi các quy trình khuếch đại khác nhau bao gồm, nhưng không bị giới hạn ở PCR và rt-PCR.

Việc tổng hợp axit nucleic trực tiếp bằng quá trình hoá học thường bao gồm bước bổ sung theo thứ tự các monome nucleotit đầu 3' bị chặn và đầu 5' bị chặn vào nhóm 5'-hydroxyl tận cùng của chuỗi polyme nucleotit đang lớn dần lên, trong đó mỗi lần bổ sung được tác động bởi sự tấn công ưa nhân của nhóm 5'-hydroxyl tận cùng của chuỗi đang lớn dần lên này lên vị trí 3' của monome đã được bổ sung, monome này thường là dẫn xuất của phospho, như

phosphotrieste, phosphoramidit, hoặc các chất tương tự. Hệ phương pháp này là đã biết với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này và được mô tả trong các sách và tài liệu thích hợp (ví dụ, xem ấn phẩm: Matteuci et al. (1980) Tet. Lett. 521:719, Patent Mỹ số 4,500,707 cấp cho Caruthers et al., và Patent Mỹ số 5,436,327 và 5,700,637 cấp cho Southern et. al.).

Có thể làm tăng mức phiên mã của axit nucleic trong vi sinh vật chủ bằng nhiều cách. Ví dụ, điều này có thể đạt được bằng cách làm tăng số bản sao của trình tự nucleotit mã hóa enzym (ví dụ, bằng cách sử dụng vector biểu hiện số bản sao cao hơn chứa trình tự nucleotit mã hóa enzym này, hoặc bằng cách đưa các bản sao bổ sung của trình tự nucleotit mã hóa enzym vào hệ gen của vi sinh vật chủ, ví dụ, bằng cách tái tổ hợp được điều hòa bởi recA, sử dụng vector "tự sát", tái tổ hợp sử dụng recombinaza của thể thực khuẩn lambda, và/hoặc chèn qua yếu tố gen nhảy (transposon) hoặc yếu tố có thể chuyển vị). Ngoài ra, điều này có thể được thực hiện bằng cách thay đổi trật tự của các vùng mã hóa trên mRNA đa cistron của operon hoặc phá vỡ operon thành các gen riêng biệt, mỗi gen có yếu tố điều hòa riêng của nó, hoặc làm tăng khả năng của yếu tố khởi đầu (trình tự kiểm soát phiên mã hoặc trình tự khởi đầu phiên mã) mà vùng mã hóa enzym được liên kết một cách linh hoạt với yếu tố khởi đầu này (ví dụ, sử dụng yếu tố khởi đầu cảm ứng arabinoza hoặc lactoza liên ứng trong vi sinh vật chủ *Escherichia coli* thay cho yếu tố khởi đầu cảm ứng lactoza đã được cải biến, như yếu tố khởi đầu được tìm thấy trong các plasmit pBluescript và pBBR1MCS), hoặc sử dụng yếu tố khởi đầu cảm ứng và cảm ứng yếu tố khởi đầu cảm ứng này bằng cách bổ sung hoá chất vào môi trường nuôi cấy. Mức dịch mã của trình tự nucleotit trong vi sinh vật chủ có thể được gia tăng theo nhiều cách, bao gồm, nhưng không bị giới hạn ở, tăng độ ổn định của mARN, cải biến trình tự của vị trí liên kết ribosom, cải biến khoảng cách hoặc trình tự giữa vị trí liên kết ribosom và codon khởi đầu của trình tự mã hóa enzym, cải biến toàn bộ vùng giữa các gen được định vị ở “ngược dòng của” hoặc liền kề với phía 5' của codon khởi đầu của vùng mã hóa enzym, làm ổn định đầu 3' của sản phẩm phiên mã mARN sử dụng các trình tự hình cặp tóc và các trình tự chuyên hoá, cải biến việc sử dụng codon của enzym, cải biến sự

biểu hiện của codon tARN hiếm, được sử dụng trong quá trình sinh tổng hợp enzym, và/hoặc làm tăng độ ổn định của enzym, như, ví dụ, thông qua đột biến trình tự mã hóa của nó. Việc xác định codon được ưu tiên và codon tRNA hiếm có thể được dựa trên việc phân tích trình tự các gen thu được từ vi sinh vật chủ.

Hoạt tính của MEV, DXP, hoặc prenyltransferaza trong vật chủ có thể được cải biến bằng nhiều cách, bao gồm, nhưng không bị giới hạn ở, biểu hiện dạng đã được cải biến của enzym thể hiện độ hoà tan tăng lên trong tế bào chủ, biểu hiện dạng cải biến của enzym thiếu vùng mà thông qua đó hoạt tính của enzym bị ức chế, biểu hiện dạng đã được cải biến của enzym mà có K_{cat} cao hơn hoặc K_m thấp hơn đối với cơ chất, hoặc biểu hiện dạng cải biến của enzym mà không bị ảnh hưởng bởi sự điều hòa ngược hoặc xuôi bởi một phân tử khác trong quá trình chuyển hoá. Các enzym biến thể như vậy còn có thể được phân lập thông qua sự phát sinh đột biến ngẫu nhiên của enzym đặc hiệu rộng hơn, như được mô tả dưới đây, và trình tự nucleotit mã hóa enzym biến thể này có thể được biểu hiện từ vectơ biểu hiện hoặc từ gen tái tổ hợp được kết hợp vào hệ gen của vi sinh vật chủ.

Vectơ của đối tượng có thể được tạo cấu trúc để sản xuất ra số lượng bản sao của enzym được mã hoá với mức mong muốn. Theo một số phương án, vectơ của đối tượng sản xuất ra ít nhất 10, từ 10 đến 20, từ 20 đến 50, từ 50 đến 100, hoặc thậm chí nhiều hơn 100 bản sao của HMG-CoA reductaza, mevalonat kinaza, hoặc cả hai. Các plasmit có số lượng bản sao thấp thường tạo ra thấp hơn khoảng 20 bản sao plasmit cho một tế bào; các plasmit có số lượng bản sao trung bình thường tạo ra từ khoảng 20 bản sao plasmit cho một tế bào đến khoảng 50 bản sao plasmit cho một tế bào, hoặc từ khoảng 20 bản sao plasmit cho một tế bào đến khoảng 80 bản sao plasmit cho một tế bào; và plasmit có số lượng bản sao cao thường tạo ra từ khoảng 80 bản sao plasmit cho một tế bào đến khoảng 200 bản sao plasmit cho một tế bào, hoặc nhiều hơn.

Vectơ biểu hiện lượng bản sao thấp thích hợp đối với *Escherichia coli* bao gồm, nhưng không bị giới hạn ở, pACYC184, pBeloBac11, pBR332,

pBAD33, pBBR1MCS và các dẫn xuất của nó, pSC101, SuperCos (cosmid), và pWE15 (cosmid). Vector biểu hiện lượng bản sao trung bình thích hợp đối với *Escherichia coli* bao gồm, nhưng không bị giới hạn ở pTrc99A, pBAD24, và vector chứa điểm khởi đầu sao chép ColEI và các dẫn xuất của nó. Vector biểu hiện lượng bản sao cao thích hợp đối với *Escherichia coli* bao gồm, nhưng không bị giới hạn ở, vector pUC, pBluescript, pGEM, và pTZ. Vector biểu hiện (thuộc tâm động) lượng bản sao thấp thích hợp đối với nấm men bao gồm, nhưng không bị giới hạn ở, pRS415 và pRS416 (xem ấn phẩm: Sikorski & Hieter (1989) *Genetics* 122:19-27). Vector biểu hiện 2 micron lượng bản sao cao thích hợp ở nấm men bao gồm, nhưng không bị giới hạn ở, pRS425 và pRS426 (xem ấn phẩm: Christainson et al. (1992) *Gene* 110:119-122). Vector biểu hiện 2 micron khác bao gồm các biến thể không có tính chọn lọc của vector 2 micron này (xem ấn phẩm: Bruschi & Ludwig (1988) *Curr. Genet* 15:83-90) hoặc plasmit 2 micron còn nguyên vẹn mang catxet biểu hiện (như được minh họa trong đơn xin cấp patent Mỹ số 20050084972) hoặc plasmit 2 micron mang gen chỉ thị chọn lọc khuyết như LEU2d (xem ấn phẩm: Erhanrt et al. (1983) *J. Bacteriol.* 156 (2): 625-635) hoặc URA3d (Okkels (1996) *Annals of New York Academy of Sciences* 782(1): 202- 207).

Các yếu tố điều hòa bao gồm, ví dụ, các yếu tố khởi đầu và các yếu tố điều khiển còn có thể được thao tác để làm tăng dòng chuyển hóa của quá trình MEV hoặc DXP bằng cách tăng sự biểu hiện của một hoặc nhiều gen đóng vai trò quan trọng trong việc quyết định sản lượng tổng của isoprenoit được sản xuất. Yếu tố khởi đầu là các trình tự nucleotit khởi đầu và kiểm soát sự phiên mã của trình tự axit nucleic bởi enzym ARN polymeaza. Yếu tố điều khiển là trình tự nucleotit liên kề với yếu tố khởi đầu thực hiện chức năng kiểm soát phiên mã của trình tự axit nucleic mong muốn. Yếu tố điều khiển này chứa vùng liên kết protein mà tại đó protein ức chế đặc hiệu có thể liên kết. Trong điều kiện không có protein ức chế thích hợp, sự phiên mã bắt đầu nhờ yếu tố khởi đầu. Trong điều kiện có mặt protein ức chế thích hợp, protein ức chế này liên kết với yếu tố điều khiển và bằng cách này ức chế phiên mã từ yếu tố khởi

đầu. Các yếu tố khởi đầu và các yếu tố điều khiển còn được gọi là các yếu tố điều hòa phiên mã.

Theo một số phương án thực hiện sáng chế, các yếu tố khởi đầu được sử dụng trong các vector biểu hiện là có khả năng cảm ứng. Theo phương án khác, các yếu tố khởi đầu được sử dụng trong vector biểu hiện là có khả năng cấu trúc. Theo một số phương án, một hoặc nhiều trình tự axit nucleic được liên kết một cách linh hoạt với yếu tố khởi đầu cảm ứng, và một hoặc nhiều trình tự axit nucleic khác được liên kết một cách linh hoạt với yếu tố khởi đầu cấu trúc.

Ví dụ không giới hạn về các yếu tố khởi đầu thích hợp để sử dụng trong tế bào chủ nhân sơ bao gồm yếu tố khởi đầu ARN polymeaza của thể thực khuẩn T7; yếu tố khởi đầu *trp*; yếu tố khởi đầu *lac operon*, yếu tố khởi đầu lai, ví dụ, yếu tố khởi đầu lai *lac/tac*, yếu tố khởi đầu lai *tac/trc*, yếu tố khởi đầu *trp/lac*, yếu tố khởi đầu T7/*lac*, yếu tố khởi đầu *trc*, yếu tố khởi đầu *tac*, và các yếu tố khởi đầu tương tự, yếu tố khởi đầu *araBAD*; yếu tố khởi đầu được điều hòa *in vivo*, như yếu tố khởi đầu *ssaG* hoặc yếu tố khởi đầu có liên quan (xem, ví dụ, Công bố Patent Mỹ số 20040131637), yếu tố khởi đầu *pagC* (xem ấn phẩm: Pulkkinen and Miller, J. Bacteriol. (1991) 173(1):86-93; xem ấn phẩm: Alpuche-Aranda et al. (1992) Proc. Natl. Acad. Sci. U S A. 89(21):10079-83), yếu tố khởi đầu *nirB* (xem ấn phẩm: Harborne et al. (1992) Mol Micro. 6:2805-2813), và các yếu tố khởi đầu tương tự (xem, ví dụ, xem ấn phẩm: Dunstan et al. (1999) Infect. Immun. 67,5133-5141; ấn phẩm: McKelvie et al. (2004) Vaccine 22:3243-3255; và xem ấn phẩm: Chatfield et al. (1992) Biotechnol. 10:888-892); yếu tố khởi đầu sigma70, ví dụ, yếu tố khởi đầu sigma70 liên ứng (xem, ví dụ, các số đăng ký ngân hàng gen AX798980, AX798961, và AX798183); yếu tố khởi đầu pha ổn định, ví dụ, yếu tố khởi đầu *dps*, yếu tố khởi đầu *spv*, và các yếu tố khởi đầu tương tự, yếu tố khởi đầu được tạo ra từ đảo sinh bệnh SPI-2 (xem, ví dụ, WO96/17951); yếu tố khởi đầu *actA* (xem, ví dụ, ấn phẩm: Shetron-Rama et al. (2002) Infect. Immun. 70:1087-1096); yếu tố khởi đầu *rpsM* (xem, ví dụ, ấn phẩm: Valdivia and Falkow (1996) Mol. Microbiol 22:367-378); yếu tố khởi đầu *tet* (xem, ví dụ, ấn phẩm: Hillen et al.

(1989) In Saenger W. and Heinemann U. (eds) Topics in Molecular and Structural Biology, Protein-Nucleic Acid Interaction. Macmillan, London, UK, Vol. 10, pp. 143-162); yếu tố khởi đầu SP6 (xem, ví dụ, ấn phẩm: Melton et al. (1984) Nucl. Acids Res. 12:7035-7056), và các yếu tố khởi đầu tương tự.

Theo một số phương án, tổng hoạt tính của enzym MEV hoặc DXP khác loại đóng vai trò lớn hơn trong tổng sản lượng isoprenoit so với các enzym khác trong các quá trình chuyển hoá tương ứng được tăng lên bằng cách biểu hiện enzym này từ yếu tố khởi đầu mạnh. Yếu tố khởi đầu mạnh thích hợp đối với *Escherichia coli* bao gồm, nhưng không bị giới hạn ở Trc, Tac, T5, T7, và P_{Lambda}. Theo một phương án khác, hoạt tính tổng của một hoặc nhiều enzym của quá trình MEV trong vật chủ được tăng lên bằng cách biểu hiện enzym này từ yếu tố khởi đầu mạnh trên plasmid có số lượng bản sao cao. Các ví dụ thích hợp, đối với *Escherichia coli* bao gồm, nhưng không bị giới hạn ở việc sử dụng yếu tố khởi đầu Trc, Tac, T5, T7, và P_{Lambda} với vector pBAD24, pBAD18, pGEM, pBluescript, pUC, và pTZ.

Ví dụ không giới hạn về yếu tố khởi đầu thích hợp để sử dụng trong tế bào chủ nhân chuẩn bao gồm, nhưng không bị giới hạn ở, yếu tố khởi đầu ngay sớm CMV, yếu tố khởi đầu HSV thymidin kinaza, yếu tố khởi đầu SV40 sớm hoặc muộn, LTR từ retrovirus, và yếu tố khởi đầu metallothionein-I ở chuột.

Ví dụ không giới hạn về yếu tố khởi đầu có cấu trúc thích hợp để sử dụng trong tế bào chủ nhân sơ bao gồm yếu tố khởi đầu sigma70 (ví dụ, yếu tố khởi đầu sigma70 liên ứng). Ví dụ không giới hạn về yếu tố khởi đầu cảm ứng thích hợp để sử dụng trong tế bào chủ vi khuẩn bao gồm pL của thể thực khuẩn λ ; Plac; Ptrp; Ptac (yếu tố khởi đầu lai Ptrp-lac); yếu tố khởi đầu cảm ứng isopropyl-beta-D44 thiogalactopyranosit (IPTG), ví dụ, yếu tố khởi đầu *lacZ*; yếu tố khởi đầu cảm ứng tetracyclin; yếu tố khởi đầu cảm ứng arabinosa, ví dụ, PBAD (xem, ví dụ, ấn phẩm: Guzman et al. (1995) J. Bacteriol. 177:4121-4130); yếu tố khởi đầu cảm ứng xyloza, ví dụ, Pxy1 (xem, ví dụ, ấn phẩm: Kim et al. (1996) Gene 181:71-76); yếu tố khởi đầu GAL1; yếu tố khởi đầu tryptophan; yếu tố khởi đầu lac; yếu tố khởi đầu cảm ứng rượu, ví dụ, yếu tố

khởi đầu cảm ứng metanol, yếu tố khởi đầu cảm ứng etanol; yếu tố khởi đầu cảm ứng rafinoza; yếu tố khởi đầu cảm ứng nhiệt, ví dụ, yếu tố khởi đầu lambda PL có khả năng cảm ứng nhiệt; yếu tố khởi đầu được kiểm soát bởi gen ức chế nhạy cảm nhiệt (ví dụ, vectơ biểu hiện dựa trên lambda ức chế CI857; xem, ví dụ, ấn phẩm: Hoffmann et al. (1999) *FEMS Microbiol Lett.* 177(2):327-34); và các yếu tố khởi đầu tương tự.

Ví dụ không giới hạn về yếu tố khởi đầu cấu trúc thích hợp để sử dụng ở tế bào chủ nấm men bao gồm yếu tố khởi đầu ADH1, ADH2, PGK, hoặc LEU2. Ví dụ không giới hạn về yếu tố khởi đầu cảm ứng thích hợp để sử dụng ở tế bào chủ nấm men bao gồm, nhưng không bị giới hạn ở, yếu tố khởi đầu cảm ứng galactosa phân hướng như yếu tố khởi đầu GAL 1 hoặc GAL 10 (Xem ấn phẩm: West et al. (1984) *Mol. Cell. Biol.* 4(11):2467-2478), hoặc yếu tố khởi đầu CUP1. Khi cần, vectơ của đối tượng chứa yếu tố khởi đầu mạnh hơn yếu tố khởi đầu Lac bẩm sinh của *E. Coli*.

Ví dụ không giới hạn về các yếu tố điều khiển để sử dụng trong tế bào chủ vi khuẩn bao gồm yếu tố điều khiển của yếu tố khởi đầu lactoza (protein ức chế LacI thay đổi hình dạng khi tiếp xúc với lactoza, bằng cách này ngăn chặn protein ức chế LacI liên kết với yếu tố điều khiển này), yếu tố điều khiển của yếu tố khởi đầu tryptophan (khi phức hợp với tryptophan, protein ức chế TrpR có hình dạng liên kết với yếu tố điều khiển; trong điều kiện không có tryptophan, protein ức chế TrpR có hình dạng mà không liên kết với yếu tố điều khiển), và yếu tố điều khiển của yếu tố khởi đầu tac (xem, ví dụ, ấn phẩm: deBoer et al. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25.).

Các gen trong vectơ biểu hiện thường cũng sẽ mã hoá vị trí liên kết ribosom để hướng sự dịch mã (tức là, tổng hợp) của sản phẩm gen mRNA đã được mã hoá bất kỳ. Đối với các vị trí liên kết ribosom thích hợp để sử dụng trong *Escherichia coli*, xem ấn phẩm: Shine et al. (1975) *Nature* 254:34, và ấn phẩm: Steitz, in *Biological Regulation and Development: Gene Expression* (ed. R. F. Goldberger), vol. 1, p. 349, 1979, Plenum Publishing, N. Y. Insertion of the ribosome binding site encoding nucleotide sequence 5'-AAAACA-3'

upstream of a coding sequence facilitates efficient translation in a yeast host microorganism (ấn phẩm: Looman et al. (1993) Nuc. Ac. Res. 21:4268-4271; ấn phẩm: Yun et. al. (1996) Mol. Microbiol. 19:1225-1239).

Các yếu tố điều hòa khác có thể được sử dụng trong vector biểu hiện bao gồm các yếu tố tăng cường phiên mã và yếu tố kết thúc phiên mã. Xem, ví dụ, Bitter et al. (1987) Methods in Enzymology, 153:516-544.

Vector biểu hiện có thể thích hợp để sử dụng trong các loại vi sinh vật chủ cụ thể và không thích hợp để sử dụng trong các loại vi sinh vật chủ khác. Tuy nhiên, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, có thể xác định một cách dễ dàng bằng thử nghiệm thông thường xem vector biểu hiện cụ thể đó có thích hợp cho vi sinh vật chủ đã cho hay không. Ví dụ, vector biểu hiện có thể được đưa vào sinh vật chủ, sau đó được kiểm tra về khả năng tồn tại và biểu hiện của các gen bất kỳ chứa trong vector này.

Vector biểu hiện này có thể còn chứa một hoặc nhiều gen chỉ thị có khả năng chọn lọc mà, trong quá trình biểu hiện, tạo ra một hoặc nhiều tính trạng kiểu hình hữu dụng cho việc chọn lọc hoặc nhận biết tế bào chủ mang vector biểu hiện này. Ví dụ không giới hạn về gen chỉ thị thích hợp có khả năng chọn lọc đối với tế bào nhân chuẩn bao gồm dihydrofolat reductaza và kháng neomycin. Ví dụ không giới hạn về gen chỉ thị có khả năng chọn lọc thích hợp đối với tế bào nhân sơ bao gồm gen kháng tetracyclin, ampicillin, chloramphenicol, carbenicillin, và kanamycin.

Để sản xuất isoprenoid trên quy mô công nghiệp, có thể không thực tế hoặc quá tốn kém nếu sử dụng gen chỉ thị có khả năng chọn lọc mà cần phải bổ sung chất kháng sinh vào môi trường lên men. Do đó, một số phương án của sáng chế sử dụng tế bào chủ mà không cần phải sử dụng gen chỉ thị có khả năng chọn lọc tạo ra tính kháng chất kháng sinh để đảm bảo sự duy trì plasmit (vector biểu hiện). Theo các phương án của sáng chế, vector biểu hiện chứa hệ duy trì plasmit như plasmit IncP (RK2) 60kb, tùy ý cùng với hệ sao chép và/hoặc phân tách RK2 plasmit, để tác động đến sự duy trì plasmit trong điều kiện không có sự chọn lọc chất kháng sinh (xem, ví dụ, ấn phẩm: Sia et al.

(1995) J. Bacteriol. 177:2789-97; ấn phẩm: Pansegrau et al. (1994) J. Mol. Biol. 239:623-63). Hệ duy trì plasmid thích hợp cho mục đích này được mã hoá bởi *parDE* operon của RK2, mà mã hóa độc tố ổn định và kháng độc tố không ổn định. Kháng độc tố này có thể ức chế tác động gây chết của độc tố nhờ sự tương tác protein-protein một cách trực tiếp. Tế bào mất vector biểu hiện chứa operon *parDE* nhanh chóng bị thiếu kháng độc tố không ổn định, tạo ra độc tố ổn định sau đó làm chết tế bào. Hệ sao chép RK2 plasmid được mã hoá bởi gen *trfA*, mã hóa protein sao chép ADN. Hệ phân tách RK2 plasmid được mã hoá bởi operon *parCBA*, mã hóa các protein thực hiện chức năng tách các multimer plasmid có thể xuất hiện từ quá trình sao chép ADN.

Các vector của đối tượng có thể được đưa vào tế bào chủ một cách ổn định hoặc tạm thời bằng nhiều kỹ thuật đã được thiết lập. Ví dụ, phương pháp liên quan đến việc xử lý bằng canxi clorua trong đó vector biểu hiện được đưa vào thông qua sự kết tủa canxi. Các muối khác, ví dụ canxi phosphat, cũng có thể được sử dụng theo quy trình tương tự. Ngoài ra, có thể sử dụng phương pháp dòng điện dịch chuyển (tức là, sử dụng dòng điện để tăng khả năng thâm nhập của tế bào với axit nucleic). Các phương pháp biến nạp khác bao gồm vi tiêm, biến nạp qua trung gian DEAE dextran, và sốc nhiệt trong điều kiện có mặt lithi axetat. Phức hợp lipid, liposom, và dendrime cũng có thể được sử dụng để chuyển nhiễm vào vi sinh vật chủ.

Trong quá trình biến nạp, có thể thực hiện nhiều phương pháp để nhận dạng tế bào chủ mà vector của đối tượng đã được đưa vào đó. Một phương pháp chọn lọc làm ví dụ bao gồm cấy truyền các tế bào riêng biệt để tạo thành các khuẩn lạc riêng biệt, sau đó kiểm tra sự biểu hiện của sản phẩm gen mong muốn. Phương pháp khác đòi hỏi việc chọn lọc các tế bào chủ đã được biến nạp dựa trên tính trạng kiểu hình được tạo ra thông qua việc biểu hiện của gen chỉ thị có khả năng chọn lọc được chứa trong vector biểu hiện. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể nhận biết tế bào chủ đã được cải biến di truyền bằng cách sử dụng các phương pháp này hoặc các phương pháp khác đã biết trong lĩnh vực kỹ thuật này.

Việc đưa các trình tự của quá trình chuyển hóa khác nhau theo sáng chế vào tế bào chủ có thể được xác nhận bằng các phương pháp như PCR, phương pháp lai thẩm tách Southern blot hoặc Northern blot. Ví dụ, axit nucleic có thể được điều chế từ tế bào chủ thu được, và các trình tự đặc hiệu quan tâm có thể được khuếch đại bằng PCR trong đó, sử dụng các môi đặc hiệu đối với các trình tự quan tâm. Sản phẩm khuếch đại được tiến hành điện di gel agarosa, điện di gel polyacrylamit hoặc điện di mao dẫn, sau đó nhuộm bằng ethidium bromua, dung dịch xanh SYBR hoặc các chất tương tự, hoặc phát hiện ADN bằng cách dò UV. Theo cách khác, các mẫu dò axit nucleic đặc hiệu đối với các trình tự quan tâm có thể được sử dụng trong phản ứng lai. Sự biểu hiện của trình tự gen đặc hiệu có thể được xác định bằng cách phát hiện mRNA tương ứng thông qua PCR cặp đôi phiên mã ngược, phương pháp lai thẩm tách Northern blot, hoặc bằng thử nghiệm miễn dịch sử dụng các kháng thể phản ứng với sản phẩm gen đã được mã hoá. Các thử nghiệm miễn dịch làm ví dụ bao gồm, nhưng không bị giới hạn ở ELISA, thử nghiệm miễn dịch phóng xạ, và thử nghiệm miễn dịch bánh kẹp.

Hoạt tính enzym của enzym của quá trình chuyển hóa nhất định có thể được thử nghiệm bằng nhiều phương pháp đã biết trong lĩnh vực kỹ thuật này. Nói chung, hoạt tính enzym có thể được xác định bằng cách tạo ra sản phẩm hoặc chuyển hoá cơ chất của phản ứng enzym cần kiểm tra. Phản ứng này có thể xảy ra in vitro hoặc in vivo. Ví dụ, hoạt tính tương đối của HMG-CoA reductaza và HMG-CoA synthaza ở tế bào có thể được xác định bằng mức trạng thái cân bằng của HMG-CoA trong tế bào. HMG-CoA có thể được chiết bằng axit trichloroacetic (Trichloroacetic Acid - TCA), sau đó phân tích nguyên liệu đã được chiết bằng sắc ký lỏng/phổ khối. Hoạt tính của mevalonat kinaza có thể được minh chứng bằng sự tạo thành mevalonat 5-phosphat. Hoạt tính tương đối của mevalonat kinaza và HMG-CoA reductaza có thể được xác định bằng mức trạng thái cân bằng của mevalonat, mức này có thể được xác định bằng sắc ký khí/phổ khối. Xem ví dụ, WO05033287, được đưa vào đây bằng cách tham khảo.

Có thể làm tăng sản lượng isoprenoit qua một hoặc nhiều quá trình chuyển hoá được bộc lộ ở đây bằng cách ức chế phản ứng mà làm trệch hướng các sản phẩm trung gian từ các bước sản xuất tiến tới sự hình thành sản phẩm isoprenoit. Việc ức chế các phản ứng không sản xuất có thể đạt được bằng cách làm giảm sự biểu hiện và/hoặc hoạt tính của enzym liên quan đến một hoặc nhiều phản ứng không sản xuất. Các phản ứng này bao gồm phản ứng phụ của chu trình TCA dẫn đến sự sinh tổng hợp axit béo, sinh tổng hợp alanin, quá trình siêu chuyển hoá aspartat, sự hình thành glucoza trong cơ thể động vật, sinh tổng hợp hem, và/hoặc sinh tổng hợp glutamat, ở mức tác động đến tổng sản lượng của isoprenoit. Ngoài ra, việc chuyển hoá axetyl-CoA thành axetat nhờ tác động của phosphotransaxetylaza là ví dụ khác về phản ứng phụ không sản xuất. Vì vậy, khi cần, "bất hoạt" hoặc "kìm hãm" gen *pta* mã hóa phosphotransaxetylaza cũng có thể được thực hiện để làm tăng sản lượng isoprenoit. Tùy thuộc vào isoprenoit cụ thể quan tâm, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể chọn nhắm đích các bước không sản xuất bổ sung. Ví dụ, nếu carotenoit là isoprenoit được chọn lựa, thì chuyên gia có thể chọn "bất hoạt" hoặc "kìm hãm" một hoặc nhiều gen được chọn từ nhóm bao gồm gen *gdhA*, *aceE*, *fdhF*, *yjiD*, *hnr* hoặc *yjfP*, *ackA*, *appY*, *aspC*, *clp*, *clpP*, *clpXP*, *crcB*, *csdA*, *cyaA*, *evgS*, *fdhA*, *fdhD*, *feoB*, *fumA*, *glnE*, *gbxR*, *gntK*, *hycI*, *lipB*, *lysU*, *modA*, *moeA*, *nadA*, *nuoC*, *nuoK*, *pflB*, *pitA*, *pst*, *pstC*, *pta*, *p-yjiD*, *sohA*, *stpA*, *yagR*, *yaiD*, *ybaS*, *ycfZ*, *ydeN*, *yebB*, *yedN*, *yfcC*, *ygiP*, *yibD*, *yjfP*, *yjhH*, hoặc *yliE*, hoặc các gen bất kỳ khác một mình hoặc ở dạng kết hợp, việc ức chế các gen này sẽ tạo ra sản lượng carotenoit cao hơn như nêu trong đơn xin cấp patent Mỹ số 20060121558, được đưa vào đây để tham khảo.

Có nhiều phương pháp khác nhau đã biết nhằm bất hoạt hoặc kìm hãm gen quan tâm. Ví dụ, có thể làm giảm biểu hiện gen bằng cách xóa bỏ, đột biến, và/hoặc sắp xếp lại gen. Điều này còn có thể được thực hiện bằng cách sử dụng ARN đối nghĩa, siARN, miARN, các ribozym, ADN sợi ba, và chất ức chế phiên mã và/hoặc dịch mã. Ngoài ra, gen nhảy (transposon) có thể được sử dụng để phá vỡ sự biểu hiện gen, ví dụ, bằng cách cài gen nhảy này vào giữa

yếu tố khởi đầu và vùng mã hóa, hoặc giữa hai gen liền kề để bất hoạt một hoặc cả hai gen này.

Năng suất sản xuất hợp chất isoprenoit cao

Sáng chế đề xuất hợp phần và phương pháp sản xuất mạnh isoprenoit bằng cách sử dụng các enzym của quá trình isopentenyl pyrophosphat được kiểm soát bởi ít nhất một yếu tố điều hòa khác loại hoặc các điều kiện lên men, hoặc riêng rẽ hoặc ở dạng kết hợp.

Theo một khía cạnh, phương pháp sản xuất isoprenoit bao gồm các bước (a) thu nhiều tế bào chủ chứa quá trình enzym để tạo ra isopentenyl pyrophosphat trong đó toàn bộ các enzym của quá trình này được kiểm soát bởi ít nhất một yếu tố điều hòa phiên mã khác loại, và (b) nuôi cấy tế bào chủ này ở môi trường trong điều kiện là dưới điểm cực thuận so với điều kiện mà sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ này. Theo một số phương án, quá trình này là quá trình mevalonat. Theo phương án khác, quá trình này là quá trình DXP. Theo phương án khác, ít nhất một trình tự điều hòa phiên mã khác loại là có khả năng cảm ứng. Theo phương án khác, các enzym của quá trình này được kiểm soát bởi một yếu tố điều hòa phiên mã. Theo phương án khác, các enzym của quá trình này được kiểm soát bởi nhiều yếu tố điều hòa khác loại.

Theo một số phương án, quá trình này chứa trình tự axit nucleic mã hóa enzym của quá trình mevalonat từ sinh vật nhân sơ có quá trình mevalonat nội sinh. Ví dụ không giới hạn về sinh vật nhân sơ thích hợp bao gồm các sinh vật nhân sơ thuộc các chi: *Actinoplanes*, *Archaeoglobus*, *Bdellovibrio*, *Borrelia*; *Chloroflexus*, *Enterococcus*, *Lactobacillus*, *Listeria*, *Oceanobacillus*; *Paracoccus*; *Pseudomonas*, *Staphylococcus*; *Streptococcus*, *Streptomyces*, *Thermoplasma*; và *Vibrio*. Ví dụ không giới hạn về các chủng đặc trưng bao gồm: *Archaeoglobus fulgidus*; *Bdellovibrio bacteriovorus*; *Borrelia burgdorferi*; *Chloroflexus aurantiacus*; *Enterococcus faecalis*; *Enterococcus faecium*; *Lactobacillus johnsonii*; *Lactobacillus plantarum*; *Lactococcus lactis*, *Listeria innocua*; *Listeria monocytogenes*; *Oceanobacillus iheyensis*,

Paracoccus zeaxanthinifaciens; *Pseudomonas mevalonii*; *Staphylococcus aureus*; *Staphylococcus epidermidis*; *Staphylococcus haemolyticus*; *Streptococcus agalactiae*; *Streptomyces griseolosporeus*; *Streptococcus mutans*; *Streptococcus pneumoniae*, *Streptococcus pyogenes*; *Thermoplasma acidophilum*, *Thermoplasma volcanium*, *Vibrio cholerae*; *Vibrioparahaemolyticus*, và *Vibrio vulnificus*;

Theo một phương án khác, trình tự axit nucleic mã hóa enzym của quá trình mevalonat được chọn từ axetyl-CoA thiolaza, HMG-CoA synthaza, HMG-CoA reductaza, và mevalonat kinaza. Theo một phương án khác, trình tự axit nucleic mã hóa enzym của quá trình mevalonat được chọn từ axetyl-CoA thiolaza, HMG-CoA synthaza, HMG-CoA reductaza, và mevalonat kinaza và là từ sinh vật nhân sơ thuộc chi *Enterococcus* hoặc chi *Pseudomonas* hoặc chi *Staphylococcus*. Theo một phương án khác, trình tự axit nucleic mã hóa enzym của quá trình mevalonat được chọn từ axetyl-CoA thiolaza, HMG-CoA synthaza, HMG-CoA reductaza, và mevalonat kinaza và là từ *Enterococcus faecalis* hoặc từ *Staphylococcus aureus*.

Theo một phương án khác, trình tự axit nucleic mã hóa enzym của quá trình mevalonat là HMG-CoA reductaza lớp II. HMG-CoA reductaza thường được phân loại thành hai lớp, hai lớp này có thể phân biệt được dựa trên độ tương đồng trình tự và/hoặc đặc tính enzym (xem, ví dụ, ấn phẩm: Hedl, et al., J. Bacteriology, 1927-1932, 2004, và ấn phẩm: Bochar, et al., Molec. Genet. Metab, 66, 122-127, 1999).

HMG-CoA reductaza lớp II có thể được đặc trưng, một phần, bởi độ nhạy cảm của chúng đối với mức statin thấp, bao gồm, nhưng không bị giới hạn ở Atorvastatin, Cerivastatin, Fluvastatin, Lovastatin, Pravastatin, Simvastatin, theo một số phương án, HMG-CoA reductaza lớp II thể hiện hằng số ức chế statin lớn hơn khoảng 1 micromol, 10 micromol, hoặc 100 micromol. Theo phương án khác, HMG-CoA reductaza lớp II có hằng số ức chế đối với Lovastatin lớn hơn hằng số ức chế đối với Lovastatin của HMG-CoA lớp I bởi hệ số ít nhất khoảng 10, 100, 1000, hoặc 10.000. Theo phương án khác, HMG-

CoA reductaza lớp II có hằng số ức chế đối với Lovastatin lớn hơn hằng số ức chế đối với Lovastatin của HMG-CoA lớp I được phân lập từ *Homo sapien* bởi hệ số ít nhất khoảng 10, 100, 1000, hoặc 10.000. Theo phương án khác, HMG-CoA reductaza lớp II là từ sinh vật nhân sơ. Theo phương án khác, HMG-CoA reductaza lớp II là từ các vi khuẩn cổ khuẩn.

HMG-CoA reductaza lớp II nguyên thủy được tạo ra từ *Pseudomonas mevalonii*. Cũng được bao gồm theo sáng chế là HMG-CoA reductaza lớp II biến thể thể hiện ít nhất khoảng 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 80%, 90%, hoặc 95% độ tương đồng với trình tự axit amin của HMG-CoA reductaza của *P. mevalonii*. Cũng được bao gồm theo sáng chế là các biến thể có thấp hơn khoảng 40%, 35%, 30%, 25%, 20%, hoặc ít hơn độ tương đồng với HMG-CoA reductaza của *H. Sapiens*. Độ tương đồng trình tự axit amin có thể được xác định bằng phương pháp được mô tả trong ấn phẩm: Bochar, et al., Molec Genet. Metab., 66, 122-127, 1999.

Các ví dụ không giới hạn về HMG-CoA reductaza lớp II bao gồm các HMG-CoA reductaza lớp II được tạo ra từ: HMG-CoA reductaza từ: *Archaeoglobus fulgidus* (NC_000917), *Bdellovibrio bacteriovorus* (BX842650); *Borrelia burgdorferi* (AE001169); *Chloroflexus aurantiacus* (AJ299212); *Enterococcus faecalis* (AAO81155); *Enterococcus faecium* (AF290094); *Lactobacillus johnsonii* (AE017204); *Lactobacillus plantarum*; *Lactococcus lactis* (AE006387); *Listeria innocua* (CAC96053); *Listeria monocytogenes* (AE017324); *Oceanobacillus iheyensis* (NC_000917), *Paracoccus zeaxanthinifaciens* (AJ431696); *Pseudomonas mevalonii* (M24015); *Staphylococcus aureus* (AF290086); *Staphylococcus epidermidis* (AF290090); *Staphylococcus haemolyticus* (AF290088); *Streptococcus agalactiae* (CAD47046); *Streptomyces griseolosporeus* (AB037907); *Streptococcus mutans* (AAN58647); *Streptococcus pneumoniae* (AF290098); *Streptococcus pyogenes* (AF290096); *Thermoplasma acidophilum* (CAC11548), *Thermoplasma volcanium* (AL935253); *Vibrio cholerae* (AAF96622); *Vibrio parahaemolyticus* (BAC62311); và *Vibrio vulnificus* (AAO07090).

Các phương pháp lên men được nêu ở đây có liên quan đến việc điều biến tốc độ sinh trưởng riêng của tế bào chủ. Thường được đặc trưng bởi tham số μ , tốc độ sinh trưởng riêng biểu thị cho tốc độ sinh trưởng của tế bào trên một đơn vị sinh khối trong một đơn vị thời gian. Tốc độ sinh trưởng riêng có các đơn vị thời gian nghịch đảo ($1/t$). Tốc độ sinh trưởng riêng tối đa đối với các tế bào trong môi trường nuôi cấy có liên quan đến sự ảnh hưởng của nồng độ cơ chất đến tốc độ sinh trưởng. Nói chung, tế bào sẽ sinh trưởng chậm ở mức cơ chất thấp, và khi mức cơ chất trong môi trường gia tăng, thì tốc độ sinh trưởng của tế bào cũng gia tăng. Tuy nhiên, tốc độ sinh trưởng tế bào không tiếp tục tăng một cách vô hạn, và ở mức cơ chất cao, việc tăng lượng cơ chất đã cho sẽ làm tăng càng ngày càng ít tốc độ sinh trưởng của tế bào. Vì vậy, tốc độ sinh trưởng cuối cùng sẽ tiến đến giới hạn, giới hạn này thường được dùng để chỉ tốc độ sinh trưởng riêng tối đa. Việc nghiên cứu về mặt lý thuyết mối tương quan giữa tốc độ sinh trưởng trong môi trường nuôi cấy là đã biết đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, và được gọi là phương trình Monod. Xem, ví dụ, ấn phẩm: Pirt, Principles of Microbe and Cell Cultivation, Wiley, NY, 1975, các trang 4-10. Trong nghiên cứu về mặt lý thuyết này, tốc độ riêng tối đa là giới hạn tiệm cận mà không bao giờ đạt tới cho đến khi mức cơ chất vô hạn đạt tới. Tuy nhiên, trên thực tế, tốc độ sinh trưởng riêng tối đa có thể được xem là đạt được khi điều kiện đang nghiên cứu (ví dụ, mức cơ chất hoặc nhiệt độ) hỗ trợ tốc độ sinh trưởng khởi đầu nhanh nhất. Ví dụ, trong thiết bị phản ứng nạp liệu theo mẻ, điều kiện ban đầu trong đó các chất dinh dưỡng được cung cấp dư thừa được coi là điều kiện cho tốc độ sinh trưởng tối đa. Xem, ví dụ, ấn phẩm: Lee et al. (1996) *Trends Biotechnol.* 14: 98-105 và Korz et al. (1995) *J Biotechnology* 39 59-65. Điều kiện trong đó cơ chất được bổ sung để hỗ trợ tốc độ sinh trưởng riêng tối đa đôi khi còn được dùng để chỉ sự sinh trưởng không giới hạn. Ngoài ra, trong khi phương trình Monod mô tả các tính chất về tốc độ lý thuyết đối với cơ chất mà tiến gần một cách tiệm cận đến tốc độ riêng tối đa, đối với nhiều cơ chất, đúng hơn là tiến đến giá trị khi nhiều cơ chất hơn được bổ sung, quan sát thấy sự giảm tốc độ ở mức cơ chất cao hơn

sau khi đạt được tốc độ tối đa, tức là, đạt được tốc độ sinh trưởng riêng tối đa sau đó là giảm tốc độ sinh trưởng.

Tốc độ sinh trưởng riêng tối đa còn có thể được áp dụng đối với nhiệt độ cũng như đối với cơ chất. Nói chung, sinh vật sẽ sinh trưởng chậm ở nhiệt độ thấp, và sẽ sinh trưởng với tốc độ nhanh hơn khi nhiệt độ tăng đến một điểm nhất định, sau điểm này, tốc độ sinh trưởng sẽ giảm xuống. Sẽ có một giá trị nhiệt độ mà ở đó tốc độ sinh trưởng sẽ đạt mức tối đa, đây là nhiệt độ ở đó đạt được tốc độ sinh trưởng tối đa. Tác giả sáng chế đã phát hiện ra rằng có thể làm tăng việc sản xuất isoprenoit bằng cách hạ nhiệt độ xuống thấp hơn nhiệt độ mà hỗ trợ tốc độ sinh trưởng riêng tối đa.

Tốc độ sinh trưởng riêng tối đa còn có thể được áp dụng đối với các chất phụ gia khác đối với quá trình lên men ngoài cơ chất. Ví dụ, đối với chất dinh dưỡng, vitamin, và khoáng chất, có thể có tốc độ thấp ở hàm lượng thấp của các thành phần này, tốc độ này sẽ đạt cao hơn khi nồng độ của các thành phần này tăng, sau đó, trong một số trường hợp, ngay cả ở nồng độ cao hơn của các thành phần này, thì tốc độ sẽ giảm xuống. Tốc độ sinh trưởng riêng tối đa đạt được khi nồng độ của các thành phần hỗ trợ tốc độ cao nhất.

Tốc độ sinh trưởng riêng tối đa của tế bào trong môi trường thường được xác định ở các giai đoạn ban đầu của quá trình lên men trước khi bị ức chế bởi sản phẩm cuối hoặc các sản phẩm trung gian, mật độ tế bào, hoặc các yếu tố khác góp phần làm giảm tốc độ sinh trưởng. Ví dụ tốc độ sinh trưởng tối đa thường được xác định trong pha sinh trưởng bùng nổ ngoài pha lag, pha giảm tốc độ, hoặc pha ổn định. Khái niệm tốc độ sinh trưởng riêng tối đa còn có thể được áp dụng ở các giai đoạn sau của quá trình lên men bằng cách tính đến các biến số thích hợp.

Như vậy, theo một số phương án, các tế bào chủ được nuôi cấy trong điều kiện sao cho tốc độ sinh trưởng thấp hơn khoảng 90% tốc độ sinh trưởng riêng tối đa. Theo phương án khác, các tế bào chủ được nuôi cấy trong điều kiện sao cho tốc độ sinh trưởng thấp hơn khoảng 80%, 75%, 60%, 50%, 40%,

30%, 25%, 20%, 10%, 5%, hoặc 1%, hoặc thấp hơn tốc độ sinh trưởng riêng tối đa.

Theo phương án khác, các tế bào chủ được nuôi cấy ở nhiệt độ môi trường thấp hơn ít nhất là khoảng 2°C, 4°C, 5°C, 6°C, 8°C, 10°C, 15°C, hoặc 20°C so với nhiệt độ mà sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Bằng cách hạ thấp nhiệt độ thì sự sinh trưởng sẽ giảm, điều này, lại làm giảm sự hình thành sản phẩm phụ gây độc trong môi trường và sinh ra nhiệt trao đổi chất. Việc hạ thấp nhiệt độ môi trường nuôi cấy còn làm giảm nhu cầu oxy của tế bào làm cho có thể thu được mật độ tế bào cao hơn.

Nhiệt độ ở đó có thể đạt được tốc độ sinh trưởng riêng tối đa của tế bào chủ sẽ tùy thuộc vào loại tế bào chủ được chọn. Có thể xác định điều này bằng cách nuôi dưỡng tế bào chủ trong các điều kiện nhiệt độ khác nhau trong một khoảng thời gian xác định để thu được đường cong sinh trưởng tương ứng. Nhiệt độ hỗ trợ tốc độ sinh trưởng riêng tối đa có thể được xác định bằng cách so sánh độ dốc sinh trưởng ở đường cong tương ứng này. Trong trường hợp của *E. Coli*, nhiệt độ cho tốc độ sinh trưởng riêng tối đa là khoảng 37°C. Như vậy, nếu *E. Coli* là tế bào chủ được dùng để sản xuất isoprenoit bằng cách lên men, thì nhiệt độ của quá trình lên men là dưới 37°C. Nếu *S. cerevisiae* được sử dụng, thì nhiệt độ cho tốc độ sinh trưởng riêng tối đa là khoảng 30°C. Như vậy, nếu *S. cerevisiae* là tế bào chủ được dùng để sản xuất isoprenoit bằng cách lên men thì nhiệt độ lên men là dưới 30°C. Nói chung, nhiệt độ mong muốn là thấp hơn khoảng 2°C, 4°C, 5°C, 6, 8, 10°C, 15°C, và 20°C so với nhiệt độ mà ở đó có thể đạt được tốc độ sinh trưởng riêng tối đa của tế bào chủ.

Theo các phương án khác, tế bào chủ được nuôi cấy trong môi trường lên men chứa nguồn cacbon có mặt với lượng thấp hơn lượng sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Theo các phương án nhất định, các tế bào chủ được nuôi cấy trong môi trường trong đó nguồn cacbon được duy trì ở mức để tạo ra thấp hơn khoảng 90%, 80%, 75%, 60%, 50%, 40%, 30%, 25%, 20%, 10%, 5%, 1%, hoặc thấp hơn, tốc độ sinh trưởng riêng tối đa. Có thể sử dụng nguồn chứa cacbon bất kỳ mà vi sinh vật chuyển hoá được. Các ví dụ không giới hạn

bao gồm hydrat cacbon như monosacarit, oligosacarit và polysacarit, axit hữu cơ như axit axetic, axit propionic; và rượu như etanol và propanol, và polyol như glyxerol.

Theo một số phương án, nguồn cacbon chủ yếu bao gồm monosacarit hoặc oligosacarit. Theo phương án khác, nguồn cacbon về cơ bản bao gồm monosacarit và disacarit. Theo các phương án khác nữa, nguồn cacbon về cơ bản không chứa xenluloza.

Monosacarit là đường đơn đóng vai trò là các khối xây dựng của hydrat cacbon. Chúng được phân loại dựa trên bộ khung xương của nguyên tử cacbon (C) của chúng: trioza có ba nguyên tử cacbon, tetroza có bốn, pentoza có năm, hexoza có sáu, và heptoza có bảy nguyên tử cacbon. Các nguyên tử cacbon này được liên kết với các nguyên tử hydro (-H), các nhóm hydroxyl (-OH), và các nhóm cacbonyl (-C=O), các cách kết hợp, thứ tự, và cấu hình của chúng tạo ra lượng lớn chất đồng phân lập thể lớn. Các pentoza bao gồm xyloza, được tìm thấy trong nguyên liệu gỗ, arabinoza, được tìm thấy trong chất gôm từ cây lá kim; riboza, là thành phần của ARN và một vài vitamin, và deoxyriboza, là thành phần của ADN. Các hexoza làm ví dụ bao gồm glucoza, galactoza, và fructoza. Monosacarit kết hợp với nhau và với các nhóm khác để tạo ra nhiều disacarit, và oligosacarit. Các oligosacarit là polyme sacarit chứa lượng nhỏ (thường từ ba đến mười) đường đơn giản. Chúng thường được tìm thấy là được liên kết hoặc *O* hoặc *N* với chuỗi bên amino axit tương thích trong các gốc protein hoặc lipid. Oligosacarit được ưu tiên dùng trong phản ứng lên men hiện nay là disacarit, bao gồm ví dụ, sucroza, hoặc trisacarit như rafinoza.

Nếu muốn sử dụng xenluloza, glycan, tinh bột, hoặc các polysacarit khác để làm nguồn cacbon cơ bản, thì các polysacarit này đầu tiên có thể được chuyển hoá thành monosacarit và oligosacarit bằng quá trình hoá học hoặc bằng phương pháp enzym. Ví dụ, xenluloza có thể được chuyển hoá thành glucoza bởi enzym xenlulaza. Như vậy, nếu polysacarit như xenluloza được tìm thấy trong sinh khối (bao gồm ví dụ, cây cải dầu canola, cỏ linh lăng, cây lúa, cây lúa mạch đen, cây cao lương, cây hướng dương, cây lúa mì, cây đậu

tương, cây thuốc lá, cây khoai tây, cây lạc, cây bông, cây khoai lang, cây sắn, cây cà phê, cây dừa, cây có múi, cây ca cao, cây chè, quả như: quả chuối, quả sung, quả dứa, quả ổi, quả xoài, hạt yến mạch, hạt lúa mạch, cây rau, cây làm cảnh, hoặc cây lá kim) được dùng làm nguồn cacbon cơ bản, nó có thể được phân giải bằng xenulaza để tạo ra các đường đơn giản hơn dùng kết hợp với quy trình lên men theo sáng chế. Theo các phương án nhất định, sau sự phân hủy polysacarit, thì monosacarit và/hoặc oligosacarit cấu thành ít nhất khoảng 50% khối lượng nguồn cacbon như được xác định vào lúc bắt đầu quá trình lên men. Theo phương án khác, monosacarit và/hoặc oligosacarit cấu thành ít nhất khoảng 80% hoặc thậm chí 90% khối lượng nguồn cacbon như được xác định vào lúc bắt đầu quá trình lên men, để cho môi trường lên men về cơ bản không chứa xenluloza.

Theo phương án khác, tế bào chủ được nuôi cấy trong môi trường lên men chứa nguồn nitơ có mặt với lượng thấp hơn lượng mà sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Không bị ràng buộc bởi lý thuyết cụ thể bất kỳ, đã biết rằng việc thay đổi mức độ của các thành phần như nitơ mà có sẵn đối với tế bào có thể thay đổi dòng liên quan thông qua nhiều quá trình hoá học khác nhau trong tế bào. Các tác giả sáng chế đã phát hiện ra bằng cách hạn chế mức nitơ có sẵn đối với vi sinh vật, thì lượng isoprenoit như amorphen được tạo ra bởi vi sinh vật được tăng lên. Mức nitơ theo sáng chế làm ví dụ bao gồm lượng mà sẽ hỗ trợ khoảng 90%, 80%, 75%, 60%, 50%, 40%, 30%, 25%, 20%, 10%, 5%, 1%, hoặc thấp hơn tốc độ sinh trưởng riêng tối đa.

Việc hạn chế nitơ có thể được thực hiện ở nhiều giai đoạn. Theo một số phương án, nitơ ở dạng amoniac được cung cấp vào lúc khởi đầu quá trình lên men để hỗ trợ sự sinh trưởng ban đầu, tuy nhiên ở các bước bổ sung tiếp theo của quá trình lên men là không có nitơ, hoặc là không có nitơ để đảm bảo cho mức amoniac cần thiết để duy trì độ pH của quá trình lên men là 7 bằng dung dịch amoniac. Đối với phần chính của quá trình lên men, mức nitơ được duy trì ở mức thấp hơn lượng mà sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Các lượng như vậy có thể là lượng mà sẽ hỗ trợ ít nhất khoảng 90%, 80%, 75%, 50%,

40%, 30%, 25%, 20%, 10%, 5% hoặc 1%, hoặc thấp hơn tốc độ sinh trưởng riêng tối đa. Quá trình lên men theo sáng chế có mức nitơ ban đầu trên 10mM như được xác định trong môi trường lên men, để hỗ trợ sự sinh trưởng ban đầu, và mức nitơ tiếp theo trong môi trường lên men dưới 50mM, 40mM, 30mM, 20mM, 10mM, hoặc 4 mM.

Nguồn nitơ có thể đồng hoá mà có thể được sử dụng trong hỗn hợp phản ứng lên men thích hợp bao gồm, nhưng không bị giới hạn ở, các nguồn nitơ đơn giản, các nguồn nitơ hữu cơ, và các nguồn nitơ phức tạp. Các nguồn nitơ như vậy bao gồm amoniak khan, muối amoni của axit vô cơ hoặc axit hữu cơ như amoni clorua, amoni sulfat, amoni axetat, amoni phosphat, các hợp chất và cơ chất chứa nitơ khác có nguồn gốc từ động vật, thực vật, và/hoặc vi khuẩn. Axit amin cũng có thể được sử dụng để làm nguồn nitơ, bao gồm leuxin, isoleuxin hoặc valin, hoặc hỗn hợp của chúng.

Có thể dùng phương pháp đã biết bất kỳ để cung cấp cơ chất cho phản ứng lên men để duy trì mức cơ chất thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Các ví dụ minh họa bao gồm phương pháp theo mẻ trong đó tất cả cơ chất cho quá trình lên men được thêm vào tại thời điểm bắt đầu phản ứng lên men; phương pháp nạp liệu liên tục; và phương pháp tốc độ nạp liệu thay đổi, trong đó, ví dụ, cung cấp lượng cơ chất khi quá trình lên men tiếp diễn để hỗ trợ tăng nồng độ tế bào trong môi trường. Việc kết hợp ba phương pháp này thường được sử dụng. Ví dụ, thường có lượng cơ chất nhất định có mặt lúc ban đầu trong quá trình lên men, để cho vi sinh vật sử dụng cạn kiệt lượng cơ chất ban đầu này, sau đó bổ sung cơ chất một cách liên tục hoặc bổ sung cơ chất một cách thay đổi sau khi lượng cơ chất ban đầu đã được sử dụng. Một khía cạnh khác theo sáng chế đề xuất lượng cơ chất cho tế bào trong môi trường lên men, lượng cơ chất này có thể có mặt ở mức tương đối cao, để tế bào chủ gần như sử dụng hết cơ chất ban đầu, sau đó cung cấp cơ chất cho tế bào chủ ở mức dưới điểm cực thuận so với lượng sẽ hỗ trợ tốc độ sinh trưởng tối đa bằng tốc độ nạp liệu liên tục hoặc thay đổi. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng vì tế bào có thể sinh trưởng ở tốc độ bùng nổ, nên có thể

có lợi nếu thay đổi tốc độ nạp liệu ở tốc độ bùng nổ để giữ cho lượng cơ chất so với mức tế bào chủ không đổi. Do đó, theo các phương án nhất định, cơ chất được bổ sung theo cách tăng bùng nổ, tuy nhiên vẫn ở mức thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa.

Theo một số phương án, phản ứng lên men được cấp nguồn cacbon giảm so với nguồn cacbon sẽ tạo ra tốc độ sinh trưởng riêng tối đa. Không bị ràng buộc bởi lý thuyết cụ thể, đã biết rằng việc thay đổi các mức chất dinh dưỡng có sẵn đối với tế bào sẽ thay đổi dòng liên quan thông qua nhiều quá trình hoá học khác nhau trong tế bào. Ví dụ, một số enzym có tính cảm ứng, và sẽ chỉ hoạt động khi có mặt các chất dinh dưỡng nhất định. Các tác giả sáng chế đã quan sát thấy rằng việc hạ thấp tốc độ nạp liệu nguồn cacbon cho vi sinh vật có thể cải thiện lượng isoprenoit được sản xuất ra trong quá trình lên men. Thực tế, nguồn cacbon có thể được cung cấp ban đầu với lượng đủ để hỗ trợ sự sinh trưởng ban đầu của tế bào chủ cho đến khi nguồn cacbon ban đầu này gần như bị cạn kiệt, sau đó nguồn cacbon được thêm vào ở tốc độ bùng nổ, tuy nhiên ở tốc độ thấp hơn mức sẽ hỗ trợ sự sinh trưởng riêng tối đa của tế bào chủ. Ví dụ, trong phương pháp theo sáng chế, nguồn cacbon được thêm vào với lượng sẽ hỗ trợ khoảng 90%, 80%, 75%, 60%, 50%, 40%, 30%, 25%, 20%, 15%, 10%, hoặc thấp hơn tốc độ sinh trưởng riêng tối đa.

Theo phương án khác, phản ứng lên men được cấp viên chứa nguồn cacbon ban đầu đủ để sinh trưởng ở tốc độ sinh trưởng riêng tối đa hoặc tiệm cận tốc độ sinh trưởng riêng tối đa (sinh trưởng không bị giới hạn) sau đó tốc độ nạp liệu giảm, ở mức thấp hơn mức cần để hỗ trợ tốc độ sinh trưởng riêng tối đa, đối với phần còn lại của quá trình lên men. Trong một số trường hợp, thời điểm thực hiện tốc độ nạp liệu nguồn cacbon giảm là thời điểm đạt được tốc độ nạp liệu đã được định trước. Theo các phương án nhất định, vi sinh vật được cung cấp nguồn cacbon đủ để sinh trưởng bùng nổ đến tốc độ nạp liệu khoảng 15 g/l/giờ, sau đó tốc độ nạp liệu giảm xuống còn 5,7g/l/giờ và giữ không đổi ở tốc độ này đối với phần còn lại của quá trình lên men.

Theo các phương án nhất định, một hoặc nhiều enzym của quá trình mevalonat hoặc DXP khác loại là có khả năng cảm ứng và được cảm ứng sau khi tốc độ nạp liệu nguồn cacbon đã được làm giảm xuống đến mức thấp hơn mức cần cho sự sinh trưởng riêng tối đa. Ví dụ, nếu vi sinh vật đã được thiết kế di truyền chứa yếu tố khởi đầu cảm ứng, thì quá trình lên men đầu tiên được tiến hành bằng cách bổ sung nguồn cacbon để đạt được sự sinh trưởng bùng nổ, tuy nhiên vẫn ở mức thấp hơn mức hỗ trợ sinh trưởng riêng tối đa, thì tốc độ nạp liệu nguồn cacbon bị giảm xuống mức thậm chí thấp hơn đối với phần còn lại của quá trình lên men, và chất cảm ứng được bổ sung sau khi tốc độ nạp liệu nguồn cacbon này giảm. Theo một số phương án, vi sinh vật được cảm ứng bằng isopropylthio-beta-D-galactosit (IPTG) sau khi bắt đầu giảm nạp nguồn cacbon.

Theo phương án khác, phản ứng lên men được thực hiện theo cách tránh tạo ra các chất độc làm giảm tốc độ sinh trưởng tế bào. Ví dụ, đã biết rằng khi bổ sung quá nhiều glucoza vào môi trường, thì các sản phẩm độc như axetat có thể tạo ra trong sinh vật. Xem, ví dụ, ấn phẩm: Kortz et al. (1995) *J. Biotechnol* 39:59-65. Do đó, bằng cách cung cấp nguồn cacbon với mức cao, bằng lượng hoặc tiệm cận lượng sẽ hỗ trợ tốc độ sinh trưởng tối đa (sinh trưởng không bị giới hạn), thì sự sinh trưởng ban đầu của các tế bào có thể cao hơn, tuy nhiên sự sinh trưởng này sẽ bị hãm lại do sự tích lũy các chất độc. Mức nguồn cacbon được thêm vào thấp hơn mức mà ở đó không tích lũy các sản phẩm độc được dùng để chỉ mức tới hạn hoặc ngưỡng ức chế. Do đó, theo các phương án nhất định, phản ứng lên men được thực hiện sao cho nguồn cacbon được giữ ở dưới mức tới hạn đối với việc tạo thành các chất độc. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ hiểu rằng nồng độ cơ chất tới hạn sẽ thay đổi theo chủng và môi trường được sử dụng.

Hỗn hợp phản ứng lên men hữu hiệu có thể chứa các hợp chất khác như muối vô cơ, vitamin, kim loại vi lượng, hoặc chất thúc đẩy sinh trưởng. Các hợp chất khác này cũng có thể có mặt trong nguồn cacbon, nitơ hoặc khoáng chất trong hỗn hợp phản ứng hữu hiệu hoặc có thể được bổ sung một cách riêng

biệt vào hỗn hợp phản ứng. Một phương án theo sáng chế đề xuất việc cung cấp các hợp chất này ở mức dưới điểm cực thuận so với mức sẽ hỗ trợ tốc độ sinh trưởng tối đa của tế bào chủ để tăng việc sản xuất isoprenoit.

Hỗn hợp phản ứng lên men còn có thể chứa nguồn phosphat thích hợp. Nguồn phosphat như vậy bao gồm cả các nguồn phosphat vô cơ và nguồn phosphat hữu cơ. Ví dụ không giới hạn về các nguồn phosphat bao gồm, nhưng không bị giới hạn ở, muối phosphat như mono hoặc dibazơ natri và kali phosphat, amoni phosphat, polyphosphat, và hỗn hợp của chúng. Hỗn hợp phản ứng lên men thích hợp cũng có thể chứa nguồn magie. Theo một số phương án, magie ở dạng muối chấp nhận được về mặt sinh lý, như magie sulfat heptahydrat, mặc dù các nguồn magie khác ở nồng độ đóng góp lượng magie tương tự có thể được sử dụng. Ngoài ra, trong một số trường hợp có thể mong muốn để cho hỗn hợp phản ứng lên men trở nên bị cạn kiệt nguồn magie trong quá trình lên men. Theo một số phương án, nguồn phospho được cung cấp với lượng dưới điểm cực thuận so với lượng sẽ hỗ trợ tốc độ sinh trưởng riêng tối đa.

Hỗn hợp phản ứng lên men còn có thể chứa chất tạo chelat chấp nhận được về mặt sinh học, như dihydrat của trinatri xitrat và axit etylendiamintetraaxetic. Hỗn hợp phản ứng lên men này ban đầu còn có thể chứa axit hoặc bazơ chấp nhận được về mặt sinh học để duy trì độ pH mong muốn của hỗn hợp phản ứng lên men. Axit chấp nhận được về mặt sinh học bao gồm, nhưng không bị giới hạn ở, axit clohydric, axit sulfuric, axit nitric, axit phosphoric và hỗn hợp của chúng. Bazơ chấp nhận được về mặt sinh học bao gồm, nhưng không bị giới hạn ở, amoni hydroxit, natri hydroxit, kali hydroxit và hỗn hợp của chúng.

Hỗn hợp phản ứng lên men còn có thể chứa nguồn canxi chấp nhận được về mặt sinh học, bao gồm, nhưng không bị giới hạn ở, canxi clorua. Hỗn hợp phản ứng lên men còn có thể chứa natri clorua. Hỗn hợp phản ứng lên men còn có thể chứa kim loại vi lượng. Kim loại này có thể được bổ sung vào hỗn hợp phản ứng lên men dưới dạng dung dịch gốc mà, để thuận tiện, có thể được

chuẩn bị một cách riêng biệt từ phần còn lại của hỗn hợp phản ứng lên men. Dung dịch kim loại thích hợp có thể bao gồm, nhưng không bị giới hạn ở natri selenat, sắt sulfat, heptahydrat, cupic sulfat, pentahydrat, kẽm sulfat, heptahydrat, natri molybđat, dihydrat, coban clorua, dung dịch selen hoặc crom, hexahydrat, và mangan sulfat monohydrat. Axit clohydric có thể được bổ sung vào dung dịch gốc để giữ muối của kim loại trong dung dịch.

Nếu chất trung gian của quá trình chuyển hoá hoặc hợp chất mà có thể được chuyển hoá thành chất trung gian của quá trình chuyển hoá được bổ sung vào môi trường lên men, thì chất trung gian hoặc hợp chất này thường có mặt với lượng dư.

Quá trình lên men có thể được tiến hành trong điều kiện yếm khí (thiếu oxy) hoặc hiếu khí (được oxy hoá). Trong điều kiện hiếu khí, vi sinh vật có thể phân hủy các loại đường thành các sản phẩm cuối như CO_2 và H_2O . Trong điều kiện yếm khí, tế bào chủ sử dụng quá trình thay thế để sản xuất CO_2 và etanol. Quá trình lên men còn có thể được dùng để chỉ sự sinh trưởng ở mức lớn của vi sinh vật trong môi trường nuôi cấy trong đó không có sự khác biệt giữa trao đổi chất hiếu khí và yếm khí. Nói chung, quá trình lên men hiếu khí được thực hiện để sản xuất các isoprenoit.

Quy trình lên men theo sáng chế có thể được thực hiện theo quy trình theo mẻ, nạp liệu theo mẻ hoặc liên tục. Quy trình theo mẻ thường là quy trình kín trong đó toàn bộ nguyên liệu thô được bổ sung vào lúc bắt đầu quá trình lên men. Quy trình nạp liệu theo mẻ thường là quy trình kín trong đó nguồn cacbon và/hoặc các cơ chất khác được bổ sung với lượng gia tăng trong suốt quy trình. Quy trình nạp liệu theo mẻ cho phép kiểm soát tốt hơn môi trường và sự sinh trưởng của vi sinh vật. Quy trình liên tục có thể được coi là hệ thống mở trong đó môi trường được bổ sung một cách liên tục và sản phẩm được lấy ra một cách đồng thời. Quy trình kết hợp ba loại này cũng có thể được sử dụng. Ví dụ, theo một phương án, quy trình lên men bắt đầu là quy trình nạp liệu theo mẻ, và lớp hữu cơ, như dodecan được đặt tiếp xúc với môi trường lên men trong khi quá trình lên men tiếp diễn. Isoprenoit, thường có độ hoà tan trong môi trường

hữu cơ lớn hơn so với trong môi trường lên men chứa nước, được chiết khỏi môi trường lên men vào lớp hữu cơ. Nếu isoprenoit được sản xuất vượt quá điểm bão hoà và hình thành lớp có thể tách riêng khỏi môi trường, thì có thể thực hiện sự phân tách đơn giản bằng cách tháo rút hoặc hút lớp pha riêng biệt. Quy trình này có đặc điểm của cả quy trình nạp liệu theo mẻ và quy trình liên tục, do việc lấy sản phẩm khỏi môi trường và việc tiến hành quá trình lên men. Quy trình nạp liệu theo mẻ và quy trình liên tục cho phép kiểm soát sự bổ sung các thành phần của quá trình lên men trong quá trình lên men. Quy trình nạp liệu theo mẻ, quy trình liên tục, hoặc kết hợp của các quy trình này thường được ưu tiên trong việc thực hiện sáng chế. Các quy trình này cho phép kiểm soát tốt hơn tốc độ bổ sung nguyên liệu cấp và các thành phần khác của quá trình lên men theo thời gian. Việc thu lấy sản phẩm trong quá trình lên men có thể có lợi, đặc biệt là khi sản phẩm được tích lũy dẫn đến sự ức chế quá trình sản xuất.

Lượng vi sinh vật trên một lít môi trường lên men, hoặc mật độ vi sinh vật, có thể được xác định bằng cách xác định khối lượng vi sinh vật được phân lập từ thể tích môi trường lên men đã cho. Thông số phổ biến là khối lượng khô của tế bào trên một lít môi trường lên men. Có thể sử dụng phương pháp khác để giám sát quá trình lên men đang tiếp diễn là đo mật độ quang của môi trường. Phương pháp phổ biến là đo mật độ quang ở bước sóng 600nm, được gọi là OD₆₀₀, hoặc OD. OD có thể tương quan với mật độ của loại sinh vật cụ thể trong môi trường cụ thể, tuy nhiên quan hệ cụ thể giữa OD và lượng vi sinh vật trên đơn vị thể tích thường sẽ không thể áp dụng được đối với tất cả các loại sinh vật trong tất cả các loại môi trường. Đường cong định chuẩn có thể được tạo ra bằng cách đo OD và khối lượng tế bào khô theo một khoảng mật độ tế bào. Trong một số trường hợp, các mối tương quan này có thể được sử dụng trong các quá trình lên men khác nhau của cùng vi sinh vật hoặc vi sinh vật tương tự trong cùng môi trường hoặc môi trường tương tự.

Theo một khía cạnh khác, sáng chế đề xuất phương pháp bao gồm các bước (i) tiến hành phản ứng lên men bao gồm môi trường lên men và nhiều tế

bào chủ đã được cải biến di truyền sản xuất ra isoprenoit ở các điều kiện sao cho (a) môi trường lên men được giữ ở nhiệt độ thấp hơn nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ này; (b) môi trường lên men chứa nguồn cacbon có mặt với lượng thấp hơn lượng sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ; và/hoặc (c) môi trường lên men chứa nguồn nitơ có mặt với lượng thấp hơn lượng sẽ tạo ra tốc độ sinh trưởng riêng tối đa của tế bào chủ; (ii) thu hồi isoprenoit được tạo ra trong một hoặc nhiều điều kiện nêu trong các mục từ (a) đến (c). Theo một phương án, phản ứng lên men được tiến hành trong điều kiện (a). Theo một phương án khác, phản ứng lên men được tiến hành trong điều kiện (a) và (b). Theo một phương án khác nữa, phản ứng lên men được tiến hành trong điều kiện (a), (b), và (c), hoặc dạng kết hợp bất kỳ của chúng.

Bằng cách sử dụng phương pháp được mô tả ở đây, các tế bào chủ sản xuất ra nhiều hơn khoảng 10 gam isoprenoit trên một lít hỗn hợp phản ứng lên men (10g/l). Theo các phương án khác, sản xuất ra nhiều hơn khoảng 15g/l, nhiều hơn khoảng 20g/l, nhiều hơn 25g/l, hoặc nhiều hơn khoảng 30g/l isoprenoit.

Theo một phương án khác, các tế bào chủ sản xuất ra nhiều hơn khoảng 50mg isoprenoit trên một gam tế bào chủ khô (50mg trên một gam khối lượng tế bào khô) được sản xuất ra. Theo các phương án khác, nhiều hơn khoảng 100mg trên một gam khối lượng tế bào khô, nhiều hơn khoảng 150mg trên một gam khối lượng tế bào khô, nhiều hơn khoảng 200mg trên một gam khối lượng tế bào khô, nhiều hơn khoảng 250mg trên một gam khối lượng tế bào khô, nhiều hơn khoảng 500mg trên một gam khối lượng tế bào khô, nhiều hơn khoảng 750mg trên một gam khối lượng tế bào khô, hoặc nhiều hơn khoảng 1000mg trên một gam khối lượng tế bào khô của isoprenoit được sản xuất ra.

Theo các phương án khác, mức sản xuất, dù ở mức gam trên một lít hay mg trên một gam khối lượng tế bào khô đạt được trong khoảng thời gian thấp hơn 150 giờ, tốt hơn là thấp hơn khoảng 96 giờ, hoặc thậm chí thấp hơn khoảng 72 giờ.

Ví dụ không giới hạn về isoprenoit thích hợp bao gồm: hemiterpen (được tạo ra từ 1 đơn vị isopren) như isopren; monoterpen (được tạo ra từ 2 đơn vị isopren) như myrxen; sesquiterpen (được tạo ra từ 3 đơn vị isopren) như amomorpha-4,11-dien; diterpen (được tạo ra từ bốn đơn vị isopren) như taxadien; triterpen (được tạo ra từ 6 đơn vị isopren) squalen; tetraterpen (được tạo ra từ 8 isoprenoit) α -caroten; và polyterpen (được tạo ra từ nhiều hơn 8 đơn vị isopren) như polyisopren. Theo một số phương án, isoprenoit không phải là carotenoit. Theo các phương án khác, isoprenoit là C_5 - C_{20} isoprenoit.

Mặc dù sáng chế đã được mô tả kết hợp với các phương án cụ thể của nó, nhưng phần mô tả trên đây chỉ nhằm mục đích minh họa và không nhằm giới hạn phạm vi của sáng chế. Các khía cạnh, các ưu điểm và các cải biến khác cũng nằm trong phạm vi của sáng chế sẽ được người có hiểu biết trung bình trong lĩnh vực kỹ thuật này biết rõ.

Ví dụ thực hiện sáng chế

Trừ khi có quy định khác, việc thực hiện sáng chế có thể sử dụng các phương pháp thông thường của ngành công nghiệp sinh tổng hợp và các ngành khác thuộc lĩnh vực kỹ thuật này. Trong trường hợp các phương pháp này không được mô tả một cách đầy đủ ở đây, có thể tham khảo các phương pháp đó một cách đầy đủ trong các tài liệu khoa học.

Trong các ví dụ sau, mặc dù đã cố gắng để đảm bảo độ chính xác đối với các con số được sử dụng (ví dụ, khối lượng, nhiệt độ, và các thông số tương tự), tuy nhiên độ biến thiên và độ lệch có thể được xem xét, và trong trường hợp tồn tại sai số ghi chép trong các con số được dùng ở đây, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể suy ra trị số đúng khi xem xét phần bộc lộ còn lại ở đây. Trừ khi có quy định khác, nhiệt độ được báo cáo là độ C, và áp suất là tại hoặc gần áp suất khí quyển ở mực nước biển. Trừ khi có quy định khác, các chất phản ứng được mua từ thị trường. Các ví dụ sau chỉ nhằm mục đích minh họa và không nhằm giới hạn phạm vi của sáng chế theo cách bất kỳ.

Ví dụ 1

Ví dụ này mô tả các phương pháp tạo ra plasmit biểu hiện mã hóa enzym bao gồm enzym của quá trình MEV từ *Saccharomyces cerevisiae* được sắp xếp trong các operon.

Plasmit biểu hiện pMevT được tạo ra bằng cách chèn operon MevT (SEQ ID NO: 1) vào vector pBAD33. Operon MevT mã hóa tập hợp các enzym của quá trình MEV mà cùng nhau chuyển hóa tiền chất thường gặp axetyl-CoA thành (R)-mevalonat, cụ thể là axetoaxetyl-CoA thiolaza, HMG-CoA synthaza, và HMG-CoA reductaza. Operon MevT được tạo ra bằng cách khuếch đại PCR từ ADN hệ gen của *Escherichia coli* trình tự mã hóa của gen *atoB* (số đăng ký ngân hàng gen NC 000913 REGION: 2324131..2325315) (mã hóa axetoaxetyl-CoA thiolaza), từ ADN hệ gen của *Saccharomyces cerevisiae* trình tự mã hóa của gen *ERG13* (số đăng ký ngân hàng gen X96617, REGION: 220..1695) (mã hóa HMG-CoA synthaza), và từ ADN hệ gen của *Saccharomyces cerevisiae* đoạn của vùng mã hóa của gen *HMG1* (số đăng ký ngân hàng gen M22002, REGION: 1660..3165) (mã hóa HMG-CoA reductaza bị cắt cụt (tHMGR)). Đoạn mỗi PCR ngược dòng được dùng để khuếch đại đoạn gen *HMG1* chứa codon khởi đầu nhân tạo. Các đoạn đã được khuếch đại này được gắn với nhau bằng cách sử dụng kỹ thuật kéo dài chồng lắp (SOEing), trong quá trình này các vị trí liên kết ribosom được đưa vào sau các trình tự mã hóa *atoB* và *ERG13*. Sau khi bổ sung đầu treo 3' A, operon MevT được gắn vào vector tách dòng TA pCR4 (Invitrogen, Carlsbad, CA), và được giải trình tự để đảm bảo độ chính xác. Operon MevT sau đó được gắn vào vị trí enzym giới hạn *XmaI* *PstI* của vector pBAD33 (xem ấn phẩm: Guzman et al. (1995) *J. Bacteriol.* 177(14): 4121-4130). Để đặt operon này dưới sự kiểm soát của yếu tố khởi đầu P_{Lac}, đoạn *araC-P_{BAD}NsiI-XmaI* của pBAD33 được thay thế bằng đoạn *NsiI-XmaI* của pBBR1MCS, tạo ra plasmit biểu hiện pMevT (xem Patent Mỹ số 7,192,751).

Plasmit biểu hiện pAM36-MevT66 được tạo ra bằng cách chèn operon MevT66 vào vector pAM36. Vector pAM36 được tạo ra bằng cách chèn catxet oligonucleotit chứa vị trí enzym giới hạn *AscI-SfiI-AsiSI-XhoI-PacI-FsII-PmeI*

vào vector pACYC184 (số đăng ký ngân hàng gen XO6403), và bằng cách loại bỏ gen kháng *tet* trong pACYC184. Operon MevT66 được tạo ra bằng cách tổng hợp nhờ sử dụng trình tự nucleotit SEQ ID NO: 1 làm khuôn, khuôn này chứa gen *atoB* từ *Escherichia coli* (số đăng ký ngân hàng gen NC_000913 REGION 2324131..2325315), gen *ERG13* từ *Saccharomyces cerevisiae* (số đăng ký ngân hàng gen X96617, REGION: 220..1695), và kiểu cắt cụt của gen *HMG* từ *Saccharomyces cerevisiae* (số đăng ký ngân hàng gen M22002, REGION 1777..3285), cả ba trình tự này được tối ưu hoá codon để biểu hiện trong *Escherichia coli*. Operon MevT66 đã được tạo ra bằng cách tổng hợp được gắn vào bên sườn bởi vị trí enzyme giới hạn 5' *EcoRI* và vị trí enzyme giới hạn 3' *Hind III*, và do đó có thể được tách dòng vào các vị trí enzyme giới hạn tương thích của vector tách dòng như vector gốc pUC hoặc pACYC chuẩn. Từ cấu trúc này, operon MevT66 được khuếch đại bằng PCR với các vị trí enzyme giới hạn *SfiI* và *AsiSI* gắn vào bên sườn, đoạn ADN đã được khuếch đại này được cắt để hoàn thành bằng cách sử dụng các enzyme giới hạn *SfiI* và *AsiSI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 4,2 kb được tách bằng gel có sử dụng bộ kit tinh chế gel Qiagen (Valencia, CA), và đoạn ADN đã phân lập được này được gắn vào vị trí enzyme giới hạn *SfiI AsiSI* của vector pAM36, tạo ra plasmid biểu hiện pAM36-MevT66.

Plasmid biểu hiện pAM25 được tạo ra bằng cách chèn operon MevT66 vào vector pAM29. Vector pAM29 được tạo ra bằng cách lắp ráp gốc sao chép p15A và gen kháng *kan* từ pZS24-MCS1 (xem ấn phẩm: Lutz and Bujard (1997) *Nucl Acids Res.* 25 1203-1210) với yếu tố khởi đầu *lacUV5* được tạo ra bởi oligonucleotit-. Cấu trúc tổng hợp ADN chứa operon MevT66 (xem ở trên) được cắt hoàn toàn bằng cách sử dụng các enzyme giới hạn *EcoRI* và *Hind III*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN 4,2kb được chiết gel, và đoạn ADN đã phân lập được này được gắn vào vị trí enzyme giới hạn *EcoRI HindIII* của pAM29, để thu được plasmid biểu hiện pAM25.

Plasmid biểu hiện pMevB-Cm được tạo ra bằng cách chèn operon MevB vào vector pBBR1MCS-1. Operon MevB mã hóa tập hợp các enzym mà cùng nhau chuyển hoá (R)-mevalonat thành IPP, cụ thể là mevalonat kinaza, phosphomevalonat kinaza, và mevalonat pyrophosphat cacboxylaza. Operon MevB được tạo ra bằng cách khuếch đại bằng PCR từ ADN hệ gen của *Saccharomyces cerevisiae* các trình tự mã hóa của gen *ERG12* (số đăng ký ngân hàng gen X55875, REGION 580 1911) (mã hóa mevalonat kinaza), gen *ERG8* (số đăng ký ngân hàng gen Z49939, REGION: 3363..4718) (mã hóa phosphomevalonat kinaza), và gen *MVD1* (số đăng ký ngân hàng gen X97557, REGION:544 1734) (mã hóa mevalonat pyrophosphat cacboxylaza), và bằng cách gắn các đoạn PCR với nhau sử dụng kỹ thuật kéo dài chồng lắp (SOEing). Bằng cách chọn các trình tự môi thích hợp, codon kết thúc của *ERG12* và *ERG8* được chuyển từ TAA thành TAG trong quá trình khuếch đại để đưa vào các vị trí liên kết ribosom. Sau khi bổ sung đầu treo 3' A, operon MevB được gắn vào vector tách dòng TA pCR4 (Invitrogen, Carlsbad, CA). Operon MevB được cắt bỏ bằng cách phân giải cấu trúc tách dòng này đến hoàn toàn bằng cách sử dụng enzym giới hạn *PstI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, chiết bằng gel đoạn ADN 4,2kb, và gắn đoạn ADN đã được tách riêng này vào vị trí enzym giới hạn *PstI* của vector pBBR1MCS-1 (xem ấn phẩm: Kovach *et al.*, *Gene* 166(1) 175-176 (1995)), để thu được plasmid biểu hiện pMevB-Cm.

Plasmid biểu hiện pMBI được tạo ra bằng cách chèn operon MBI vào vector pBBR1MCS-3. Operon MBI mã hóa cùng enzym như operon MevB, cũng như isopentenyl pyrophosphataza isomeraza, enzym này xúc tác sự chuyển hoá IPP thành DMAPP. Operon MBI được tạo ra bằng cách khuếch đại bằng PCR từ ADN hệ gen của *Escherichia coli* trình tự mã hóa của gen *idi* (số đăng ký ngân hàng gen AF 119715) sử dụng các môi chứa vị trí enzym giới hạn *XmaI* ở đầu 5' của chúng, cắt đoạn ADN đã được khuếch đại này đến hoàn toàn bằng cách sử dụng enzym giới hạn *XmaI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, chiết bằng gel đoạn 0,5kb, và gắn đoạn ADN đã được tách riêng này vào vị trí enzym giới hạn *XmaI* của plasmid biểu hiện

pMevB-Cm, bằng cách này đặt *idi* ở đầu 3' của operon MevB. Operon MBI được tách dòng phụ vào vị trí enzym giới hạn *SalI* và *SacI* của vector pBBR1MCS-3 (xem ấn phẩm: Kovach *et al.*, *Gene* 166(1) 175-176 (1995)), để thu được plasmit biểu hiện pMBI (xem Patent Mỹ số 7, 192,751).

Plasmit biểu hiện pMBIS được tạo ra bằng cách chèn gen *ispA* vào pMBI. Gen *ispA* mã hóa farnesyl pyrophosphat synthaza xúc tác cho việc ngưng tụ của hai phân tử IPP với một phân tử DMAPP để tạo ra farnesyl pyrophosphat (FPP). Trình tự mã hóa của gen *ispA* (số đăng ký ngân hàng gen D00694, REGION 484 1383) được khuếch đại bằng PCR từ ADN của hệ gen *Escherichia coli* sử dụng mỗi xuôi với vị trí enzym giới hạn *SacII* và mỗi ngược với vị trí enzym giới hạn *SacI*. Sản phẩm PCR đã được khuếch đại được phân giải hoàn toàn bằng các enzym giới hạn *SacII* và *SacI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, và đoạn ADN 0,9 kb được tách bằng gel. Đoạn ADN đã được tách riêng được gắn vào vị trí enzym giới hạn *SacII* *SacI* của pMBI, bằng cách này đặt gen *ispA* ở đầu 3' của *idi* và operon MevB, và tạo ra plasmit biểu hiện pMBIS (xem Patent Mỹ số 7, 192,751).

Plasmit biểu hiện pMBIS-gpps được tạo ra từ plasmit biểu hiện pMBIS bằng cách thay thế trình tự mã hóa *ispA* bằng trình tự nucleotit mã hóa geranyl diphosphat synthaza ("gpps"). Đoạn ADN chứa trình tự nucleotit mã hóa geranyl diphosphat synthaza này được tạo ra bằng cách tổng hợp sử dụng trình tự mã hóa của gen *gpps* của *Arabidopsis thaliana* (số đăng ký ngân hàng gen Y17376, REGION 52..1320), được tối ưu hoá codon để biểu hiện trong *Escherichia coli*, để làm khuôn. Trình tự nucleotit được gắn bên sườn bằng vị trí enzym giới hạn *SacII* dẫn đầu và vị trí enzym giới hạn *SacI* cuối cùng, và có thể được tách dòng vào các vị trí enzym giới hạn tương thích của vector tách dòng như vector gốc pUC hoặc pACYC chuẩn. Trình tự geranyl diphosphat synthaza đã được tạo ra bằng cách tổng hợp này được phân lập bằng cách phân giải cấu trúc tổng hợp ADN đến hoàn toàn bằng cách sử dụng các enzym giới hạn *SacII* và *SacI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, tách bằng gel đoạn ADN khoảng 1,3kb, và gắn đoạn ADN đã được tách

riêng vào vị trí enzym giới hạn *SacII* *SacI* của plasmit biểu hiện pMBIS, để thu được plasmit biểu hiện pMBIS-gpps (xem Fig.3 là bản đồ plasmit).

Plasmit biểu hiện pAM45 được tạo ra bằng cách chèn operon MBIS vào pAM36-MevT66 và bổ sung yếu tố khởi đầu *lacUV5* vào trước hai operon này. Operon MBIS được khuếch đại bằng PCR từ pMBIS sử dụng các mồi chứa vị trí enzym giới hạn 5'*XhoI* và vị trí enzym giới hạn 3' *PacI*. Sản phẩm PCR đã được khuếch đại này được phân giải hoàn toàn bằng cách sử dụng các enzym giới hạn *XhoI* và *PacI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN 5,4kb được tách bằng gel, và đoạn ADN đã được tách riêng được gắn vào vị trí enzym giới hạn *XhoI* *PacI* của pAM36-MevT66, để thu được plasmit pAM43. Đoạn ADN chứa trình tự nucleotit mã hóa yếu tố khởi đầu *lacUV5* được tổng hợp từ oligonucleotit và được tách dòng phụ vào vị trí enzym giới hạn *AscI* *SfiI* và *AsiSI* *XhoI* của pAM43, để thu được plasmit biểu hiện pAM45.

Ví dụ 2

Ví dụ này mô tả phương pháp tạo ra vector biểu hiện mã hóa enzym bao gồm enzym của quá trình MEV từ *Staphylococcus aureus* được sắp xếp trong các operon.

Plasmit biểu hiện pAM41 được tạo ra từ plasmit biểu hiện pAM25 bằng cách thay thế trình tự mã hóa của gen *HMG1I*, gen này mã hóa HMG-CoA reductaza của *Saccharomyces cerevisiae*, bằng trình tự mã hóa của gen *mvaA*, gen này mã hóa HMG-CoA reductaza của *Staphylococcus aureus* (số đăng ký ngân hàng gen BA000017, REGION: 2688925..2687648). Trình tự mã hóa của gen *mvaA* được khuếch đại bằng PCR từ ADN hệ gen của *Staphylococcus aureus* subsp. aureus (ATCC 70069) sử dụng các mồi 4-49 *mvaA* *SpeI* (SEQ ID NO: 2) và 4-49 *mvaAR* *XbaI* (SEQ ID NO: 3), đoạn ADN đã được khuếch đại này được phân giải hoàn toàn bằng cách sử dụng enzym giới hạn *SpeI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, và đoạn ADN khoảng 1,3kb được tách bằng gel. Trình tự mã hóa *HMG1* được loại bỏ khỏi pAM25 bằng cách phân giải plasmit đến hoàn toàn sử dụng enzym giới hạn

HindIII. Đầu treo tận cùng của đoạn ADN mạch thẳng tạo thành được cắt cùn bằng cách sử dụng T4 ADN polymeraza. Sau đó, đoạn ADN này được phân giải một phần sử dụng enzym giới hạn *SpeI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, và đoạn ADN 4,8kb được tách bằng gel. Đoạn ADN đã được tách riêng được gắn với sản phẩm PCR *mvaA* đã được phân giải bằng *SpeI*, để thu được plasmit biểu hiện pAM41. Trình tự nucleotit của operon *atoB(opt):ERG13(opt):mvaA* được chứa trong pAM41 là SEQ ID NO: 41.

Plasmit biểu hiện pAM52 được tạo ra từ plasmit biểu hiện pAM41 bằng cách thay thế trình tự mã hóa của gen *ERG13*, gen này mã hóa HMG-CoA synthaza của *Saccharomyces cerevisiae* bằng trình tự mã hóa của gen *mvaS*, gen này mã hóa HMG-CoA synthaza của *Staphylococcus aureus* (số đăng ký ngân hàng gen BA000017, REGION: 2689180..2690346). *ERG13* còn được gọi là HMGS hoặc HMG-CoA synthaza. Trình tự mã hóa của gen *mvaS* được khuếch đại bằng PCR từ ADN hệ gen của *Staphylococcus aureus* subsp. *aureus* (ATCC 70069) sử dụng các môi HMGS 5' *Sa mvaS-S* (SEQ ID NO: 4) và HMGS 3' *Sa mvaS-AS* (SEQ ID NO: 5), và đoạn ADN đã khuếch đại này được dùng làm môi PCR để thay thế trình tự mã hóa của gen *HMG1* trong pAM41 theo phương pháp của Geiser et al (*BioTechniques* 31:88-92 (2001)), để thu được plasmit biểu hiện pAM52. Trình tự nucleotit của operon *atoB(opt):mvaS:mvaA* được chứa trong pAM52 là SEQ ID NO: 42.

Plasmit biểu hiện pAM97 được tạo ra từ plasmit biểu hiện pAM45 bằng cách thay thế operon *MevT66* bằng operon (*atoB(opt):mvaS:mvaA*) của plasmit biểu hiện pAM52. Plasmit biểu hiện pAM45 được phân giải hoàn toàn bằng cách sử dụng các enzym giới hạn *AsiSI* và *SfiI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, và đoạn ADN 8,3kb khuyết operon *MevT66* được tách bằng gel. Operon (*atoB(opt):mvaS:mvaA*) của pAM52 được khuếch đại bằng PCR sử dụng các môi 19-25 *atoB SfiI-S* (SEQ ID NO: 6) và 19-25 *mvaA-AsiSI-AS* (SEQ ID NO: 7), sản phẩm PCR này được phân giải hoàn toàn sử dụng các enzym giới hạn *SfiI* và *AsiSI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN 3,7kb được tách bằng gel,

và đoạn ADN đã được tách riêng này được gắn vào vị trí enzym giới hạn *AsiSI* *SfiI* của plasmit biểu hiện pAM45, để thu được plasmit biểu hiện pAM97.

Plasmit biểu hiện pAM97-MBI được tạo ra từ plasmit biểu hiện pAM97 và pAM45 bằng cách thay thế operon MBIS của pAM97 bằng operon MBI của pAM45. Operon MBI được khuếch đại bằng PCR từ pAM45 sử dụng các mồi 9-70C (SEQ ID NO: 8) và 26-39B (SEQ ID NO: 9), hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN 4,5kb được tách bằng gel, và đoạn ADN đã được tách riêng này được phân giải hoàn toàn sử dụng enzym giới hạn *SacI* và *XhoI*. Plasmit biểu hiện pAM97 được phân giải hoàn toàn sử dụng enzym giới hạn *SacI* và *XhoI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn 7,6kb được tách bằng gel, và đoạn ADN đã được tách riêng này được gắn với sản phẩm PCR operon MBI, để thu được plasmit biểu hiện pAM97-MBI.

Plasmit biểu hiện pAM97-MevB được tạo ra từ plasmit biểu hiện pAM97 và pAM45 bằng cách thay thế operon MBIS của pAM97 bằng operon MevB của pAM45. Operon MevB được khuếch đại bằng PCR từ pAM45 sử dụng mồi 9-70C (SEQ ID NO:8) và 26-39A (SEQ ID NO: 10), hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN 3,9kb được tách bằng gel, và đoạn ADN đã được tách riêng này được phân giải hoàn toàn sử dụng enzym giới hạn *SacI* và *XhoI*. Plasmit biểu hiện pAM97 được phân giải hoàn toàn sử dụng enzym giới hạn *SacI* và *XhoI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn 7,6kb được tách bằng gel, và đoạn ADN đã được tách riêng này được gắn với sản phẩm PCR của operon MevB, để thu được plasmit biểu hiện pAM97-MevB.

Plasmit biểu hiện pAM128 được tạo ra bằng cách chèn các operon (*atoB(opt):mvaS:mvaA*) và MBIS của plasmit biểu hiện pAM97 vào vector chứa hệ duy trì, phân tách và sao chép của plasmit RK2, hệ này ngăn ngừa nhu cầu liên tục đối với việc chọn chất kháng sinh của thể biến nạp tế bào chủ. Plasmit RK2 được phân giải hoàn toàn sử dụng enzym giới hạn *PstI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 6,3kb chứa

toàn bộ locus *par* được tách bằng gel, và đoạn ADN đã được tách riêng này được tách dòng phụ vào vị trí enzym giới hạn *PstI* của mini RK2 replicon pRR10 (xem ấn phẩm: Roberts et al. (1990) *J Bacteriol.* **172**(11): 6204-6216), để thu được vector pAM132. Plasmid biểu hiện pAM97 được phân giải hoàn toàn sử dụng enzym giới hạn *AscI* và *SacI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 9,4kb được tách bằng gel, và đoạn ADN đã được tách riêng này được gắn vào vị trí enzym giới hạn *MluI* *SacI* của pAM132, để thu được plasmid biểu hiện pAM128.

Ví dụ 3

Ví dụ này mô tả phương pháp tạo ra vector biểu hiện mã hóa enzym bao gồm enzym của quá trình MEV từ *Enterococcus faecalis* được sắp xếp trong các operon.

Plasmid pAM16 được tạo ra bằng cách chèn trình tự mã hóa của gen *mvaE* của *Enterococcus faecalis* (số đăng ký ngân hàng gen AF290092 REGION: 1479..3890) (mã hóa axetyl-CoA axetyltransferaza/HMG-CoA reductaza (HMGR)) vào vector pBlueScripII-KS(+). Trình tự mã hóa của gen *mvaE* được khuếch đại bằng PCR từ ADN hệ gen của *Enterococcus faecalis* (ATCC 700802) sử dụng các mồi được phosphoryl hoá đầu 5' 4-40 mvaEF BamHI (SEQ ID NO: 11) và 4-40 mvaERHindIII (SEQ ID NO: 12). (Lưu ý rằng mồi 4-40 mvaEF BamHI chuyển codon khởi đầu của gen *mvaE* từ TTG thành ATG trong sản phẩm PCR đã được khuếch đại). Sản phẩm PCR tạo ra này được gắn vào vị trí enzym giới hạn *SmaI* của pBlueScripII-KS(+) (Stratagene, La Jolla, CA), để thu được plasmid biểu hiện pAM16.

Plasmid pAM18 được tạo ra bằng cách chèn trình tự mã hóa của gen *mvaS* của *Enterococcus faecalis* (số đăng ký ngân hàng gen AF290092 REGION: 142..1293) (mã hóa HMG-CoA synthaza (HMGS)) vào vector pBlueScripII-KS(+). Trình tự mã hóa của gen *mvaS* được khuếch đại bằng PCR từ ADN hệ gen của *Enterococcus faecalis* (ATCC 700802) sử dụng các mồi được phosphoryl hoá đầu 5' 4-40 mvaSF BglIII (SEQ ID NO: 13) và 4-39 mvaSR BamHI (SEQ ID NO: 14), và sản phẩm PCR này được gắn vào vị trí

enzym giới hạn *SmaI* của pBlueScripII-KS(+) (Stratagene, La Jolla, CA), để thu được plasmit biểu hiện pAM18.

Plasmit biểu hiện pAM22 được tạo ra bằng cách chèn trình tự mã hóa của gen *mvaE* của plasmit biểu hiện pAM16 vào vector pZE21-P_{L-lacO1}. Vector pZE21-P_{L-lacO1} là dẫn xuất của vector pZE21-MCS-1 trong đó yếu tố khởi đầu tet được thay thế bằng yếu tố khởi đầu P_{L-lacO1} (xem ấn phẩm: Lutz and Bujard (1997) *Nucl Acids Res.* 25:1203-1210). Plasmit biểu hiện pAM 16 được phân giải hoàn toàn sử dụng enzym giới hạn *BamHI* và *HindIII*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 2,4kb chứa trình tự mã hóa *mvaE* được tách bằng gel, và đoạn ADN đã được tách riêng này được chèn vào vị trí enzym giới hạn *BamHI HindIII* của pZE21-P_{L-lacO1}, để thu được plasmit biểu hiện pAM22.

Plasmit biểu hiện pAM33 được tạo ra bằng cách chèn trình tự mã hóa của gen *mvaS* của plasmit biểu hiện pAM18 vào plasmit biểu hiện pAM22. Plasmit biểu hiện pAM18 được phân giải hoàn toàn sử dụng enzym giới hạn *BglII* và *BamHI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 1,2kb chứa trình tự mã hóa của gen *mvaS* được tách bằng gel, và đoạn ADN đã được tách riêng này được chèn vào vị trí *BamHI* của plasmit biểu hiện pAM22, để thu được plasmit biểu hiện pAM33.

Plasmit biểu hiện pAM34 được tạo ra bằng cách chèn operon *mvaS-mvaE* của plasmit biểu hiện pAM33 vào vector pAM29. Operon *mvaS-mvaE* được phân lập bằng cách phân giải một phần pAM33 bằng cách sử dụng enzym giới hạn *EcoRI*, phân giải đoạn ADN mạch thẳng tạo ra bằng cách sử dụng enzym giới hạn *MluI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, và tách bằng gel đoạn ADN khoảng 3,6kb. Khung vector của pAM29 thu được bằng cách phân giải hoàn toàn vector biểu hiện pAM25 sử dụng enzym giới hạn *MluI* và *EcoRI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, và tách bằng gel đoạn ADN khoảng 2,1kb. Hai đoạn ADN đã được phân lập này được gắn, để thu được plasmit biểu hiện pAM34.

Ví dụ 4

Ví dụ này mô tả phương pháp tạo ra plasmid biểu hiện mã hóa enzym, bao gồm enzym của quá trình DXP từ *Escherichia coli* được sắp xếp trong các operon.

Plasmid biểu hiện pAM408 được tạo ra bằng cách chèn gen mã hóa enzym của quá trình DXP "đỉnh" vào vector pAM29. Enzym của quá trình DXP "đỉnh" bao gồm 1-deoxy-D-xyluloza-5-phosphat synthaza (được mã hoá bởi gen *dxs* của *Escherichia coli*), 1-deoxy-D-xyluloza-5-phosphat reductoisomerase (được mã hoá bởi gen *dxr* của *Escherichia coli*), 4-diphosphoxytidyl-2C-metyl-D-erythritol synthaza (được mã hoá bởi gen *ispD* của *Escherichia coli*), và 4-diphosphoxytidyl-2C-metyl-D-erythritol synthaza (được mã hoá bởi gen *ispE* của *Escherichia coli*), cùng nhau chuyển hóa pyruvat và D-glyxeraldehyt-3-phosphat thành 4-diphosphoxytidyl-2C-metyl-D-erythritol-2-phosphat. Các đoạn ADN chứa trình tự nucleotit mã hóa enzym của quá trình DXP "đỉnh" này được tạo ra bằng cách khuếch đại bằng PCR trình tự mã hóa của các gen *dxs* (số đăng ký ngân hàng gen U00096 REGION 437539.. 439401), *dxr* (số đăng ký ngân hàng gen U00096 REGION 193521..194717), *ispD* (số đăng ký ngân hàng gen U00096 REGION 2869803.. 2870512), và *ispE* (số đăng ký ngân hàng gen U00096 REGION 1261249..1262100) từ chủng *Escherichia coli* DH1 (ATCC #33849) bằng các trình tự Shine Dalgarno tối ưu được bổ sung và vị trí enzym giới hạn đầu 5' và 3' sử dụng mỗi PCR được thể hiện trong SEQ ID các số 15-18. Các sản phẩm PCR này được phân tách bằng phương pháp điện di gel, được tách bằng gel sử dụng bộ kit tinh sạch gel Qiagen (Valencia, CA), được phân giải hoàn toàn sử dụng enzym giới hạn thích hợp (*XhoI* và *KpnI* đối với sản phẩm PCR chứa gen *dxs*, *KpnI* và *ApaI* đối với sản phẩm PCR chứa gen *dxr*, *ApaI* và *NdeI* đối với sản phẩm PCR chứa gen *ispD*, *NdeI* và *MluI* đối với sản phẩm PCR chứa gen *ispE*), và được tinh chế bằng cách sử dụng bộ kit tinh chế PCR Qiagen (Valencia, CA). Sau đó, lượng xấp xỉ đẳng mol của mỗi sản phẩm PCR được bổ sung vào phản ứng ghép nối để lắp ráp các gen riêng rẽ vào trong operon. Từ phản ứng ghép nối này, 1µl hỗn hợp phản ứng thường dùng để khuếch đại bằng PCR 2 catxet gen riêng biệt, cụ thể là catxet gen *dxs-dxr* và *ispD-ispE*.

Catxet gen *dxs-dxr* được khuếch đại bằng PCR bằng cách sử dụng các mồi 67-1A-C (SEQ ID NO 15) và 67-1D-C (SEQ ID NO 18), và catxet gen *ispD-ispE* được khuếch đại bằng PCR sử dụng các mồi 67-1E-C (SEQ ID NO 19) và 67-1H-C (SEQ ID NO: 22). Hai sản phẩm PCR này được phân tách bằng phương pháp điện di gel, và được tách bằng gel. Sản phẩm PCR chứa catxet gen *dxs-dxr* được phân giải hoàn toàn sử dụng enzym giới hạn *XhoI* và *ApaI*, và sản phẩm PCR chứa catxet gen *ispD-ispE* được phân giải hoàn toàn sử dụng enzym giới hạn *ApaI* và *MluI*, và hai sản phẩm PCR này được tinh chế. Vectơ pAM29 được phân giải hoàn toàn sử dụng enzym giới hạn *Sall* và *MluI*, và hai sản phẩm PCR đã được phân giải chứa operon của quá trình DXP "đỉnh" được gắn vào vị trí enzym giới hạn *Sall MluI* của vectơ pAM29, để thu được plasmit biểu hiện pAM408 (xem Fig.4 về bản đồ plasmit).

Plasmit biểu hiện pAM409 được tạo ra bằng cách chèn gen mã hóa enzym của quá trình DXP "đáy" vào vectơ pAM369. Enzym của quá trình DXP "đáy" bao gồm 2C-metyl-D-erythritol 2,4-xyclodiphosphat synthaza (được mã hoá bởi gen *ispF* của *Escherichia coli*), 1-hydroxy-2-metyl-2-(E)-butenyl-4-diphosphat synthaza (được mã hoá bởi gen *ispG* của *Escherichia coli*), và isopentenyl/dimetylalyl diphosphat synthaza (được mã hoá bởi gen *ispH* của *Escherichia coli*), cùng nhau chuyển hóa 4-diphosphoxytidyl-2C-metyl-D-erythritol-2-phosphat thành IPP và DMAPP. IPP cũng được chuyển hoá thành DMAPP nhờ hoạt động của isopentyl diphosphat isomeraza (được mã hoá bởi gen *idi* của *Escherichia coli*). DMAPP có thể được chuyển hoá tiếp thành FPP nhờ hoạt động của farnesyl diphosphat synthaza (được mã hoá bởi gen *ispA* của *Escherichia coli*). Operon mã hóa enzym của quá trình DXP "đáy" cũng như isopentyl diphosphat isomeraza và farnesyl diphosphat synthaza được tạo ra bằng cách khuếch đại bằng PCR các gen *ispF* (số đăng ký ngân hàng gen U00096 REGION: 2869323..2869802), *ispG* (số đăng ký ngân hàng gen U00096 REGION: 2638708..2639826), *ispH* (số đăng ký ngân hàng gen U00096 REGION: 26277..27227), *idi* (số đăng ký ngân hàng gen AF 119715), và *ispA* (số đăng ký ngân hàng gen D00694 REGION: 484..1383) từ chủng *Escherichia coli* DH1 (ATCC #33849) bằng các trình tự Shine Dalgarno tối ưu

được bổ sung và vị trí enzym giới hạn đầu 5' và 3' sử dụng các môi PCR đầu 5' và 3'. Sản phẩm PCR này được phân tách bằng phương pháp điện di gel, được tách bằng gel, được phân giải bằng enzym giới hạn thích hợp (*BamHI* và *ApaI* đối với sản phẩm PCR chứa gen *ispF*; *KpnI* và *ApaI* đối với sản phẩm PCR chứa gen *ispG*; *Sall* và *KpnI* đối với sản phẩm PCR chứa gen *ispH*; *Sall* và *HindIII* đối với sản phẩm PCR chứa gen *idi*; *HindIII* và *NcoI* đối với sản phẩm PCR chứa gen) *ispA*, và được tinh chế. Sau đó, lượng xấp xỉ bằng mol của mỗi sản phẩm PCR được bổ sung vào phản ứng ghép nối để lắp ráp các gen riêng rẽ trong operon. Từ phản ứng nối này, 1 μ l hỗn hợp phản ứng thường dùng để khuếch đại bằng PCR 2 catxet gen riêng biệt, cụ thể là catxet gen *ispF-ispG* và *ispH-idi-ispA*. Catxet gen *ispF-ispG* được khuếch đại bằng PCR sử dụng các môi 67-2A-C (SEQ ID NO: 23) và 67-2D-C (SEQ ID NO: 26), và catxet gen *ispH-idi-ispA* được khuếch đại bằng PCR sử dụng các môi 67-2E-C (SEQ ID NO: 27) và 67-2J-C (SEQ ID NO: 32). Hai sản phẩm PCR này được phân tách bằng phương pháp điện di gel, và được tách bằng gel. Sản phẩm PCR chứa catxet gen *ispF-ispG* được phân giải hoàn toàn sử dụng enzym giới hạn *BamHI* và *KpnI*, và sản phẩm PCR chứa catxet gen *ispH-idi-ispA* được phân giải hoàn toàn sử dụng enzym giới hạn *KpnI* và *NcoI*, và hai sản phẩm PCR này được tinh chế. Vector pAM369 được tạo ra bằng cách lắp ráp gốc sao chép p15A từ pAM29 và gen beta-lactamaza kháng ampicilin từ pZE12-luc (xem ấn phẩm: Lutz and Bujard (1997) *Nucl Acids Res.* 25:1203-1210) với oligonucleotit-yếu tố khởi đầu *lacUV5* đã được tạo ra. Vector pAM369 được phân giải hoàn toàn sử dụng enzym giới hạn *BamHI* và *NcoI*, và 2 sản phẩm PCR đã được phân lập chứa operon của quá trình DXP “đáy” được gắn vào vị trí enzym giới hạn *BamHI NcoI* của vector pAM369, để thu được plasmit biểu hiện pAM409.

Plasmit biểu hiện pAM424, là dẫn xuất của plasmit biểu hiện pAM409 chứa gốc sao chép RK2 phổ vật chủ rộng, được tạo ra bằng cách chuyển yếu tố khởi đầu *lacUV5* và operon *ispFGH-idi-ispA* của pAM409 vào vector pAM257. Vector pAM257 được tạo ra như sau: locus RK2 *par* được khuếch đại từ ADN plasmit RK2 (xem ấn phẩm: Meyer et al. (1975) *Science* 190 1226-1228) sử dụng các môi 9-156A (SEQ ID NO: 33) và 9-156B (SEQ ID NO. 34), sản

phẩm PCR 2,6 kb được phân giải hoàn toàn sử dụng enzym giới hạn *AatII* và *XhoI*, và đoạn ADN được gắn vào plasmit chứa gốc sao chép p15 và gen kháng cloramphenicol từ vector pZA31-luc (xem ấn phẩm: Lutz and Bujard (1997) *Nucl Acid Res* 25:1203-1210), để thu được plasmit pAM37-*par*, pAM37-*par* được phân giải hoàn toàn sử dụng enzym giới hạn *SacI* và *HindIII*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN chứa locus RK2 *par* và gen kháng cloramphenicol được tách bằng gel, và đoạn ADN đã được tách riêng được gắn vào vị trí *SacI HindIII* của mini-RK2 replicon pRR10 (xem ấn phẩm: Roberts et al. (1990) *J Bacteriol.* **172**:6204-6216), để thu được vector pAM133; pAM133 được phân giải hoàn toàn sử dụng enzym giới hạn *BglII* và *HindIII*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 6,4 kb khuyết gen kháng ampicilin và gốc tiếp hợp *oriT* được tách bằng gel, và đoạn ADN đã được tách riêng này được gắn với đoạn ADN được tạo ra bằng cách tổng hợp chứa nhiều vị trí tách dòng mà chứa vị trí enzym giới hạn *PciI* và *XhoI*, để thu được vector pAM257. Plasmit biểu hiện pAM409 được phân giải hoàn toàn sử dụng enzym giới hạn *XhoI* và *PciI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, và đoạn ADN khoảng 4,4 kb được tách bằng gel. Vector pAM257 được phân giải hoàn toàn sử dụng enzym giới hạn *XhoI* và *PciI*, và đoạn ADN đã được tách riêng chứa yếu tố khởi đầu *lacUV5* và operon *ispFGH-idi-ispA* được gắn vào vị trí enzym giới hạn *XhoI PciI* của vector pAM257, để thu được plasmit biểu hiện pAM424 (xem Fig.5 thể hiện bản đồ plasmit).

Ví dụ 5

Ví dụ này mô tả phương pháp tạo ra plasmit biểu hiện mã hóa enzym chuyển hoá FPP hoặc GPP.

Plasmit biểu hiện pTrc99A-ADS được tạo ra bằng cách chèn trình tự nucleotit mã hóa amorpha-4,11-dien synthaza ("ADS") vào vector pTrc99A. Trình tự amorpha-4,11-dien synthaza được tạo ra bằng cách tổng hợp, sao cho trong quá trình dịch mã, trình tự axit amin sẽ giống với trình tự được mô tả trong ấn phẩm Merke et al. (2000) *Ach. Biochem. Biophys.* **381**:173-180, sao

cho trình tự nucleotit mã hóa amorpho-4,11-dien synthaza được tối ưu hóa để biểu hiện trong *Escherichia coli*, và sao cho trình tự nucleotit tạo sườn bởi vị trí enzym giới hạn 5' *NcoI* và 3' *XmaI* (xem Patent Mỹ số 7,192,751). Trình tự nucleotit được phân giải hoàn toàn sử dụng enzym giới hạn *NcoI* và *XmaI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 1,6 kb được tách bằng gel, và đoạn ADN đã được tách riêng này được chèn vào vị trí enzym giới hạn *NcoI XmaI* của vector pTrc99A (xem ấn phẩm: Amman et al. (1985) *Gene* 40:183-190), để thu được plasmit biểu hiện pTrc99A-ADS (xem Fig.6 thể hiện bản đồ plasmit).

Plasmit biểu hiện pAM113 là dẫn xuất kháng cloramphenicol của pTrc99A-ADS. Plasmit này được tạo ra bằng cách khuếch đại bằng PCR gen kháng cloramphenicol từ vector pZA31-luc (xem ấn phẩm: Lute and Bujard (1997) *Nucl Acids Res* 25: 1203-1210) sử dụng các môi được phosphoryl hoá đầu 5' 19-137 cml-pAM37-AS (SEQ ID NO: 35) và 19-137 cml-pAM37-S (SEQ ID NO: 36), và chèn sản phẩm PCR 920bp vào vị trí enzym giới hạn *FspI* của plasmit biểu hiện pTrc99A-ADS, để thu được plasmit biểu hiện pAM113.

Plasmit biểu hiện pC9 được tạo ra bằng cách chèn đoạn ADN hệ gen của *Bacillus subtilis* 6051 chứa trình tự mã hóa của gen *nudF* và các trình tự hệ gen ngược dòng (số đăng ký ngân hàng gen Z99116 REGION: 49364..48548) vào vector pTrc99A (xem ấn phẩm: Amann et al. (1988) *Gene* 69:301-315). Plasmit biểu hiện pNudF-H được tạo ra bằng cách chèn trình tự mã hóa của gen *nudF* của *Bacillus subtilis* 6051 (số đăng ký ngân hàng gen Z99116 REGION: 49105..48548) vào vector pTrc99A. Plasmit biểu hiện pyhfR được tạo ra bằng cách chèn trình tự mã hóa của gen *yhfR* của *Bacillus subtilis* 6051 (số đăng ký ngân hàng gen Z99109 REGION: 97583..97002) vào vector pTrc99A.

Plasmit biểu hiện pAM373 được tạo ra bằng cách chèn trình tự nucleotit mã hóa β -farnesen synthaza ("FSB") của *Artemisia annua* (số đăng ký ngân hàng gen AY835398), được tối ưu hoá codon để biểu hiện trong *Escherichia coli*, vào vector pTrc99A. Trình tự nucleotit mã hóa β -farnesen synthaza này

được tạo ra bằng cách tổng hợp, và được khuếch đại bằng PCR từ cấu trúc tổng hợp ADN của nó nhờ sử dụng các môi thích hợp. Để tạo ra vị trí enzym giới hạn *NcoI* dẫn đầu trong sản phẩm PCR chứa trình tự mã hóa β -farnesen synthaza, codon mã hóa axit amin thứ hai trong trình tự polypeptit ban đầu (TCG mã hóa serin) được thay thế bằng codon mã hóa axit aspartic (GAC) trong mỗi PCR đầu 5' (SEQ ID NO: 37). Sản phẩm PCR tạo ra này được phân giải một phần bằng cách sử dụng enzym giới hạn *NcoI*, và được phân giải hoàn toàn sử dụng enzym giới hạn *SacI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 1,7 kb chứa trình tự mã hóa β -farnesen synthaza được tách bằng gel, và đoạn ADN đã được tách riêng được gắn vào vị trí enzym giới hạn *NcoI SacI* của vector pTrc99A, để thu được plasmit biểu hiện pAM373 (xem Fig.6 là bản đồ plasmit).

Các plasmit biểu hiện pTrc99A-FSA, pTrc99A-GTS, pTrc99A-PS, pTrc99A-TS được tạo ra bằng cách chèn đoạn ADN chứa trình tự nucleotit mã hóa α -farnesen synthaza ("FSA"), γ -terpinen synthaza ("GTS"), α -pinen synthaza ("APS"), hoặc terpinolen synthaza ("TS") vào vector pTrc99A. Đoạn ADN chèn được tạo ra bằng cách tổng hợp, dùng làm khuôn, ví dụ, trình tự mã hóa của gen α -farnesen synthaza của *Picea abies* (số đăng ký ngân hàng gen AY473627, REGION: 24..1766), trình tự mã hóa của gen β -farnesen synthaza của *Artemisia annua* (số đăng ký ngân hàng gen AY835398), trình tự mã hóa của gen γ -terpinen synthaza của *Citrus limon* (số đăng ký ngân hàng gen AF514286 REGION: 30..1832), trình tự mã hóa của gen α -pinen synthaza của *Abies grandis* (số đăng ký ngân hàng gen U87909, REGION: 6..1892) hoặc của *Pinus taeda* (số đăng ký ngân hàng gen AF543530 REGION: 1..1887), hoặc trình tự mã hóa của gen terpinolen synthaza của *Ocimum basilicum* (số đăng ký ngân hàng gen AY693650) hoặc của *Pseudotsuga menziesii* (số đăng ký ngân hàng gen AY906866 REGION: 10..1887) hoặc của *Abies grandis* (số đăng ký ngân hàng gen AF 139206), tất cả các trình tự nucleotit này được tối ưu hoá codon để biểu hiện trong *Escherichia coli*. Các đoạn ADN cho FSA được khuếch đại bằng PCR từ cấu trúc tổng hợp ADN của nó sử dụng các trình tự

mỗi SEQ ID NO: 39 và SEQ ID NO: 40. Sản phẩm PCR tạo ra này được phân giải hoàn toàn sử dụng enzym giới hạn *NcoI* và *SacI*, hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 1,7kb chứa trình tự mã hóa α -farnesen synthaza được tách bằng gel, và đoạn ADN đã được tách riêng được gắn vào vị trí enzym giới hạn *NcoI SacI* của vector pTrc99A, để thu được plasmit biểu hiện pTrc99A-FSA (xem Fig.6 thể hiện bản đồ plasmit). Các đoạn ADN cho GTS, APS, và TS được thiết kế để tạo sườn bởi vị trí enzym giới hạn *XmaI* dẫn đầu và vị trí enzym giới hạn cuối cùng, và được tách dòng vào các vị trí enzym giới hạn tương thích của vector tách dòng như vector gốc pUC hoặc pACYC chuẩn, từ vector này chúng có thể lại được giải phóng bằng cách phân giải hoàn toàn cấu trúc tổng hợp ADN sử dụng enzym giới hạn *XbaI* và *XmaI*, phân tách hỗn hợp phản ứng này bằng phương pháp điện di gel, và tách bằng gel đoạn ADN mã hóa từ terpen synthaza 1,7 đến 1,9. Các đoạn ADN đã được phân tách được gắn vào vị trí enzym giới hạn *XmaI XbaI* của vector pTrc99A (xem ấn phẩm: Amman et al., *Gene* 40:183-190 (1985)), để thu được plasmit pTrc99A-GTS, pTrc99A-APS, hoặc pTrc99A-TS (xem Fig.6 thể hiện bản đồ plasmit).

Các plasmit biểu hiện pRS425-FSA và pRS425-FSB được tạo ra bằng cách chèn lần lượt trình tự nucleotit mã hóa α -farnesen synthaza ("FSA") hoặc β -farnesen synthaza ("FSB"), vào vector pRS425-Gal1 (xem ấn phẩm: Mumberg et. al. (1994) *Nucl Acids Res* 22(25): 5767-5768). Đoạn chèn trình tự nucleotit được tạo ra bằng cách tổng hợp, sử dụng để làm khuôn, ví dụ, trình tự mã hóa của gen α -farnesen synthaza của *Picea abies* (số đăng ký ngân hàng gen AY473627, REGION:24..1766) hoặc của gen β -farnesen synthaza của *Artemisia annua* (số đăng ký ngân hàng gen AY835398), được tối ưu hoá codon để biểu hiện trong *Saccharomyces cerevisiae*. Trình tự nucleotit đã được tạo ra bằng cách tổng hợp tạo sườn bởi vị trí 5' *BamHI* và vị trí 3' *XhoI*, và do đó có thể được tách dòng vào các vị trí enzym giới hạn tương thích của vector tách dòng như vector gốc pUC hoặc pACYC chuẩn. Trình tự nucleotit đã được tạo ra bằng cách tổng hợp được phân lập bằng cách phân giải hoàn toàn cấu

trúc tổng hợp ADN sử dụng enzym giới hạn *BamHI* và *XhoI*. Hỗn hợp phản ứng này được phân tách bằng phương pháp điện di gel, đoạn ADN khoảng 1,7kb chứa trình tự mã hóa α -farnesen synthaza hoặc β -farnesen synthaza được tách bằng gel, và đoạn ADN đã được tách riêng này được gắn vào vị trí enzym giới hạn *BamHI* *XhoI* của vector pRS425-Gal1, để thu được lần lượt plasmit biểu hiện pRS425-FSA hoặc pRS425-FSB.

Các plasmit biểu hiện pTrc99A-LLS, pTrc99A-LMS, pTrc99A-BPS, pTrc99A-PHS, pTrc99A-CS, và pTrc99A-SS được tạo ra bằng cách chèn trình tự nucleotit mã hóa linalool synthaza ("LLS"), limonen synthaza ("LMS"), β -pinen synthaza ("BPS"), β -phellandren ("PHS"), caren synthaza ("CS"), hoặc sabinin synthaza ("SS") vào vector pTrc99A. Đoạn chèn trình tự nucleotit này được tạo ra bằng cách tổng hợp, dùng làm khuôn, ví dụ, trình tự mã hóa của gen linalool synthaza của *Artemisia annua* (số đăng ký ngân hàng gen AF154124, REGION: 13..1764), trình tự mã hóa của gen limonen synthaza của *Abies grandis* (số đăng ký ngân hàng gen AF006193 REGION: 73..1986), trình tự mã hóa của β -pinen synthaza của *Artemisia annua* (số đăng ký ngân hàng gen AF276072 REGION: 1..1749), trình tự mã hóa của gen β -phellandren synthaza của *Abies grandis* (số đăng ký ngân hàng gen AF139205 REGION: 34..1926), trình tự mã hóa của gen caren synthaza của *Salvia stenophylla* (số đăng ký ngân hàng gen AF527416 REGION: 78..1871), hoặc trình tự mã hóa của gen sabinin synthaza của *Salvia officinalis* (số đăng ký ngân hàng gen AF051901 REGION: 26..1798). Trình tự nucleotit mã hóa β -pinen, sabinin, và β -phellandren synthaza tạo sườn bởi vị trí enzym giới hạn *XmaI* dẫn đầu và vị trí enzym giới hạn *XbaI* cuối cùng, trình tự nucleotit mã hóa linalool và caren synthaza tạo sườn bởi vị trí enzym giới hạn *NcoI* dẫn đầu và vị trí enzym giới hạn *XmaI* cuối cùng, và trình tự nucleotit mã hóa limonen synthaza tạo sườn bởi vị trí enzym giới hạn *NcoI* dẫn đầu và vị trí enzym giới hạn *PstI* cuối cùng. Các cấu trúc tổng hợp ADN được phân giải hoàn toàn sử dụng enzym giới hạn là *XmaI* và *XbaI* (đối với các cấu trúc β -pinen, sabinin, và β -phellandren synthaza), enzym giới hạn là *NcoI* và *XmaI* (đối với các cấu trúc linalool và

caren synthaza), hoặc enzym *XbaI* và *PstI* (đối với cấu trúc limonen synthaza). Hỗn hợp phản ứng được phân tách bằng phương pháp điện di gel, các đoạn ADN khoảng từ 1,7 đến 1,9 kb được tách bằng gel, và các đoạn ADN đã được phân lập được gắn vào vị trí enzym giới hạn *XmaI XbaI* (đối với đoạn chèn β -pinen, và β -phellandren synthaza), vị trí enzym giới hạn *NcoI XmaI* (đối với đoạn chèn linalool và caren synthaza), hoặc vị trí enzym giới hạn *XbaI PstI* (đối với đoạn chèn limonen synthaza) của vector pTrc99A, để thu được các plasmit biểu hiện pTrc99A-LLS, pTrc99A-LMS, pTrc99A-BPS, pTrc99A-PHS, pTrc99A-CS, và pTrc99A-SS (xem Fig.6 là bản đồ plasmit).

Ví dụ 6

Ví dụ này mô tả việc tạo ra các chủng chủ *Escherichia coli* để sử dụng theo sáng chế.

Như được nêu chi tiết trong Bảng 1, các chủng chủ này được tạo ra bằng cách biến nạp bằng quá trình hoá học tế bào *Escherichia coli* cha mẹ có khả năng bị biến nạp với một hoặc nhiều plasmit biểu hiện của các Ví dụ từ 1 đến 5.

Bảng 1. Các chủng chủ *E. coli*

Chủng chủ	Chủng E.coli cha mẹ	Các plasmit biểu hiện	Chọn chất kháng sinh
B32	DH1	pMevT	100µg/ml carbenixilin
B292	B	pMBIS	5µg/ml tetracyclin
B210	DP	pTrc99A-ADS	34µg/ml cloramphenicol
B153	DH1	pAM97	100µg/ml carbenixilin
B282	DP	pTrc99A-ADS	34µg/ml cloramphenicol
B255	DH1	pAM128	100µg/ml carbenixilin
B256	DP	pAM113	34µg/ml cloramphenicol
B86	DH1	pAM52 pMBIS pTrc99A-ADS	50µg/ml kanamycin 100µg/ml carbenixilin 5µg/ml tetracyclin
B61	DH1	pAM25 pBBR1MCS-3	

		pTrc99A	
B62		pAM34 pBBR1MCS-3 pTrc99A	
B003	DH10B	pTrc99A-ADS	100µg/ml carbenixilin
B617		pAM408 pTrc99A-ADS-	100µg/ml carbenixilin 50µg/ml kanamycin
B618		pAM424 pTrc99A-ADS	100µg/ml carbenixilin 35µg/ml cloramphenicol
B619		pAM408 pAM424 pTrc99A-ADS	100µg/ml carbenixilin 50µg/ml kanamycin 35µg/ml cloramphenicol
B650	DH10B	pAM373	100µg/ml carbenixilin
B651		pAM408 pAM373	100µg/ml carbenixilin 50µg/ml kanamycin
B652		pAM424 pAM373	100µg/ml carbenixilin 35µg/ml cloramphenicol
B653		pAM408 pAM424 pAM373	100µg/ml carbenixilin 50µg/ml kanamycin 35µg/ml cloramphenicol
B286	DH1	pAM97-MevB pC9	100µg/ml carbenixilin 34µg/ml
B287		pAM97-MevB pnudF-H	cloramphenicol.
B288		pAM97-MevB pyhfR	
B291		pAM97-MBI pyhfR	
B592	DH1	pMevT pMBIS pTrc99A-FSA	100µg/ml carbenixilin 34µg/ml cloramphenicol
B552		pMevT pMBIS pAM373	5µg/ml tetraxyclin
Tế bào chủ của ví dụ 21 (sản xuất GTS, APS, TS)		pMevT pMBIS-gpps pTrc99A-GTS hoặc – APS hoặc -TS	
Tế bào chủ của ví		pMevT	100µg/ml carbenixilin

dự 21 (sản xuất LLS, LMS, BPS, PHS, CS, SS)		pMBIS-gpps pTrc99A-LLS hoặc – LMS hoặc –BPS hoặc – PHS hoặc –CS hoặc – SS	34µg/ml cloramphenicol 5µg/ml tetracyclin
---	--	---	---

Các thể biến nạp tế bào chủ được chọn trên thạch Luria Bertoni (LB) chứa thuốc kháng sinh như được nêu chi tiết trong Bảng 1. Các khuẩn lạc đơn được chuyển từ thạch LB sang ống nuôi cấy chứa 5ml môi trường LB lỏng và thuốc kháng sinh. Các thể biến nạp tế bào chủ B003, B617, B618, B619, B650, B651, B652, và B653 được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút trong 30 giờ. Tất cả các thể biến nạp tế bào chủ khác được ủ ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi sự sinh trưởng đạt tới pha ổn định. Các tế bào này được làm thích nghi với môi trường tối thiểu bằng cách chuyển các tế bào này qua từ 4 đến 5 vòng liên tiếp của môi trường M9-MOPS chứa glucoza 0,8% và thuốc kháng sinh (xem Bảng 2 mô tả thành phần của môi trường M9-MOPS). Các tế bào được bảo quản ở nhiệt độ -80°C trong lọ bảo quản lạnh sâu với lượng nhỏ giống nuôi cấy gốc 1ml được điều chế từ 400µl glyxerol 50% vô trùng và 600µl dịch nuôi cấy lỏng.

Bảng 2 – Thành phần của môi trường nuôi cấy M9-MOPS

Thành phần	Khối lượng (trên một lít)
Na ₂ HPO ₄ 7H ₂ O	12,8g
KH ₂ PO ₄	3g
NaCl	0,5g
NH ₄ Cl	1g
MgSO ₄	2mmol
CaCl ₂	0,1mmol
Thiamin	0,1µg
Đệm MOPS pH 7,4	100mmol
(NH ₃) ₆ Mo ₇ O ₂₄ 4H ₂ O	3,7µg
H ₃ BO ₄	25µg

CoCl ₂	7,1µg
CuSO ₄	2,4µg
MnCl ₂	16µg
ZnSO ₄	2,9µg
FeSO ₄	0,28µg

Ví dụ 7

Ví dụ này chứng minh độ ổn định của plasmit biểu hiện trong điều kiện không có thuốc kháng sinh trong chủng chủ *Escherichia coli* chứa plasmit biểu hiện chứa hệ duy trì, phân tách và sao chép của plasmit RK2.

Giống nuôi cấy mầm của chủng chủ được tạo nên bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của chủng này vào bình 125ml chứa 40ml M9-MOPS, glucoza 2%, dịch chiết nấm men 0,5%, và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và bằng cách nuôi các dịch nuôi cấy này qua đêm.

Giống nuôi cấy mầm này được sử dụng để cấy truyền ở giá trị OD₆₀₀ ban đầu khoảng 0,05, vào hai bình 250ml, mỗi bình chứa 40 ml môi trường M9-MOPS, glucoza 2%, và dịch chiết nấm men 0,5%. Dịch nuôi cấy #1 còn chứa 100µg/ml carbenixilin và 34µg/ml cloramphenicol. Dịch nuôi cấy #2 không nhận bất kỳ thuốc kháng sinh nào. Cả hai dịch nuôi cấy này được ủ ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi chúng đạt tới OD₆₀₀ khoảng 0,2, ở thời điểm này việc sản xuất amorpho-4,11-dien ở tế bào chủ được cảm ứng bằng cách bổ sung 40µl IPTG 1M vào môi trường nuôi cấy. Tại thời điểm cảm ứng, các dịch nuôi cấy này được phủ bằng 8ml chất phủ hữu cơ để thu amorpho-4,11-dien. Các mẫu được lấy định kỳ trong thời gian tổng cộng khoảng 72 giờ. Việc sản xuất amorpho-4,11-dien bởi chủng chủ trong 2 dịch nuôi cấy được xác nhận bằng GC/MS như nêu trong Ví dụ 10.

Để đánh giá độ ổn định của plasmit trong hai dịch nuôi cấy tế bào, mẫu của mỗi dịch nuôi cấy được lấy ra ở thời điểm 72 giờ và được cấy thành vạch vào đĩa thạch LB (không chứa thuốc kháng sinh). Sau khi ủ qua đêm ở nhiệt độ 37°C, 50 khuẩn lạc riêng rẽ thu được từ mỗi dịch nuôi cấy được cấy lặp lại trên

đĩa thạch LB chứa thuốc kháng sinh (34 µg/ml cloramphenicol, 100 µg/ml carbenixilin) và đĩa thạch LB trừ thuốc kháng sinh (không chứa chất kháng sinh). Sau khi ủ qua đêm thêm lần nữa ở nhiệt độ 37°C, mỗi đĩa thạch LB chứa thuốc kháng sinh và đĩa thạch LB không chứa thuốc kháng sinh được phát hiện thấy chứa khoảng 50 khuẩn lạc, chứng tỏ rằng sự lưu giữ plasmit cả trong điều kiện chứa và không chứa thuốc kháng sinh trong môi trường nuôi cấy là khoảng 100%.

Ví dụ 8

Ví dụ này chứng minh hoạt tính đặc hiệu và độ ổn định của HMGR của *Enterococcus faecalis* tăng lên so với của tHMGR của *Saccharomyces cerevisiae* ở chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ B61 và B62 được tạo nên bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình 125 ml chứa 20ml môi trường M9-MOPS, glucoza 0,8%, và thuốc kháng sinh như được nêu chi tiết trong Bảng 5, và bằng cách nuôi cấy các dịch nuôi cấy này đến bão hoà. Các giống nuôi cấy mầm được pha loãng tỉ lệ 1 : 100 vào 140ml môi trường mới trong bình 500ml, và lại được sinh trưởng đến giá trị OD₅₅₀ khoảng 0,1, ở thời điểm này việc sản xuất amorpho-4,11-dien được cảm ứng bằng cách bổ sung 140µl IPTG 1M vào mỗi dịch nuôi cấy. Tại thời điểm 4, 12, 20, 28, 36, và 49 giờ sau khi cảm ứng, các mẫu được lấy ra từ mỗi dịch nuôi cấy, và các tế bào được kết hạt bằng cách ly tâm. Các hạt nhỏ tế bào được làm đông lạnh nhanh trên đá khô, và sau đó được bảo quản ở nhiệt độ -80°C.

Để tiến hành thử nghiệm enzym, các hạt nhỏ tế bào được làm tan đông trên đá, và sau đó được cho dung giải bằng cách sử dụng Bugbuster (Novagen, Madison, WI) chứa hỗn hợp chất ức chế proteaza #3 (Calbiochem, San Diego, CA), benzonaza (20 µL over 5 ml bugbuster; Novagen, Madison, WI), và lysozym (30 µg/ml). Hoạt tính enzym của tHMGR ở *Saccharomyces cerevisiae* được thử nghiệm trong Tris HCl 50 mM (độ pH bằng 7,5), NADPH 0,2mM (Sigma, St. Louis, MO), và muối natri DL-3-hydroxy-3-methylglutaryl coenzym A (DL-3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA)) 0,3 mM

(Sigma, St. Louis, MO). Thử nghiệm này được bắt đầu bằng cách bổ sung dịch giải tế bào, và kiểm soát sự biến mất của NADPH bằng cách đo độ hấp thụ ở 340nm. Để giải thích về sự biến mất không đặc hiệu của NADPH, các kết quả thu được trong thử nghiệm đối chứng khuyết HMG-CoA được trừ đi kết quả thu được trong mẫu thử nghiệm. Hoạt tính enzym của HMGR ở *Enterococcus faecalis* được xác định một cách tương tự chỉ khác là dung dịch đệm thử nghiệm chứa đệm kali phosphat 100mM (độ pH bằng 6,5), NADPH 0,4mM, EDTA 1,0mM, và KCl 100mM.

Thử nghiệm protein được tiến hành bằng phương pháp Bradford ((1976) *Anal Biochem.* 72:248-254). Hoạt tính đặc hiệu được tính là Δ nmol NADPH/phút/mg protein.

Như được thể hiện trên Fig.8, HMGR ở *Enterococcus faecalis* thể hiện hoạt tính đặc hiệu cao hơn và độ ổn định tăng so với tHMGR ở *Saccharomyces cerevisiae*.

Ví dụ 9

Ví dụ này mô tả mối tương quan giữa OD₆₀₀ với khối lượng tế bào khô (Dry Cell Weight ("DCW")).

Để thu được mối tương quan giữa DCW và OD₆₀₀, chủng đại diện, B32, được sinh trưởng trong các quy trình mật độ tế bào cao tương tự với các quy trình được mô tả trong các ví dụ từ 10 đến 14. Các mẫu được lấy trong toàn bộ quá trình chạy, và OD₆₀₀ và DCW được xác định đối với mỗi mẫu. Để xác định DCW, các tế bào được kết hạt và dịch nổi được loại bỏ. Các hạt nhỏ tế bào được rửa một lần bằng nước, và sau đó được sấy khô trong lò ở nhiệt độ 80°C trong ít nhất 3 ngày. Các ống nghiệm chứa các hạt nhỏ tế bào được cân, khối lượng của ống nghiệm được trừ đi khối lượng đã được xác định, và khối lượng còn lại được chia cho thể tích ban đầu của mỗi mẫu (0,0015l) để thu được DCW.

Fig.9 thể hiện tương quan giữa DCW và OD₆₀₀ được xác định trong các thử nghiệm này.

Ví dụ 10

Ví dụ này chứng minh việc sản xuất amorph-4,11-dien tăng trong các chủng chủ *Escherichia coli* biểu hiện HMGR và HMGS ở *Staphylococcus aureus* so với các chủng chủ biểu hiện tHMGR và HMGS ở *Saccharomyces cerevisiae*.

Giống nuôi cấy mầm của các chủng chủ B32, B153, B210, B282, B292, B86, B255, và B256 được tạo nên bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình 125ml riêng biệt chứa 25ml môi trường M9-MOPS, 0,8% glucoza, và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và bằng cách nuôi các dịch nuôi cấy này qua đêm.

Giống nuôi cấy mầm được sử dụng để cấy truyền vào các bình 250ml riêng biệt có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml môi trường M9-MOPS, glucoza 2%, và thuốc kháng sinh. Các dịch nuôi cấy này được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi chúng đạt tới OD₆₀₀ vào khoảng 0,2, ở thời điểm này việc sản xuất amorph-4,11-dien ở tế bào chủ được cảm ứng bằng cách bổ sung 40µl IPTG 1M vào môi trường nuôi cấy. Dịch nuôi cấy được phủ bằng 8ml chất phủ hữu cơ (ví dụ, dodecan, metyl oleat hoặc isopropyl myristat). Mẫu của chất phủ hữu cơ và nước dùng nuôi cấy được lấy một lần mỗi ngày trong 72 giờ. Các mẫu canh nuôi cấy được xác định OD₆₀₀. Nồng độ amorph-4,11-dien được xác định bằng cách chuyển 5µl chất phủ hữu cơ vào bình thủy tinh sạch chứa 500 µl etyl axetat được gắn beta hoặc trans-caryophylen vào để làm tiêu chuẩn bên trong.

Các mẫu chất phủ hữu cơ/etyl axetat được phân tích trên sắc ký khí/máy đo phổ khối Hewlett-Packard 6890 (GC/MS) bằng cách chỉ quét đối với hai ion, ion phân tử (204 m/z) và ion 189 m/z, như nêu trong ấn phẩm Martin et al. (2001) *Biotechnol. Bioeng.* 75:497-503. Để xúc tiến thời gian tiến hành, chương trình nhiệt độ và ma trận cột được biến đổi để đạt được độ phân giải đỉnh tối ưu và tổng thời gian tiến hành ngắn nhất. 1 µl mẫu được phân tách trên GC sử dụng cột DB-XLB (có sẵn từ Agilent Technologies, Inc., Palo Alto, CA) và sử dụng khí chất mang heli. Chương trình nhiệt độ cho phép phân tích là

như sau: 100°C trong 0,75 phút, tăng nhiệt độ với tốc độ 60°C/phút đến nhiệt độ khoảng 300°C, và duy trì ở nhiệt độ 300°C trong 0,5 phút. Các mẫu đã được phân tách được phân tích bằng máy dò chọn lọc khối 5973 mô hình Hewlett-Packard mà kiểm tra các ion 189 và 204 m/z. Phổ khối trước chứng tỏ rằng sản phẩm amorph-4,11-dien synthaza là amorph-4,11-dien, và chúng tỏ rằng amorph-4,11-dien có thời gian lưu là 3,7 phút bằng cách sử dụng Protocol GC này. Beta- hoặc trans-caryophylen được sử dụng để làm tiêu chuẩn bên trong cho việc định lượng. Độ chuẩn amorph-4,11-dien được tính toán bằng cách sử dụng tỷ lệ của diện tích tiêu chuẩn bên trong so với amorph-4,11-dien dựa trên đường cong định chuẩn định lượng của amorph-4,11-dien tinh khiết (0,63-10mg/l KJF17- 109-3) trong etyl axetat được gắn caryophylen.

Như được thể hiện trên các Fig.10A và 10B, các chủng B153 và B282, biểu hiện HMGR và HMGS ở *Staphylococcus aureus*, mức amorph-4,11-dien cao được sản xuất ra so với các chủng B32, B210, B255, B256, và B292, biểu hiện tHMGR và HMGS ở *Saccharomyces cerevisiae*.

Ví dụ 11

Ví dụ này chứng minh việc sản xuất amorph-4,11-dien tăng bởi chủng chủ ở *Escherichia coli* được sinh trưởng ở nhiệt độ dưới điểm cực thuận.

Giống nuôi cấy mầm của chủng chủ B32 được tạo nên bằng cách bổ sung 0,5ml lượng nhỏ giống nuôi cấy gốc của chủng vào bình 250ml chứa 50ml môi trường M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và bằng cách nuôi các dịch nuôi cấy này qua đêm ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút.

Giống nuôi cấy mầm được sử dụng để cấy truyền vào bốn bình 250ml có giá trị OD₆₀₀ ban đầu khoảng 0,05, mỗi bình chứa 40ml môi trường lên men theo mẻ (xem Bảng 6 về thành phần môi trường), đệm MOPS 100 mM pH 7,1, và thuốc kháng sinh. Dịch nuôi cấy được ủ trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút ở nhiệt độ 30°C hoặc 37°C cho đến khi chúng đạt tới OD₆₀₀ nằm trong khoảng từ 0,18 đến 0,22, ở thời điểm này việc sản xuất amorph-4,11-dien ở tế bào chủ được cảm ứng bằng cách bổ sung 40µl IPTG 1M vào môi

trường nuôi cấy. Vào thời điểm cảm ứng, dịch nuôi cấy được phủ bằng 8ml chất phủ hữu cơ để thu amorpho-4,11-dien. Các mẫu được lấy một lần mỗi ngày, và được phân tích như nêu trong Ví dụ 10.

Như được thể hiện trên Fig.11A và 11B, quá trình lên men ở nhiệt độ 30°C không ảnh hưởng đến sự sinh trưởng tế bào, tuy nhiên lại làm tăng khoảng 50% việc sản xuất đặc trưng amorpho-4,11-dien bởi chủng chủ *Escherichia coli*.

Ví dụ 12

Ví dụ này chứng minh việc sản xuất lượng amorpho-4,11-dien tăng lên bởi chủng chủ *Escherichia coli* được sinh trưởng trong điều kiện nguồn cacbon bị hạn chế.

Giống nuôi cấy mầm của chủng chủ B32 cho các mẻ lên men 050608-1 và 050629-1 được tạo nên bằng cách bổ sung 0,25µl lượng nhỏ giống nuôi cấy gốc của chủng vào bình 250ml chứa 50ml môi trường M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và bằng cách ủ dịch nuôi cấy này ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi nó đạt tới OD₆₀₀ nằm trong khoảng từ 1 đến 2.

Giống nuôi cấy mầm của chủng chủ B32 cho các mẻ lên men 060403-3 được tạo nên bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của chủng vào bình 250ml chứa 50 ml môi trường M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và bằng cách ủ dịch nuôi cấy qua đêm ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút. Giống nuôi cấy mầm được sử dụng để cấy truyền vào trong các bình 250ml giá trị OD₆₀₀ ban đầu khoảng 1 chứa 40ml môi trường M9-MOPS và thuốc kháng sinh, và dịch nuôi cấy này lại được ủ ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi nó đạt tới OD₆₀₀ nằm trong khoảng từ 3 đến 5.

Đối với tất cả các quy trình lên men, KH₂PO₄, K₂HPO₄·3H₂O, EDTA, axit xitric, và (NH₄)₂SO₄ được vô trùng bằng nhiệt trong bình phản ứng sinh học (2L Applikon Bioconsole ADI 1025s chứa bộ điều khiển ADI 1010,

Applikon Biotechnology, Foster City, CA). Các thành phần môi trường còn lại được vô trùng bằng cách lọc để làm dung dịch gốc và được tiệt trùng qua đĩa đỡ. Bảng 3 thể hiện thành phần môi trường cuối cùng đối với mẻ lên men 050608-1 và 050629-1. Bảng 4 thể hiện thành phần môi trường cuối cùng đối với mẻ lên men 060403-3. Thể tích ban đầu đối với mẻ 050608-1 là 0,8 l, thể tích ban đầu đối với mẻ 050629-1 là 1,2l và thể tích ban đầu đối với mẻ 060403-3 là 1l. Tất cả các mẻ đều được cấy truyền bằng cách nhiễm 50ml của giống nuôi cấy nấm thông qua đĩa đỡ.

Bảng 3 – Thành phần môi trường lên men của quá trình lên men

Mẻ 050608-1 và 050629-1

Thành phần	Môi trường theo mẻ (cho một lít)	Dung dịch cấp (cho một l)
Glucosa	5g	590-650g
KH ₂ PO ₄	4,2 g	-
K ₂ HPO ₄ 3H ₂ O	15,7g	-
Axit xitric	1,7g	-
(NH ₄) ₂ SO ₄	2g	-
MgSO ₄ 7H ₂ O	1,2g	12g
EDTA	8,4mg	13g
CoCl ₂ 6H ₂ O	0,25mg	0,4mg
MnCl ₂ 4H ₂ O	1,5mg	2,35mg
CuCl ₂ 2H ₂ O	0,15mg	0,25mg
H ₃ BO ₄	0,3mg	0,5mg
Na ₂ MoO ₄ 2H ₂ O	0,25mg	0,4mg
Zn(CH ₃ COO) ₂ 2H ₂ O	1,3mg	1,6mg
Sắt (III)xitrat hydrat	10,0mg	4,0mg
Thiamin HCl	4,5mg	-
Carbenixilin	100µg	100µg
Tetracyclin	5µg	5µg

Cloramphenicol	34µg	34µg
----------------	------	------

Bảng 4 – Thành phần môi trường lên men của quá trình lên men 060403-3

Thành phần	Môi trường theo mẻ (cho một lít)	Dung dịch cấp (cho một l)
Glucosa	15 g	650 g
KH ₂ PO ₄	4,2 g	-
K ₂ HPO ₄ 3H ₂ O	15,7 g	-
Axit xitric	1,7 g	-
(NH ₄) ₂ SO ₄	2 g	-
MgSO ₄ 7H ₂ O	1,2 g	12 g
EDTA	8,4 g	13 mg
CoCl ₂ 6H ₂ O	2,5 mg	4 mg
MnCl ₂ 4H ₂ O	15 mg	23,5 mg
CuCl ₂ 2H ₂ O	1,5 mg	2,5 mg
H ₃ BO ₄	3 mg	5 mg
Na ₂ MoO ₄ 2H ₂ O	2,5 mg	4 mg
Zn(CH ₃ COO) ₂ 2H ₂ O	13 mg	16 mg
Sắt (III)xitrat hydrat	100 mg	40 mg
Thiamin HCl	4,5 mg	-
Carbenixilin	100 µg	100 µg
Tetracyclin	5 µg	5 µg
Cloramphenicol	34 µg	34 µg

Đối với mẻ lên men 050608-1 (dư cacbon), sự nạp liệu được bắt đầu ở sự cảm ứng, và tốc độ nạp liệu được điều chỉnh bằng tay để cung cấp glucosa ở các nồng độ được thể hiện trên Fig.12C. Đối với mẻ lên men 050629-1 (hạn chế cacbon), sự nạp liệu được phân phối vào bình lên men theo quy trình được thể hiện trong Bảng 5. Đối với mẻ lên men 060403-3 (cacbon thấp nhất), sự

nạp liệu được bắt đầu một cách tự động khi viên glucoza ban đầu (15g) được sử dụng hết và oxy hoà tan hết. Đến giá trị tốc độ tối đa 27,6g/giờ, tốc độ nạp được tính theo phương trình sau:

$$m_s(t) = S(t_0)\mu e^{\mu(t-t_0)}$$

$$\mu = 0,12$$

$$S(t_0) = 15 \text{ g}$$

trong đó t_0 là thời gian ở đó glucoza ban đầu bị làm cạn kiệt. Trong quá trình tiến đến tốc độ tối đa, việc nạp glucoza bị giới hạn ở tốc độ 9,5g/giờ, và được giữ không đổi ở tốc độ này đối với phần còn lại của quá trình.

Bảng 5 – Quy trình nạp liệu đối với mẻ lên men 050629-1

Thời gian tiến hành (giờ)	Tốc độ nạp liệu glucoza (g/giờ)
0	0
7	0,37
10	0,74
12	1,11
14	1,48
16	2,22
18	2,96
20	3,69
22	4,80
24	5,91
31	7,39
33	5,54
47	3,69

Các mẻ 050608-1 và 050629-1 được tiến hành ở nhiệt độ 37°C. Dòng khí trong bình phản ứng sinh học được đặt ở 1-2l/phút; trị số độ pH được duy

trì ở 7 bằng cách sử dụng amoni hydroxit và/hoặc natri hydroxit; tốc độ lắc ban đầu là 500-600 vòng/phút; bọt được kiểm soát bằng chất chống tạo bọt B (Sigma-Aldich, St. Louis, MO); nồng độ oxy hoà tan được duy trì trên 30% bằng cách sử dụng tầng khuấy. Sau 5-6 giờ nuôi cấy, việc sản xuất amorpha-4,11-dien bởi tế bào chủ được cảm ứng bằng cách bổ sung 0,8 ml IPTG 1M vào mẻ 050608-1 và 1,2ml IPTG vào mẻ 050629-1. Trong quá trình cảm ứng, nhiệt độ nuôi cấy được giảm xuống 30°C.

Mẻ 060403-3 được thực hiện ở nhiệt độ 30°C. Dòng khí trong bình phản ứng sinh học được đặt ở 1-2 lít/phút; pH được duy trì ở 7 bằng cách sử dụng amoniac hydroxit. Oxy hoà tan được duy trì trên 30% bằng tầng khuấy và làm giàu oxy. Ở giá trị OD₆₀₀ khoảng 28 (19 giờ sau khi cấy truyền), việc sản xuất amorpha-4,11-dien bởi tế bào chủ được kích thích bằng cách bổ sung 1ml IPTG 1M.

Amorpha-4,11-dien được thu và được chiết theo hai quy trình khác nhau. Đối với các mẻ 050608-1 và 050629-1, amorpha-4,11-dien dễ bay hơi có mặt trong khí thải được thu bằng cách thải khí thải thông qua thiết bị rửa khí chứa 200ml heptanol. Sau đó, heptanol được pha loãng vào etyl axetat cho đến khi nồng độ amorpha-4,11-dien trong mẫu nằm trong khoảng từ 0,63mg/l đến 20mg/l. Đối với mẻ 060403-3, amorpha-4,11-dien được thu trong bình phản ứng sinh học bằng cách bổ sung 200ml chất phủ hữu cơ vào bình lên men ở thời điểm cảm ứng. Nồng độ sản phẩm được xác định bằng cách kết hợp 25µl cả hai mẻ cộng với chất phủ hữu cơ với 975µl axetonitril, lắc mẫu ở tốc độ lớn nhất trên thiết bị trộn Fisher Vortex Genie 2™ (Scientific Industries, Inc., Bohemia, NY) trong ít nhất 3 phút, loại tế bào khỏi mẫu bằng cách ly tâm, và pha loãng dung dịch axetonitril vào etyl axetat cho đến khi nồng độ amorpha-4,11-dien trong mẫu nằm trong khoảng từ 0,63 đến 20mg/l. Các mẫu etyl axetat được phân tích bằng GC/MS như nêu trong ví dụ 10.

Như được thể hiện trên các Fig.12A và 12B, mẻ lên men 050608-1 (dur cacbon) tạo ra mật độ tế bào tối đa thấp và việc sản xuất amorpha-4,11-dien thấp, một cách lần lượt, tương quan ít nhất một phần với sự tăng tương đối

nhánh các mức axetat (Fig.12D). Để so sánh, mẽ lên men 050629-1 (hạn chế cacbon) tạo ra việc sản xuất amorph-4,11-dien gia tăng (Fig.12B), và làm chậm sự bắt đầu sản xuất axetat. Các kết quả này thống nhất với giả thuyết rằng việc nạp glucoza dư dẫn đến việc sản xuất axetat nhanh và việc chết tế bào sớm.

Việc hạn chế glucoza khác như thu được bởi mẽ lên men 060403-3 (cacbon thấp nhất) tạo ra việc sản xuất axetat thấp trong khoảng hơn 100 giờ (Fig.12D), và mật độ tế bào tối đa và việc sản xuất amorph-4,11-dien cao hơn một cách đáng kể (các Fig.12A và 12B).

Ví dụ 13

Ví dụ này chứng tỏ việc sản xuất amorph-4,11-dien bởi chủng chủ *Escherichia coli* được sinh trưởng trong điều kiện nguồn cacbon bị hạn chế và ở nhiệt độ dưới điểm cực thuận.

Giống nuôi cấy mầm của chủng chủ B153 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của chủng vào bình 250ml chứa 50ml môi trường M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy này ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút đến OD₆₀₀ nằm trong khoảng từ 3,5 đến 4,5.

Bình phản ứng sinh học loại 2l (Thiết bị kiểm soát sinh học Biocontroller ADI 1010 có Bioconsole ADI 1025, Applikon Biotechnology, Foster City, CA) được thiết lập và được thực hiện theo cùng cách như nêu trong Ví dụ 12 đối với mẽ 060403-3, ngoại trừ là chủng và thời gian cảm ứng được thay đổi.

Việc sản xuất amorph-4,11-dien ở tế bào chủ được cảm ứng bằng cách bổ sung 1 ml IPTG 1 M vào môi trường nuôi cấy. Trong mẽ lên men được thể hiện trên Fig.13A, việc tổng hợp amorph-4,11-dien được cảm ứng ở giá trị OD₆₀₀ khoảng 2, trong khi bình lên men vẫn chứa glucoza dư. Trong mẽ lên men được thể hiện trên Fig.13B, việc tổng hợp amorph-4,11-dien được cảm

ứng ở giá trị OD₆₀₀ khoảng 33, giá trị này là sau khi bắt đầu nạp glucoza hạn chế.

Amorpha-4,11-dien được thu và được chiết theo hai quy trình khác nhau. Đối với mẻ lên men được thể hiện trên Fig.13 A, amorpha-4,11-dien dễ bay hơi có mặt trong khí thải được thu bằng cách thải khí thải đến thiết bị rửa khí chứa 200ml heptanol. Sau đó, heptanol được pha loãng vào etyl axetat cho đến khi nồng độ amorpha-4,11-dien trong mẫu nằm trong khoảng từ 0,63 đến 20 mg/l. Đối với mẻ lên men được thể hiện trên Fig.13B, amorpha-4,11-dien được thu bằng cách bổ sung 200ml chất phủ hữu cơ vào bình lên men ở thời điểm cảm ứng.

Amorpha-4,11-dien được chiết từ môi trường nuôi cấy bằng cách kết hợp 25 μ l cả hai mẻ với 975 μ l axetonitril, lắc mẫu ở tốc độ lớn nhất trên thiết bị trộn Fisher Vortex Genie 2TM (Scientific Industries, Inc., Bohemia, NY) trong ít nhất 3 phút, loại tế bào khỏi mẫu bằng cách ly tâm, và pha loãng dung dịch axetonitril vào etyl axetat cho đến khi nồng độ amorpha-4,11-dien trong mẫu nằm trong khoảng từ 0,63 đến 20mg/l. Mẫu etyl axetat được phân tích bằng GC/MS như nêu trong ví dụ 10. Đối với mẻ lên men được thể hiện trên Fig.13 A, tổng lượng amorpha-4,11-dien thu được bằng cách cộng các lượng có mặt trong khí thải và trong môi trường nuôi cấy, và chia tổng này cho thể tích bình lên men.

Quá trình lên men được thể hiện trên Fig.13 A đạt tới giá trị OD₆₀₀ tối đa là 93 và nồng độ amorpha-4,11-dien tối đa là 3,2g/l. Trong khi đó, quá trình lên men được thể hiện trên Fig.13B đạt tới giá trị OD₆₀₀ tối đa là 245 và nồng độ amorpha-4,11-dien tối đa là 15g/l. Sự giải thích phù hợp về sự khác nhau trong sự sinh trưởng của giống nuôi cấy và mức sản xuất amorpha-4,11-dien được quan sát thấy ở hai giống nuôi cấy là ở chỗ trong mẻ lên men được thể hiện trên Fig.13A việc sản xuất amorpha-4,11-dien được cảm ứng trước khi glucoza dư được tiêu thụ, và ở chỗ tính sẵn có không bị hạn chế của glucoza gây chết tế bào bởi khả năng hình thành các mức độc tố của các sản phẩm trung gian của quá trình mevalonat. Trong mẻ lên men được thể hiện trên Fig.13B, việc cảm

ứng xảy ra sau khi việc phân phối glucoza bị giới hạn, điều này ngăn ngừa sự hình thành các sản phẩm trung gian của quá trình, dẫn đến mật độ tế bào và mức sản xuất amorpho-4,11-dien cao hơn.

Ví dụ 14

Ví dụ này chứng tỏ việc sản xuất amorpho-4,11-dien tăng bởi chủng chủ *Escherichia coli* được sinh trưởng trong điều kiện nguồn cacbon và nguồn nitơ bị hạn chế ở nhiệt độ dưới điểm cực thuận.

Giống nuôi cấy mầm của chủng chủ B86 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của chủng vào bình 250ml chứa 50 ml môi trường M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1. Dịch nuôi cấy được sinh trưởng qua đêm ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút, được cấy truyền vào sáng hôm sau vào cùng môi trường ở OD₆₀₀ khoảng 1, và lại được sinh trưởng ở nhiệt độ 37°C và 250 vòng/phút đến OD₆₀₀ nằm trong khoảng từ 3 đến 5.

Bốn bình phản ứng sinh học loại 2l (thiết bị điều khiển sinh học - Biocontroller ADI 1010 có Bioconsole ADI 1025, Applikon Biotechnology, Foster City, CA) được thiết lập và được thực hiện theo cùng cách như nêu trong ví dụ 12 đối với mẻ 060403-3, ngoại trừ là các mẻ bị hạn chế nitơ không chứa amoniac sulfat trong nguyên liệu nạp.

Việc nạp liệu glucoza tăng theo hàm số mũ với thời gian gấp đôi là 6 giờ được bắt đầu một cách tự động khi viên glucoza ban đầu (15 g) được sử dụng hết và oxy hoà tan hết. Đến giá trị tối đa 30,4 g/giờ, tốc độ nạp được tính theo phương trình sau:

$$m_s(t) = S_0 \mu e^{\mu(t-t_0)}$$

$$\mu = 0,12 \text{ min}^{-1}$$

$$S_0 = 15 \text{ g}$$

trong đó μ là tốc độ sinh trưởng riêng, và t_0 là thời gian ở đó viên glucoza ban đầu bị làm suy kiệt. Trong quá trình tiến tới tốc độ tối đa, việc nạp glucoza được làm giảm đến tốc độ 11,4 g/giờ, và giữ không đổi ở tốc độ này đối với phần còn lại của mẻ. Trong mẻ lên men 060710-4, 060724-5, và 060619-5 (cacbon và nitơ bị giới hạn), việc nạp glucoza bị làm giảm tiếp khi việc hạn chế amoniac dẫn đến sự tích lũy glucoza trong môi trường.

Quá trình lên men được thực hiện ở nhiệt độ giảm khoảng 30°C. Lưu lượng khí trong bình phản ứng sinh học được đặt ở 1 vvm; khuấy ban đầu ở tốc độ 700 vòng/phút; bọt được kiểm soát bằng chất chống tạo bọt B (Sigma-Aldrich, St. Louis, MO); và áp lực oxy hoà tan được kiểm soát ở nồng độ 40% sử dụng tầng khuấy (700-1200 vòng/phút) và làm giàu oxy. Trong mẻ lên men 060327-3 (bị hạn chế cacbon), pH được duy trì ở 7 bằng cách sử dụng NH_4OH 20%; trong mẻ lên men 060710-4, 060724-5, và 060619-5 (cacbon và nitơ bị hạn chế), pH được duy trì ở 7 bằng cách ban đầu sử dụng NH_4OH 20%, và bắt đầu ở 72 giờ sử dụng hỗn hợp 50/50 chứa NaOH 2,5 N và NH_4OH 10 N, để hạn chế tiếp lượng amoniac đưa vào bình lên men.

Việc sản xuất amorph-4,11-dien ở tế bào chủ được kích thích ở giá trị OD_{600} khoảng 30 bằng cách bổ sung 1ml IPTG 1M vào môi trường nuôi cấy.

Amorph-4,11-dien được thu bằng cách phủ môi trường bằng 10% (theo thể tích) chất phủ hữu cơ.

Amorph-4,11-dien sau đó được chiết bằng cách kết hợp 25 μl cả hai mẻ với 975 μl metanol, lắc mẫu ở tốc độ lớn nhất trên thiết bị trộn Fisher Vortex Genie 2™ (Scientific Industries, Inc., Bohemia, N. Y.) trong ít nhất 15 phút, loại tế bào khỏi mẫu bằng cách ly tâm, và bổ sung 10 μl dung dịch metanol vào 990 μl etyl axetat chứa 10 $\mu\text{l/l}$ trans-caryophyllen.

Mẫu được phân tích bằng GC/MS như nêu trong ví dụ 10.

Các Fig.14A-E thể hiện dữ liệu từ mẻ lên men 060327-3 (hạn chế cacbon). Quá trình lên men tạo ra nồng độ amorph-4,11-dien tối đa là 16g/l (Fig.14A). Hiệu suất thể tích tối đa của chủng chủ là cao hơn 200mg/l/giờ

(Fig.14B). Hiệu suất riêng tối đa của chủng chủ là $>2\text{mg/l/giờ}/\text{OD}_{600}$ (Fig.14C). Nồng độ của amoniac trong môi trường nuôi cấy là khoảng 30mM vào lúc bắt đầu của mẻ lên men, tăng khoảng 76mM trong quá trình bổ sung dung dịch cấp trong pha sinh trưởng bùng nổ, và giữ ở mức trên 60mM đối với phần còn lại của mẻ (Fig.14D). OD_{600} tối đa đạt tới là khoảng 290 (Fig.14D), tương ứng với 116g DCW/L. Nồng độ glucoza trong môi trường nuôi cấy giảm từ 15g/l xuống dưới 1g/l trong ít hơn 20 giờ, và vẫn giữ mức thấp (Fig.14E). Mức axetat thấp trong quá trình lên men (Fig.14E).

Các Fig.15A-E thể hiện dữ liệu từ mẻ lên men 060710-4, 060724-5, và 060619-5 (cacbon và nitơ bị hạn chế). Quá trình lên men tạo ra nồng độ amorpho-4,11-dien tối đa nằm trong khoảng từ khoảng 20g/l đến 30g/l (Fig.15A). Hiệu suất thể tích tối đa của chủng chủ là cao hơn 400mg/l/giờ ở cả ba mẻ lên men (Fig.15B), cao hơn một cách đáng kể so với hiệu suất thể tích tối đa thu được trong quá trình lên men không hạn chế nitơ (Fig.14B). Hiệu suất riêng tối đa của chủng chủ là $>2\text{mg/l/giờ}/\text{OD}_{600}$ đối với tất cả các mẻ, và duy trì ở mức cao trong suốt các mẻ (Fig.15C). Nồng độ của amoniac trong môi trường nuôi cấy nằm trong khoảng từ 35mM đến 50mM vào lúc bắt đầu của mẻ lên men, giảm trong quá trình bổ sung dung dịch nạp trong pha sinh trưởng bùng nổ, và giữ ở mức dưới 10mM đối với phần còn lại của mẻ (Fig.15D). (Mức amoniac bị giảm so với mẻ lên men 060327-3 (Fig.14D) là do thiếu amoniac trong dung dịch nạp và amoniac giảm in bazơ được sử dụng để duy trì độ pH. Mẻ lên men 060710-4 và 060619-5 thể hiện sự hết nồng độ amoniac vào cuối mẻ, tuy nhiên sự hết này xuất hiện giai đoạn chính của việc sản xuất amorpho-4, 11-dien). Giá trị OD_{600} tối đa đạt được nằm trong khoảng từ 170 đến 220 (Fig.15D), tương ứng với từ 68g đến 88g DCW/L. Nồng độ glucoza trong môi trường nuôi cấy giảm từ 15g/l xuống dưới 1g/l trong ít hơn 20 giờ, và vẫn giữ ở mức thấp (Fig.15E). Mức axetat thấp trong toàn bộ các quá trình lên men (Fig.15E).

Ví dụ 15

Ví dụ này mô tả việc sản xuất amorpha-4,11-dien qua quá trình DXP trong chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ B003, B617, B618, và B619 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình 125ml riêng biệt chứa 25ml M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy qua đêm.

Giống nuôi cấy mầm được sử dụng để cấy truyền trong bình 250ml riêng biệt có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml môi trường M9-MOPS, thiamin 45µg/ml, chất vi dinh dưỡng, FeSO₄ 1,00E-5 mol/l, MOPS 0,1 M, dịch chiết nấm men 0,5%, D-glucoza 20g/l, và thuốc kháng sinh. Dịch nuôi cấy được ủ ở nhiệt độ 30°C trong thiết bị lắc ủ được làm ấm ở tốc độ 250 vòng/phút cho đến khi chúng đạt tới giá trị OD₆₀₀ nằm trong khoảng từ 0,2 đến 0,3, ở thời điểm này việc sản xuất amorpha-4,11-dien ở tế bào chủ được kích thích bằng cách bổ sung 40µl IPTG 1M vào môi trường nuôi cấy.

Vào thời điểm cảm ứng, dịch nuôi cấy được phủ bằng 8ml chất phủ hữu cơ để thu amorpha-4,11-dien. Các mẫu được lấy tại các thời điểm khác nhau, và amorpha-4,11-dien được chiết và được phân tích bằng GC/MS như nêu trong ví dụ 10. Các thử nghiệm được thực hiện bằng cách sử dụng 2 dòng vô tính độc lập của mỗi chủng chủ, và các kết quả được tính giá trị trung bình. Độ lệch giữa các mẫu được thấy là nhỏ hơn 10%.

Như được thể hiện trên Fig.16, chủng chủ *Escherichia coli* B619, chủng này chứa trình tự nucleotit mã hóa enzym của quá trình DXP được thao tác di truyền đầy đủ, tạo ra khoảng 45mg/g DCW amorpha-4,11-dien.

Ví dụ 16

Ví dụ này mô tả việc sản xuất 3-metyl-but-3-en-1-ol và 3-metyl-but-2-en-1-ol trong các chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ B286, B287, B288, và B291 được tạo ra bằng cách cấy thành vệt lượng nhỏ giống nuôi cấy gốc của mỗi chủng này trên thạch LB chứa thuốc kháng sinh như được nêu chi tiết trong

Bảng 1. Ba khuẩn lạc độc lập được chọn cho mỗi chủng, và mỗi khuẩn lạc được cấy truyền vào 7 ml môi trường LB chứa thuốc kháng sinh. Dịch nuôi cấy được sinh trưởng qua đêm ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút đến tận pha bùng nổ muộn. Sau đó, dịch nuôi cấy này được cấy truyền ở giá trị OD₆₀₀ là khoảng 0,05, vào bình 250ml chứa 40ml M9-MOPS, glucoza 2%, dịch chiết nấm men 0,5%, và thuốc kháng sinh. Dịch nuôi cấy này được sinh trưởng qua đêm ở nhiệt độ 37°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi chúng đạt tới giá trị OD₆₀₀ là khoảng 0,2, ở thời điểm này chúng được cảm ứng bằng cách bổ sung 40μl IPTG 1M. Dịch nuôi cấy được sinh trưởng trong 72 giờ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút. Giá trị OD₆₀₀ của mỗi dịch nuôi cấy được đo một đến hai lần một ngày, và 700μl mẫu được lấy ra. Để chiết 3-metyl-but-3-en-1-ol và 3-metyl-but-2-en-1-ol từ nước dùng nuôi cấy, 600μl etyl axetat được bổ sung vào 300 μl của mỗi mẫu đã được lấy ra. Sau đó các mẫu được trộn trong 15 phút, và 400μl của pha etyl axetat phía trên được chuyển sang lọ thủy tinh sạch để phân tích.

Các mẫu được phân tích trên thiết bị sắc ký khí/máy đo phổ khối Hewlett-Packard 6890 (GC/MS). 1μl mẫu được phân tách trên GC bằng cách sử dụng cột DB-5 (Agilent Technologies, Inc., Palo Alto, CA) và khí mang heli. Chương trình nhiệt độ cho việc phân tích là như sau: nhiệt độ 60°C trong 3 phút, tăng nhiệt độ với tốc độ 60°C/phút đến nhiệt độ là 300°C, và giữ ở nhiệt độ 300°C trong 2 phút. Tổng thời gian tiến hành là 9 phút. Các mẫu đã được phân tách được phân tích bằng máy dò chọn lọc khối 5973 mẫu Hewlett-Packard. Phổ khối trước chứng tỏ rằng 3-metyl-3-buten-1-ol và 3-metyl-2-buten-1-ol có thời gian lưu là 2,067 phút bằng cách sử dụng quy trình GC này. Để dò tiêu điểm trên 3-metyl-but-3-en-1-ol và 3-metyl-but-2-en-1-ol, sử dụng phương pháp kiểm soát ion chọn lọc, phương pháp này chỉ kiểm soát các ion 56 và 68 trong 3-metyl-but-3-en-1-ol và 3-metyl-but-2-en-1-ol.

Ví dụ 17

Ví dụ này mô tả việc sản xuất amorpho-4,11-dien bởi chủng chủ *Saccharomyces cerevisiae*.

Việc tạo ra chủng chủ EPY224 được mô tả trong ấn phẩm Ro *et al.* (*Nature* 440 940-943, 2006) và trong WO2007/005604. Chủng chủ EPY224 được xử lý plasmit biểu hiện pRS425ADS bằng cách cho sinh trưởng trong môi trường YPD (xem ấn phẩm: *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*, 2005 ed., ISBN 0-87969-728-8), dàn mỏng các khuẩn lạc đơn trên thạch YPD, và sau đó tạo đốm các khuẩn lạc đơn lên trên thạch CSM-Met His và thạch CSM-Met Leu. Các dòng vô tính sinh trưởng trên thạch CSM-Met His nhưng không sinh trưởng trên thạch CSM-Met Leu được xử lý (tức là, mất plasmit pRS425ADS). Một dòng vô tính như vậy được chỉ định là EPY300. EPY300 được biến nạp với plasmit biểu hiện pRS425-ADS-LEU2d, là plasmit giống với pRS425-ADS ngoại trừ là thay vì chứa LEU2, nó lại chứa gen chỉ thị chọn lọc LEU2d (xem ấn phẩm: Erhart and Hollenberg (1983) *J Bacteriol.* 156 625-635) để thu được chủng chủ Y185.

Thế biến nạp tế bào chủ Y185 được chọn trên môi trường tổng hợp xác định, chứa glucoza 2% và tất cả các axit amin trừ histidin, leuxin, và methionin (CSM-glucoza, MP Biomedicals, Solon, OH). Chủng chủ EPY300 có tính dinh dưỡng thụ động đối với sự sinh tổng hợp leuxin (*leu2*), nhưng plasmit biểu hiện pRS425-ADS-LEU2d trong Y185 lại khôi phục lại nguyên dưỡng leuxin (LEU2). Các khuẩn lạc đơn được tạo đốm trên môi trường chọn lọc (CSM-ghicoza-histidin, leuxin, methionin), và được sinh trưởng trong 2 ngày. Các tế bào được cạo khỏi đĩa và được chuyển vào 1ml glyxerol 25% (theo thể tích) trong ống nghiệm nhiệt độ thấp. Dịch huyền phù được trộn, và sau đó được bảo quản ở nhiệt độ -80°C.

Bình giống mầm của chủng chủ Y185 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của chủng vào bình 125ml chứa 25ml CSM-glucoza khuyết leuxin và methionin, và bằng cách nuôi dịch nuôi cấy qua đêm. Dịch nuôi cấy được sử dụng để cấy truyền trong bình có vách ngăn dung tích 250ml có giá trị OD₆₀₀ khoảng 0,05 chứa 40ml môi trường tổng hợp xác định

thiếu leuxin, và chứa glucoza 0,2%, galactoza 1,8%, và methionin 1mM. Dịch nuôi cấy này được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn ở 200 vòng/phút. Vì việc có mặt glucoza trong môi trường ngăn chặn sự cảm ứng của yếu tố khởi đầu *GAL1* bởi galactoza, việc sản xuất amorpho-4,11-dien không được cảm ứng cho đến khi tế bào đã sử dụng hết glucoza trong môi trường và chuyển sang sử dụng galactoza để làm nguồn cacbon chính của chúng. Tại thời điểm cấy truyền, dịch nuôi cấy được phủ bằng 8 ml chất phủ hữu cơ để thu amorpho-4,11-dien. Các mẫu được lấy tại thời điểm 72 giờ bằng cách chuyển 5 µl của lớp dung môi hữu cơ vào lọ thủy tinh sạch chứa 500µl etyl axetat chứa beta- hoặc trans-caryophylen với nồng độ đã biết để làm tiêu chuẩn bên trong.

Các mẫu chất phủ hữu cơ/etyl axetat được phân tích trên thiết bị sắc ký khí/phổ khối Hewlett-Packard 6890 (GC/MS) như nêu trong ví dụ 10.

Sau 72 giờ sinh trưởng, 3 dịch nuôi cấy nấm men được phát hiện thấy là tạo ra 60,68, 54,48, và 59,25 mg/l amorpho-4,11-dien.

Ví dụ 18

Ví dụ này mô tả việc sản xuất amorpho-4,11-dien trong chủng chủ *Saccharomyces cerevisiae* trong đó chủng chủ chứa quá trình mevalonat nguyên thể cũng như quá trình mevalonat khác loại được kiểm soát bởi chất kiểm soát điều hòa khác loại.

Các chủng nấm men CEN.PK2-1C (Y002) (MATA; ura3-52; trp1-289; leu2-3,112; his3ΔI, MAL2-8C, SUC2) và CEN.PK2-1D (Y003) (MATalpha; ura3-52; trp1-289; leu2-3,112; his3ΔI; MAL2-8C; SUC2) (xem ấn phẩm: J. P. van dijken *et al.*, *Enzyme Microb Technol* 26, 706 (Jun 1, 2000) được nuôi cấy trong môi trường chuẩn giàu dinh dưỡng (YPD) hoặc trong môi trường tổng hợp xác định (xem ấn phẩm: . D. Rose, F. Winston, P. Heiter, *Methods in yeast genetics, a laboratory course manual*. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y., 1990) khuyết các chất dinh dưỡng thích hợp cho phép chọn các thể biến nạp nguyên, việc giữ lại plasmit, và thể hệ con giảm phân.

Sự biến nạp qua trung gian ADN vào *S. cerevisiae* được thực hiện bằng cách sử dụng quy trình lithi axetat như được mô tả bởi R. H. Schiestl, R. D. Gietz, *Curr Genet* 16, 339 (Tháng 12, 1989). Tất cả các thay thế và phá vỡ gen được xác nhận bằng phân tích kiểu hình, phản ứng chuỗi polymeaza ("PCR") và đọc trình tự ADN của hệ gen đã được khuếch đại. Plasmid pAM489-pAM498 được cấu trúc bằng cách sử dụng pCR 2.1 (Invitrogen, Carlsbad CA) và được thể hiện bằng sơ đồ trên Fig.7A-C và Bảng 6. Các trình tự gen chỉ thị HISMX được mô tả trong ấn phẩm M. S. Longtine *et al.*, *Yeast* 14, 953 (Tháng 7, 1998). Sự nhân lên của ADN plasmid được thực hiện trong chủng *Escherichia coli* DH5 α .

Bảng 6

Chủng	5'HR	Gen #1	Yếu tố khởi đầu Crick	Yếu tố khởi đầu Watson	Gen #2	Chất đánh dấu di truyền	3'HR
pAM489	TRP1	tHMGR	GAL1	GAL10	ERG20	TRP1	TRP1
pAM490	TRP1	tHMGR	CUP1	CUP1	ERG20	TRP1	TRP1
pAN491	URA3	tHMGR	GAL1	GAL10	ERG13	URA3	URA3
pAM492	URA3	IDI1	CUP1	CUP1	tHMGR	URA3	URA3
pAM493	ADE1	tHMGR	GAL1	GAL10	IDI1	ADE1	URA3
pAM494	ADE1	tHMGR	CUP1	CUP1	IDI1	ADE1	ADE1
pAM495	HIS3	ERG12	GAL1	GAL10	ERG10	HISMX	HIS3
pAM496	HIS3	ERG12	CUP1	CUP1	ERG10	HISMX	HIS3
pAM497	LEU2	ERG19	GAL1	GAL1	ERG8	HISMX	LEU2
pAM498	LEU2	ERG19	CUP1	CUP1	ERG8	HISMX	LEU2

Các chủng *S. cerevisiae* Y002 và Y003 được chuẩn bị cho việc đưa các gen của quá trình mevalonat có khả năng cảm ứng bằng phương pháp sau. Yếu tố khởi đầu ERG9 được thay thế bằng yếu tố khởi đầu *MET3* ở *S. cerevisiae* bằng cách khuếch đại bằng PCR của vùng KanMX-PMET3 từ pAM328 (SEQ

ID NO: 43) sử dụng các môi 50-56-pw100-G (SEQ ID NO:44) và 50-56-pw101-G (SEQ ID NO: 45) chứa 45 cặp bazơ tương đồng với yếu tố khởi đầu ERG9 nguyên. 10µg sản phẩm PCR tạo ra được biến nạp vào các chủng Y002 và Y003 đang sinh trưởng bùng nổ bằng cách sử dụng polyetylen glycol 3350 40% theo khối lượng (Sigma-Aldrich St Louis, MO), lithi axetat 100mM (Sigma), 10µg ADN sẹ cá hồi (Invitrogen) và ủ ở nhiệt độ 30°C trong 30 phút sau đó gây sốc nhiệt ở nhiệt độ 42°C trong 30 phút (như được mô tả bởi Schiestl & Gietz, *Curr. Genet.* 16: 339 (1989)). Thể tái tổ hợp dương tính được nhận biết nhờ khả năng sinh trưởng của chúng trên môi trường giàu dinh dưỡng chứa 0,5µg/ml genetixin (Invitrogen Co, Carlsbad, CA) và được xác nhận bằng PCR chẩn đoán. Các dòng vô tính tạo ra được chỉ định là Y93 (MAT A) và Y94 (MAT alpha). Tiếp theo, khung đọc mở *ADE1* được thay thế bằng gen *LEU2* ở *Candida glabrata* (*CgLEU2*). Locus hệ gen *CgLEU2* 3,5kb được khuếch đại từ ADN của hệ gen *C. glabrata* (ATCC, Manassas, VA) sử dụng các môi 61-67-CPK066-G (SEQ ID NO: 46) và 61-67-CPK067-G (SEQ ID NO: 47) chứa 50 cặp bazơ có độ tương đồng bên với khung đọc mở (ORF) *ADE1*. 10µg sản phẩm PCR tạo ra được biến nạp vào Y93 và Y94 đang sinh trưởng bùng nổ như được mô tả ở trên. Các chủng *ade1* được chọn để sinh trưởng trong điều kiện không bổ sung leuxin và được xác nhận bằng PCR chẩn đoán. Các dòng vô tính tạo ra được chỉ định là Y176 (MAT A) và Y177 (MAT alpha).

Để tạo ra chủng *S. cerevisiae* Y188, 2µg ADN plasmit từ pAM491 (SEQ ID NO: 48) và pAM495 (SEQ ID NO:49), một cách lần lượt, được phân giải bằng PmeI qua đêm (New England Biolabs, Beverly, MA) và được đưa vào Y176 đang sinh trưởng bùng nổ như được mô tả ở trên. Thể tái tổ hợp dương tính được chọn cho sự sinh trưởng trên môi trường khuyết uraxil và histidin. Sự hợp nhất vào locus hệ gen chính xác được xác nhận bằng PCR chẩn đoán.

Để tạo ra chủng *S. cerevisiae* Y189, 2µg ADN plasmit từ pAM489 (SEQ ID NO: 50) và pAM497 (SEQ ID NO: 51), một cách lần lượt, được phân

giải bằng PmeI qua đêm và được đưa vào Y177 đang sinh trưởng bùng nổ như được mô tả ở trên. Thể tái tổ hợp dương tính được chọn cho sự sinh trưởng trên môi trường khuyết tryptophan và histidin. Sự hợp nhất vào locus hệ gen chính xác được xác nhận bằng PCR chẩn đoán.

Khoảng 1×10^7 tế bào từ Y188 và Y189 được trộn trên đĩa chứa môi trường YPD trong 6 giờ ở nhiệt độ phòng để lai. Sau đó, tế bào nuôi cấy đã được trộn được dàn mỏng vào môi trường khuyết histidin, uraxil và tryptophan để chọn sự sinh trưởng của các tế bào lưỡng bội. 2 μ g ADN plasmit từ pAM493 (SEQ ID NO: 52) được phân giải qua đêm bằng PmeI và được đưa vào các tế bào lưỡng bội đang sinh trưởng bùng nổ như được mô tả ở trên. Thể tái tổ hợp dương tính được chọn sinh trưởng trên môi trường thiếu adenin. Sự hợp nhất vào locus hệ gen chính xác được xác nhận bằng PCR chẩn đoán. Chúng tạo ra được chỉ định là Y238.

Để tạo ra các chủng đơn bội chứa sợi bổ sung đầy đủ của gen đã được đưa vào, Y238 được tạo bào tử trong môi trường lỏng chứa kali axetat 2% và rafinoza 0,02%. Bộ bốn khoảng 200 gen (bộ bốn là sản phẩm giảm phân bốn bào tử) được phân lập bằng cách sử dụng thiết bị vi thao tác loại Singer Instruments MSM300 (Singer Instrument Co, LTD. Somerset, UK). Các thể phân lập gen độc lập chứa sợi bổ sung thích hợp của nguyên liệu di truyền đã được đưa vào được nhận biết nhờ khả năng sinh trưởng của chúng trong điều kiện khuyết adenin, histidin, uraxil, và tryptophan. Sự hợp nhất của tất cả các ADN đã được đưa vào được xác nhận bằng PCR chẩn đoán. Các chủng tạo ra được chỉ định là Y210 (MAT A) và Y211 (MAT alpha).

2 μ g ADN plasmit từ pAM426 (SEQ ID NO:53), chứa Amorphadein Synthaza (ADS) được tối ưu hoá codon của *S. cerevisiae* được biểu hiện từ yếu tố khởi đầu *GALI* *S. cerevisiae*, được đưa vào Y210 và Y211 đang sinh trưởng bùng nổ như được mô tả ở trên. Các chủng *S. cerevisiae* chứa plasmit pAM426 được chọn về khả năng sinh trưởng của chúng trong điều kiện không bổ sung leuxin. Các chủng tạo ra được chỉ định là Y225 (MAT A) và Y227 (MAT alpha).

2 μ g ADN plasmit từ pAM322 (SEQ ID NO: 54), chứa Amorphadin Synthaza (ADS) được tối ưu hoá codon của *S. cerevisiae* và cytochrom P450 monooxygenaza (AMO) được biểu hiện từ *GAL1* ở *S. cerevisiae* và cytochrom P450 oxidoreductaza (CPR) được biểu hiện từ yếu tố khởi đầu GAL10 của *S. cerevisiae*, được đưa vào Y210 và Y211 đang sinh trưởng bùng nổ như được mô tả ở trên. Các chủng *S. cerevisiae* chứa plasmit pAM322 được chọn về khả năng sinh trưởng của chúng trong điều kiện không bổ sung leuxin. Các chủng tạo ra được chỉ định là Y222 (MAT A) và Y224 (MAT alpha).

Ví dụ 19

Ví dụ này mô tả việc sản xuất α -farnesen hoặc β -farnesen trong các chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ B552 và B592 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào bình 125 ml chứa 25 ml M9-MOPS, glucoza 0,8%, dịch chiết nấm men 0,5%, và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy qua đêm.

Các giống nuôi cấy mầm được sử dụng để cấy truyền vào bình 250ml có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml M9-MOPS, glucoza 2%, dịch chiết nấm men 0,5%, và thuốc kháng sinh. Dịch nuôi cấy được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút cho đến khi chúng đạt tới OD₆₀₀ khoảng 0,2, ở thời điểm này việc sản xuất α -farnesen hoặc β -farnesen ở tế bào chủ được cảm ứng bằng cách bổ sung 40 μ l IPTG 1 M. Vào thời điểm cảm ứng, dịch nuôi cấy được phủ bằng 8 ml chất phủ hữu cơ để thu α -farnesen. Các mẫu được lấy 24 giờ một lần đến 120 giờ (tổng là 5 thời điểm) bằng cách chuyển từ 2 μ l đến 10 μ l lớp chất phủ hữu cơ vào lọ thuỷ tinh sạch chứa 1ml etyl axetat có gắn trans-caryophylen để làm tiêu chuẩn bên trong. Ngoài ra, 1ml lượng nhỏ dịch nuôi cấy được quay tròn, các hạt nhỏ tế bào được tái tạo huyền phù trong 250 μ l nước vô trùng, và huyền phù tế bào được chuyển vào lọ thuỷ tinh chứa 1ml etyl axetat có gắn trans-caryophylen để làm tiêu chuẩn bên trong. Ngoài ra, 0,5ml lượng nhỏ của toàn bộ canh nuôi cấy được bổ sung vào lọ thuỷ tinh chứa 1ml etyl axetat có gắn trans-caryophylen để làm tiêu

chuẩn bên trong. Toàn bộ các mẫu canh nuôi cấy mẫu đều được chiết trong etyl axetat bằng cách xoáy lọ thủy tinh trong 10 phút, sau đó 600 μ l dịch chiết etyl axetat được chuyển sang lọ thủy tinh sạch.

Các mẫu chất phủ hữu cơ/etyl axetat và toàn bộ các mẫu canh nuôi cấy được chiết bằng etyl axetat được phân tích trên sắc ký Agilent 6890N được trang bị máy đo phổ khối Agilent 5975 (GC/MS) với phương pháp quét đầy đủ (50-500 m/z). Để xúc tiến thời gian tiến hành, chương trình nhiệt độ và ma trận cột được biến đổi để đạt được độ phân giải đỉnh tối ưu và tổng thời gian tiến hành ngắn nhất. 1 μ l mẫu được phân tách bằng cách sử dụng cột HP-5MS (Agilent Technologies, Inc., Palo Alto, CA) và khí chất mang heli. Chương trình nhiệt độ cho phép phân tích là như sau: nhiệt độ 150°C được duy trì trong khoảng 3 phút, tăng nhiệt độ với tốc độ 25°C/phút đến nhiệt độ 200°C, tăng nhiệt độ với tốc độ 60°C/phút đến nhiệt độ 300°C, và giữ ở nhiệt độ 300°C trong 1 phút. Phổ khối trước chứng tỏ rằng sản phẩm β -farnesen synthaza là β -farnesen, và chứng tỏ rằng β -farnesen có thời gian lưu là 4,33 phút bằng cách sử dụng quy trình GC này. Độ chuẩn farnesen được tính toán bằng cách so sánh diện tích đỉnh được tạo ra so với đường cong định chuẩn định lượng của β -farnesen đã được tinh chế (Sigma-Aldrich Chemical Company, St. Louis, MO) trong etyl axetat có gắn trans-caryophyllen.

Chủng chủ B592 sản xuất ra khoảng 400mg/l α -farnesen tại thời điểm 120 giờ (được tính trung bình qua 3 dòng vô tính độc lập), và có năng suất riêng tối đa khoảng 46mg/l/OD₆₀₀). Chủng chủ B552 sản xuất khoảng 1,1g/l β -farnesen tại thời điểm 120 giờ (được tính trung bình qua 3 dòng vô tính độc lập), và có năng suất riêng tối đa khoảng 96mg/l/OD₆₀₀ (1 dòng vô tính đại diện).

Ví dụ 20

Ví dụ này mô tả việc sản xuất β -farnesen qua quá trình DXP ở chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ B650, B651, B652, và B653 được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình dung tích 125ml riêng biệt chứa 25ml M9-MOPS và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy này qua đêm.

Giống nuôi cấy mầm được sử dụng để cấy truyền trong bình 250ml riêng biệt có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml môi trường M9-MOPS tối thiểu, thiamin 45µg/ml, chất vi dinh dưỡng, FeSO₄ 1,00E-5mol/l, MOPS 0,1M, dịch chiết nấm men 0,5%, D-glucosa 20g/l, và thuốc kháng sinh. Dịch nuôi cấy này được ủ ở nhiệt độ 30°C trong máy lắc ủ được làm ẩm ở tốc độ 250 vòng/phút cho đến khi chúng đạt tới giá trị OD₆₀₀ nằm trong khoảng từ 0,2 đến 0,3, tại thời điểm này việc sản xuất β-farnesen ở tế bào chủ được cảm ứng bằng cách bổ sung 40µl IPTG 1M vào môi trường nuôi cấy. Vào thời điểm cảm ứng, dịch nuôi cấy được phủ bằng 8 ml chất phủ hữu cơ để thu β-farnesen. Các mẫu được lấy tại các thời điểm khác nhau bằng cách chuyển 100µl mẫu chứa lớp chất phủ hữu cơ phía trên vào ống nghiệm sạch. Ống nghiệm này được ly tâm để tách riêng tế bào hoặc môi trường bất kỳ còn lại, và 10µl mẫu chất phủ hữu cơ được chuyển vào 500µl etyl axetat có gắn beta- hoặc trans-caryophylen để làm tiêu chuẩn bên trong trong các lọ thủy tinh GC sạch. Hỗn hợp được trộn trong 30 giây, và sau đó được phân tích như nêu trong ví dụ 18. Chủng chủ *Escherichia coli* B653 sản xuất khoảng 7mg/g DCW β-farnesen.

Ví dụ 21

Ví dụ này mô tả việc sản xuất α-farnesen hoặc β-farnesen ở chủng chủ *Saccharomyces cerevisiae*.

Chủng EPY300 được tạo ra bằng cách loại bỏ plasmit biểu hiện từ chủng *Saccharomyces cerevisiae* EPY224 (xem ấn phẩm: Ro *et al.* (2006) *Nature* 440 940-943, WO2007/005604) bằng cách nuôi cấy trong môi trường giàu dinh dưỡng. Sau đó, chủng EPY300 được biến nạp với plasmit biểu hiện pRS425-FSA hoặc pR425-FSB, để thu được các chủng chủ Y166 và Y164 tương ứng.

Thế biến nạp tế bào chủ được chọn trên môi trường tổng hợp xác định, chứa glucoza 2% và tất cả các axit amin trừ leuxin (SM-glu). Chủng chủ EPY300 có tính dinh dưỡng thụ động đối với sinh tổng hợp leuxin (Leu2), nhưng plasmit biểu hiện pRS425-FSA hoặc pRS425-FSB lại khôi phục lại nguyên dưỡng leuxin (LEU2). Các khuẩn lạc đơn được chuyển vào lọ dịch nuôi cấy chứa 5ml SM-glu lỏng khuyết leuxin. Dịch nuôi cấy này được ủ bằng cách lắc ở nhiệt độ 30°C cho đến khi sự sinh trưởng đạt đến pha ổn định. Các tế bào được bảo quản ở nhiệt độ -80°C trong các lọ bảo quản lạnh sâu trong lượng nhỏ đóng băng 1 ml được làm từ 400µl glyxerol 50% và 600µl dịch nuôi cấy lỏng.

Giống nuôi cấy mầm được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc vào bình 125ml chứa 25ml SM-glu khuyết leuxin, và nuôi dịch nuôi cấy này qua đêm. Chủng giống nuôi cấy được sử dụng để cấy truyền vào trong bình có vách ngăn dung tích 250ml có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml môi trường tổng hợp xác định thiếu leuxin, glucoza 0,2%, và galactoza 1,8%. Dịch nuôi cấy được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn ở tốc độ 200 vòng/phút. Vì sự có mặt của glucoza trong môi trường ngăn cản sự cảm ứng của yếu tố khởi đầu *Gall* bởi galactoza, nên việc sản xuất farnesen không được cảm ứng cho đến khi tế bào sử dụng hết glucoza trong môi trường và chuyển sang sử dụng galactoza làm nguồn cacbon chủ yếu của chúng. Dịch nuôi cấy được được phủ bằng 8ml metyl oleat hoặc isopropyl myristat. Các mẫu được lấy 24 giờ một lần bằng cách chuyển 2-10µl lớp dung môi hữu cơ vào lọ thủy tinh sạch chứa 500µl etyl axetat chứa beta- hoặc trans-caryophylen đã biết nồng độ để làm tiêu chuẩn bên trong. Ngoài ra, lượng nhỏ 0,5ml của toàn bộ canh nuôi cấy được bổ sung vào lọ thủy tinh chứa 1ml etyl axetat có gắn trans-caryophylen để làm tiêu chuẩn bên trong. Toàn bộ các mẫu canh nuôi cấy được chiết trong etyl axetat bằng cách trộn lọ thủy tinh trong 10 phút, sau đó 600µl của dịch chiết etyl axetat được chuyển sang lọ thủy tinh sạch.

Chủng chủ Y166 sản xuất khoảng 9,8mg/l α -farnesen tại thời điểm 120 giờ (được tính trung bình qua 3 dòng vô tính độc lập), và có năng suất riêng tối

đa khoảng 3 mg/l/OD₆₀₀ (1 dòng vô tính đại diện). Chủng chủ Y164 sản xuất ra khoảng 56mg/l β -farnesen tại thời điểm 120 giờ (được tính trung bình qua 3 dòng vô tính độc lập), và có năng suất riêng tối đa khoảng 20 mg/l/OD₆₀₀ (1 dòng vô tính đại diện).

Ví dụ 22

Ví dụ này mô tả việc sản xuất γ -terpinen, α -pinen, và terpinolen trong các chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm của các chủng chủ để sản xuất γ -terpinen (*E. coli* DH1-T1r [pMevT, pMevB-Gpps, pAM445]), α -pinen (*E. coli* DH1-T1r [pMevT, pMevB-Gpps, pAM443 hoặc pAM442]) hoặc terpinolen (*E. coli* DH1-TI1 [pMevT, pMevB-Gpps, pAM444] được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình riêng biệt dung tích 125 ml chứa 25 ml M9-MOPS, glucoza 2%, dịch chiết nấm men 0,5%, và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy này qua đêm đến pha sinh trưởng theo hàm mũ.

Giống nuôi cấy mầm được sử dụng để cấy truyền vào trong bình 250ml có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml M9-MOPS, glucoza 2%, dịch chiết nấm men 0,5%, và thuốc kháng sinh. Tại thời điểm cấy truyền, dịch nuôi cấy còn được phủ bằng 4ml hexadecan. Dịch nuôi cấy này được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn ở tốc độ 200 - 250 vòng/phút cho đến khi chúng đạt tới giá trị OD₆₀₀ khoảng 0,2, tại thời điểm này việc sản xuất hợp chất quan tâm ở tế bào chủ được cảm ứng bằng cách bổ sung 40 μ l IPTG 1M. Các mẫu được lấy một lần mỗi ngày trong 96 giờ bằng cách chuyển 200 μ l lớp hexadecan vào 0,6ml ống vi ly tâm. Để phân tích, lớp phủ hexadecan được pha loãng theo tỷ lệ 1:1 hoặc 1:10 với etyl axetat đã được gắn trans-caryophylen để làm tiêu chuẩn bên trong trong lọ GC 1,8 ml. Ngoài ra, lượng nhỏ 1 ml của dịch nuôi cấy được xoay tròn, các hạt nhỏ tế bào được tái tạo huyền phù trong 250 μ l nước vô trùng, và hỗn dịch tế bào được chuyển vào lọ thủy tinh chứa 1ml etyl axetat có bông trans-caryophylen để làm tiêu chuẩn bên trong. Các hạt

nhỏ tế bào được chiết trong etyl axetat bằng cách trộn các lọ thủy tinh trong 15 phút, sau đó 500µl dịch chiết etyl axetat được chuyển sang lọ thủy tinh sạch.

Các mẫu hexadecan/etyl axetat và các mẫu hạt nhỏ tế bào đã được chiết bằng etyl axetat được phân tích trên sắc ký khí Agilent 6890N được trang bị máy đo phổ khối Agilent 5975 (GC/MS) theo cách quét đầy đủ (50-500 m/z). Để xúc tiến thời gian tiến hành, chương trình nhiệt độ và ma trận cột được biến đổi để đạt được độ phân giải đỉnh tối ưu và tổng thời gian tiến hành ngắn nhất. 1 µl mẫu được tách (tỷ lệ tách nằm trong khoảng từ 1:2 đến 1:50 được chọn dựa trên nồng độ mẫu) và sau đó được phân tách bằng cách sử dụng cột HP-5MS (Agilent Technologies, Inc., Palo Alto, CA) và khí chất mang heli. Chương trình nhiệt độ cho phép phân tích là như sau: nhiệt độ 75°C được duy trì trong 3 phút, tăng nhiệt độ với tốc độ 20°C/phút đến nhiệt độ 115°C, tăng nhiệt độ với tốc độ 60°C/phút đến nhiệt độ 300°C, và giữ ở nhiệt độ 300°C trong 0,5 phút. Các sản phẩm khác nhau, γ -terpinen, α -pinen, và terpinolen lần lượt được quan sát thấy ở 5,4, 4,1, 5,4, và 5,9 phút. Độ chuẩn được tính toán bằng cách so sánh diện tích đỉnh được tạo ra với đường cong định chuẩn định lượng của các chuẩn đã được tinh chế trong etyl axetat được gắn trans-caryophyllen.

Ví dụ 23

Ví dụ này mô tả việc sản xuất linalool, limonen, β -pinen, β -phellandren, caren, hoặc sabinin trong các chủng chủ *Escherichia coli*.

Giống nuôi cấy mầm được tạo ra bằng cách bổ sung lượng nhỏ giống nuôi cấy gốc của mỗi chủng vào các bình riêng biệt dung tích 125ml chứa 25 ml M9-MOPS, dịch chiết nấm men 0,5%, glucoza 2%, và thuốc kháng sinh như được nêu chi tiết trong Bảng 1, và nuôi dịch nuôi cấy này qua đêm.

Giống nuôi cấy mầm được sử dụng để cấy truyền vào trong bình có vách ngăn dung tích 250ml có giá trị OD₆₀₀ ban đầu khoảng 0,05 chứa 40ml M9-MOPS, dịch chiết nấm men 0,5%, glucoza 2%, và thuốc kháng sinh. Dịch nuôi cấy được ủ ở nhiệt độ 30°C trên thiết bị lắc xoay tròn với tốc độ 250 vòng/phút

cho đến khi chúng đạt tới giá trị OD_{600} khoảng 0,2, vào thời điểm này việc sản xuất hợp chất quan tâm ở tế bào chủ được cảm ứng bằng cách bổ sung 40 μ l IPTG 1M vào môi trường nuôi cấy. Hợp chất quan tâm được tách khỏi môi trường nuôi cấy qua sự chiết dung môi-dung môi, hoặc bằng cách lắng và gạn nếu độ chuẩn của hợp chất quan tâm là đủ lớn để bão hoà môi trường và để tạo thành pha thứ hai.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất isoprenoit bao gồm các bước:

(a) thu nhiều tế bào chủ vi khuẩn, nấm hoặc nấm men chứa quá trình mevalonat hoặc deoxyxyluloza 5-phosphat (DXP) để tạo ra isopentenyl pyrophosphat và trong đó một hoặc nhiều enzym của quá trình này được kiểm soát bởi ít nhất một yếu tố điều hòa phiên mã khác loại; và

(b) nuôi cấy các tế bào chủ trong môi trường ở các điều kiện dưới điểm cực thuận so với các điều kiện sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ này,

trong đó các điều kiện dưới điểm cực thuận bao gồm nuôi cấy trong môi trường có nguồn cacbon, nhiệt độ, nguồn nitơ, hoặc sự kết hợp bất kỳ của chúng, là thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ, và

trong đó việc nuôi cấy tế bào chủ trong môi trường ở các điều kiện dưới điểm cực thuận này dẫn đến làm tăng sản xuất isoprenoit so với việc nuôi cấy tế bào chủ ở các điều kiện sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ này.

2. Phương pháp theo điểm 1, trong đó quá trình này là quá trình mevalonat.

3. Phương pháp theo điểm 1, trong đó quá trình này là quá trình DXP.

4. Phương pháp theo điểm 1, trong đó ít nhất một yếu tố điều hòa phiên mã khác loại này có thể cảm ứng được.

5. Phương pháp theo điểm 1, trong đó các enzym của quá trình này được kiểm soát bởi yếu tố điều hòa phiên mã đơn.

6. Phương pháp theo điểm 1, trong đó các enzym của quá trình này được kiểm soát bởi yếu tố điều hòa phiên mã đa thành phần.

7. Phương pháp theo điểm 1, trong đó tế bào chủ này là *Escherichia coli*.
8. Phương pháp theo điểm 1, trong đó tế bào chủ này là nấm.
9. Phương pháp theo điểm 1, trong đó tế bào chủ này là *Saccharomyces cerevisiae*.
10. Phương pháp theo điểm 2, trong đó quá trình này chứa trình tự axit nucleic mã hóa enzym của quá trình mevalonat từ vi khuẩn có quá trình mevalonat nội sinh.
11. Phương pháp theo điểm 10, trong đó vi khuẩn này là vi khuẩn thuộc chi được chọn từ nhóm bao gồm *Enterococcus*, *Pseudomonas*, và *Staphylococcus*.
12. Phương pháp theo điểm 11, trong đó trình tự axit nucleic mã hóa enzym của quá trình mevalonat này được chọn từ nhóm bao gồm acetyl-CoA thiolaza, HMG-CoA synthaza, HMG-CoA reductaza, và mevalonat kinaza.
13. Phương pháp theo điểm 10, trong đó trình tự axit nucleic này mã hóa HMG reductaza lớp II.
14. Phương pháp theo điểm 1, trong đó tế bào chủ được nuôi cấy trong môi trường trong đó dinh dưỡng hoặc nhiệt độ hoặc cả hai được duy trì ở mức thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa cho tế bào chủ này.
15. Phương pháp theo điểm 1, trong đó nhiệt độ của môi trường này thấp hơn ít nhất khoảng từ 2 đến 20°C so với nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa.
16. Phương pháp theo điểm 15, trong đó tế bào chủ được nuôi cấy ở nhiệt độ thấp hơn ít nhất là 5°C so với nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa.
17. Phương pháp theo điểm 15, trong đó nhiệt độ của môi trường này thấp hơn ít nhất là 10°C so với nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa.

18. Phương pháp theo điểm 1, trong đó môi trường này chứa nguồn cacbon và lượng nguồn cacbon này sẽ tạo ra khoảng 90% hoặc thấp hơn 90% tốc độ sinh trưởng riêng tối đa.

19. Phương pháp theo điểm 18, trong đó lượng nguồn cacbon này sẽ tạo ra khoảng 75% hoặc thấp hơn 75% tốc độ sinh trưởng riêng tối đa.

20. Phương pháp theo điểm 18, trong đó lượng nguồn cacbon này sẽ tạo ra khoảng 50% hoặc thấp hơn 50% tốc độ sinh trưởng riêng tối đa.

21. Phương pháp theo điểm 18, trong đó lượng nguồn cacbon này sẽ tạo ra khoảng 25% hoặc thấp hơn 25% tốc độ sinh trưởng riêng tối đa.

22. Phương pháp theo điểm 18, trong đó lượng nguồn cacbon này sẽ tạo ra khoảng từ 75% đến 10% tốc độ sinh trưởng riêng tối đa.

23. Phương pháp theo điểm 1, trong đó môi trường này chứa nguồn nitơ có mặt với lượng thấp hơn lượng sẽ tạo ra khoảng 90% hoặc thấp hơn 90% tốc độ sinh trưởng riêng tối đa.

24. Phương pháp theo điểm 23, trong đó lượng nguồn nitơ này sẽ tạo ra khoảng 75% hoặc thấp hơn 75% tốc độ sinh trưởng riêng tối đa.

25. Phương pháp theo điểm 23, trong đó lượng nguồn nitơ này sẽ tạo ra 50% hoặc thấp hơn 50% tốc độ sinh trưởng riêng tối đa.

26. Phương pháp theo điểm 23, trong đó lượng nguồn nitơ này sẽ tạo ra khoảng 25% hoặc thấp hơn 25% tốc độ sinh trưởng riêng tối đa.

27. Phương pháp theo điểm 23, trong đó lượng nguồn nitơ này sẽ tạo ra khoảng từ 75% đến 10% tốc độ sinh trưởng riêng tối đa.

28. Phương pháp theo điểm 1, trong đó isoprenoit được sản xuất ra với lượng lớn hơn 10 gam trên mỗi lít môi trường.

29. Phương pháp theo điểm 1, trong đó isoprenoit được sản xuất ra với lượng lớn hơn khoảng 50 mg trên mỗi gam khối lượng tế bào khô.
30. Phương pháp theo điểm bất kỳ trong số các điểm 28 hoặc 29, trong đó lượng isoprenoit được sản xuất ra trong khoảng thời gian nhỏ hơn 150 giờ.
31. Phương pháp theo điểm bất kỳ trong số các điểm 28 hoặc 29, trong đó lượng isoprenoit được sản xuất ra trong khoảng thời gian nhỏ hơn 96 giờ.
32. Phương pháp theo điểm bất kỳ trong số các điểm 28 hoặc 29, trong đó lượng isoprenoit được sản xuất ra trong khoảng thời gian nhỏ hơn 72 giờ.
33. Phương pháp theo điểm 1, trong đó tế bào chủ này được nuôi cấy ở nhiệt độ thấp hơn nhiệt độ sẽ tạo ra tốc độ sinh trưởng riêng tối đa, và trong đó tế bào chủ được nuôi cấy ở môi trường trong đó nguồn cacbon được duy trì ở mức thấp hơn mức sẽ tạo ra tốc độ sinh trưởng riêng tối đa này.
34. Phương pháp theo điểm 1, trong đó isoprenoit này được chọn từ nhóm bao gồm hemiterpen, monoterpen, diterpen, triterpen, tetraterpen, và polyterpen.
35. Phương pháp theo điểm 1, trong đó isoprenoit này không phải là carotenoit.
36. Phương pháp theo điểm 1, trong đó isoprenoit này là C₅-C₂₀ isoprenoit.
37. Phương pháp theo điểm 1, trong đó isoprenoit này được lựa chọn từ nhóm bao gồm abietadien, amorphadien, caren, α -farnesen, β -farnesen, farnesol, geraniol, geranylgeraniol, isopren, linalool, limonen, myrcen, nerolidol, ocimen, patchoulol, β -pinen, sabinen, γ -terpinen, terpinolen và valencen.

Fig.1A

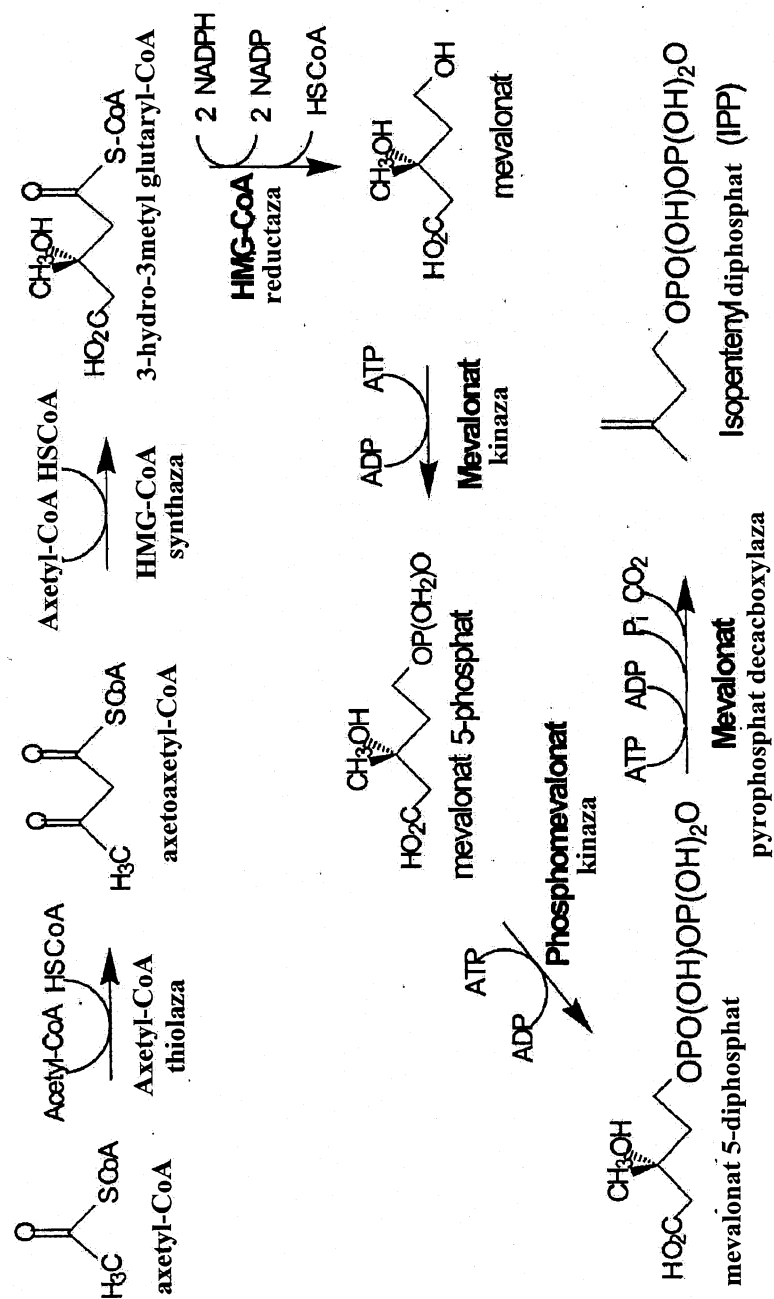


Fig.1B

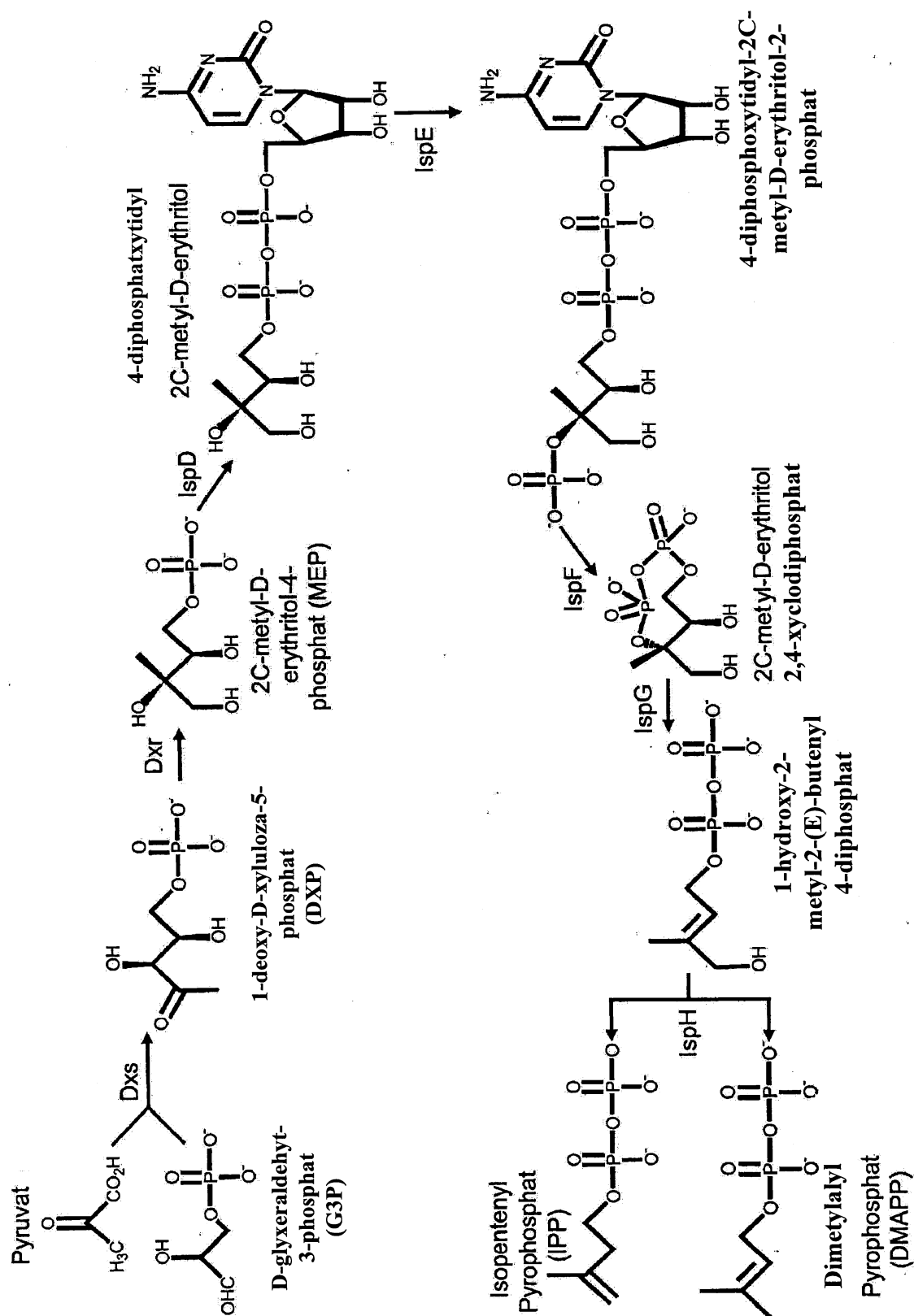


Fig. 2

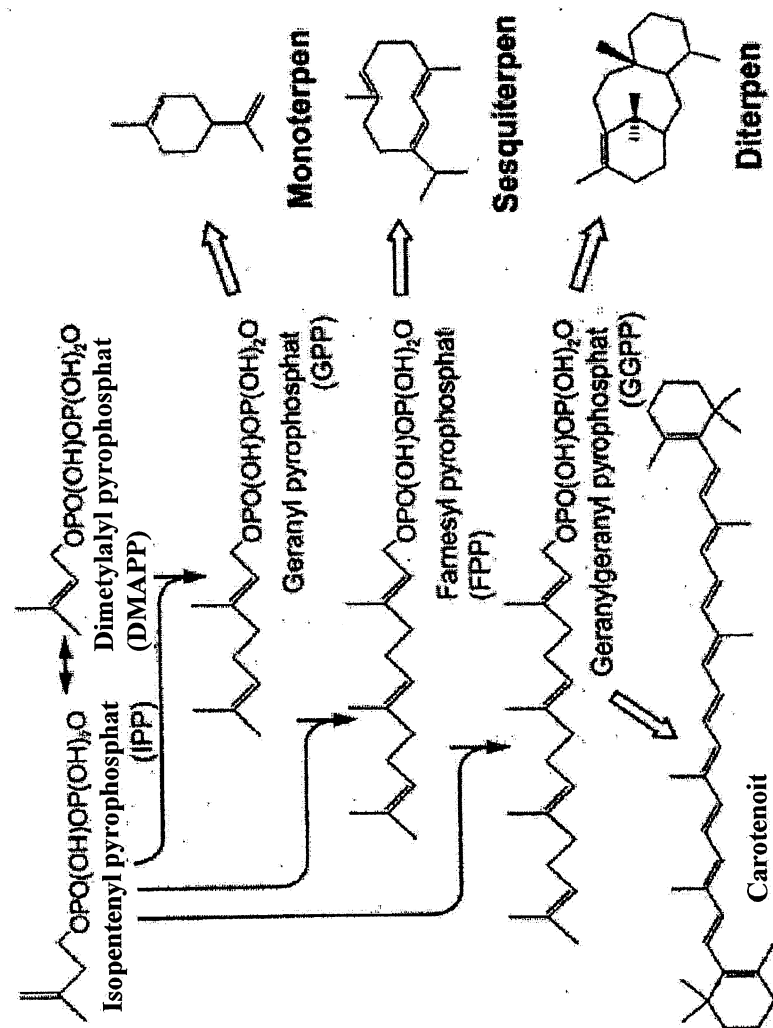


Fig. 3

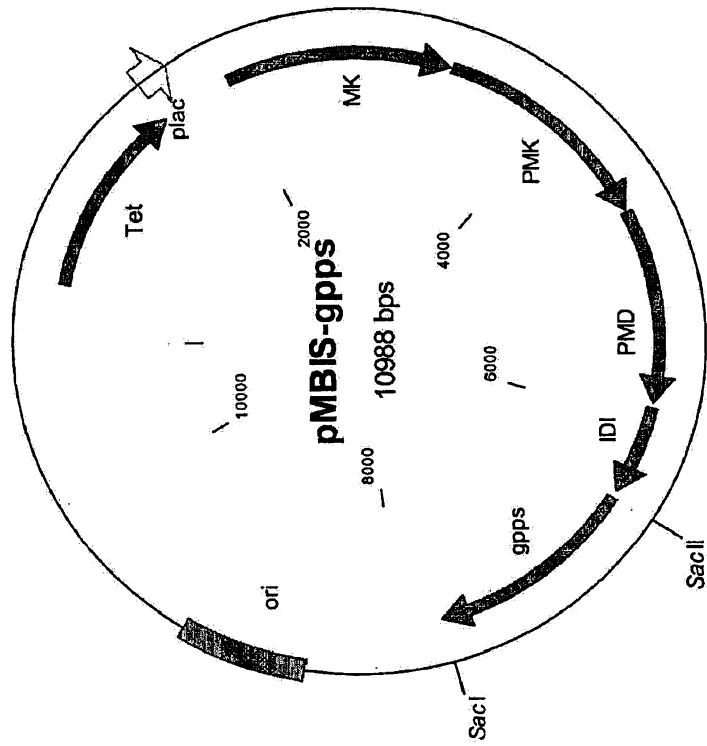


Fig. 4

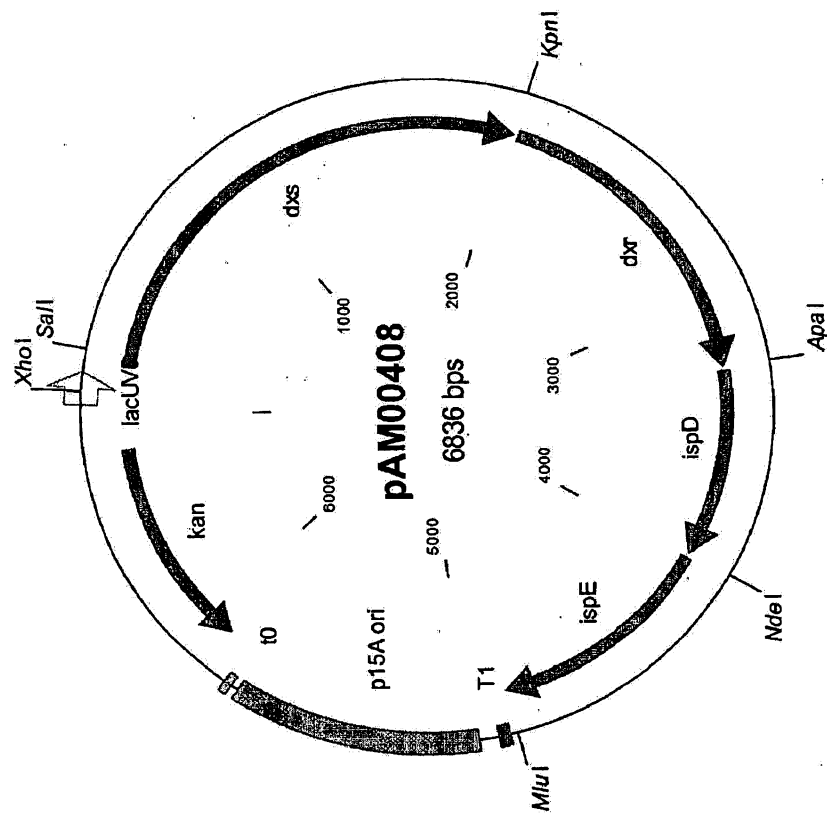


Fig. 5

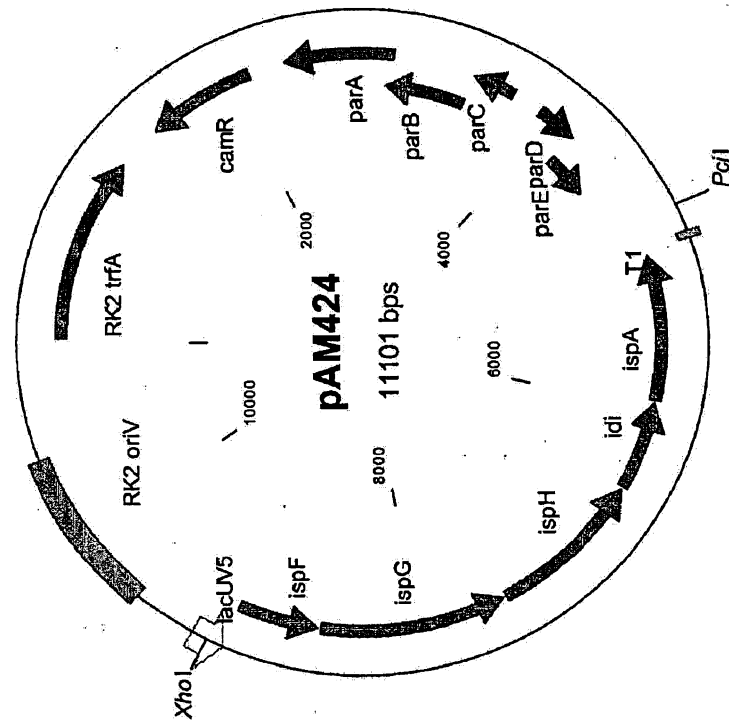
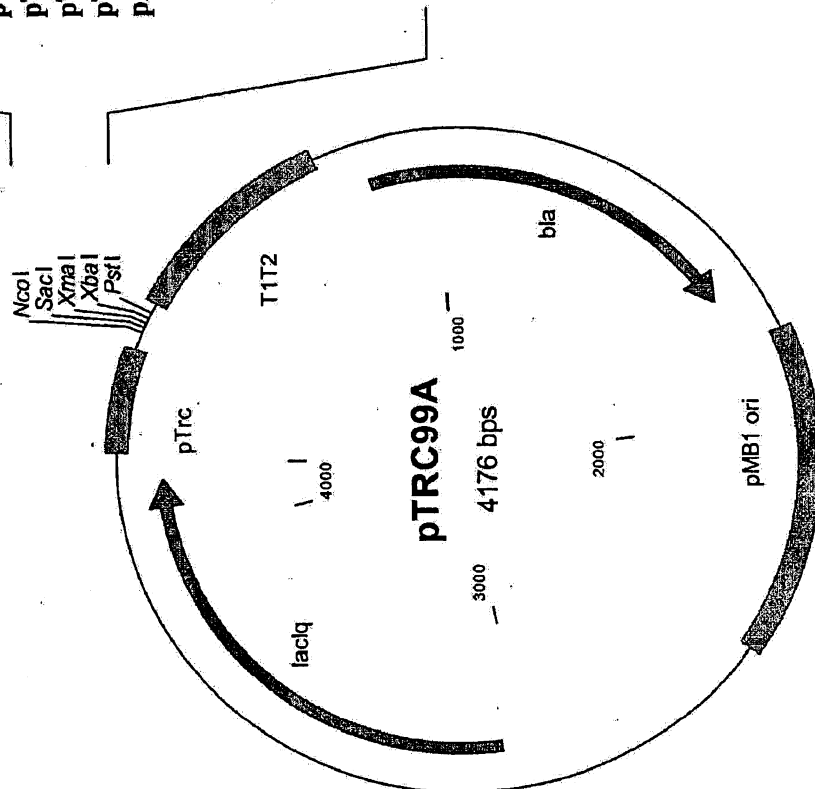


Fig. 6



pTrc99A-ADS: Đoạn chèn ADS ~1,6 kb ở vị trí *NcoI* *XmaI*
 pTrc99A-FSA: Đoạn chèn FSA ~1,7 kb ở vị trí *NcoI* *SacI*
 pTrc99A-GTS: Đoạn chèn GTS ~1,8 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-APS: Đoạn chèn APS ~1,9 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-TS: Đoạn chèn TS ~1,9 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-BPS: Đoạn chèn BPS ~1,7 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-PHS: Đoạn chèn PHS ~1,9 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-SS: Đoạn chèn SS ~1,8 kb ở vị trí *XmaI* *XbaI*
 pTrc99A-LLS: Đoạn chèn LLS ~1,7 kb ở vị trí *NcoI* *XmaI*
 pTrc99A-CS: Đoạn chèn CS ~1,8 kb ở vị trí *NcoI* *XmaI*
 pTrc99A-LMS: Đoạn chèn LMS ~1,9 kb ở vị trí *XbaI* *PstI*
 pAM373: Đoạn chèn FSA ~1,7 kb ở vị trí *NcoI* *SacI*

Fig.7

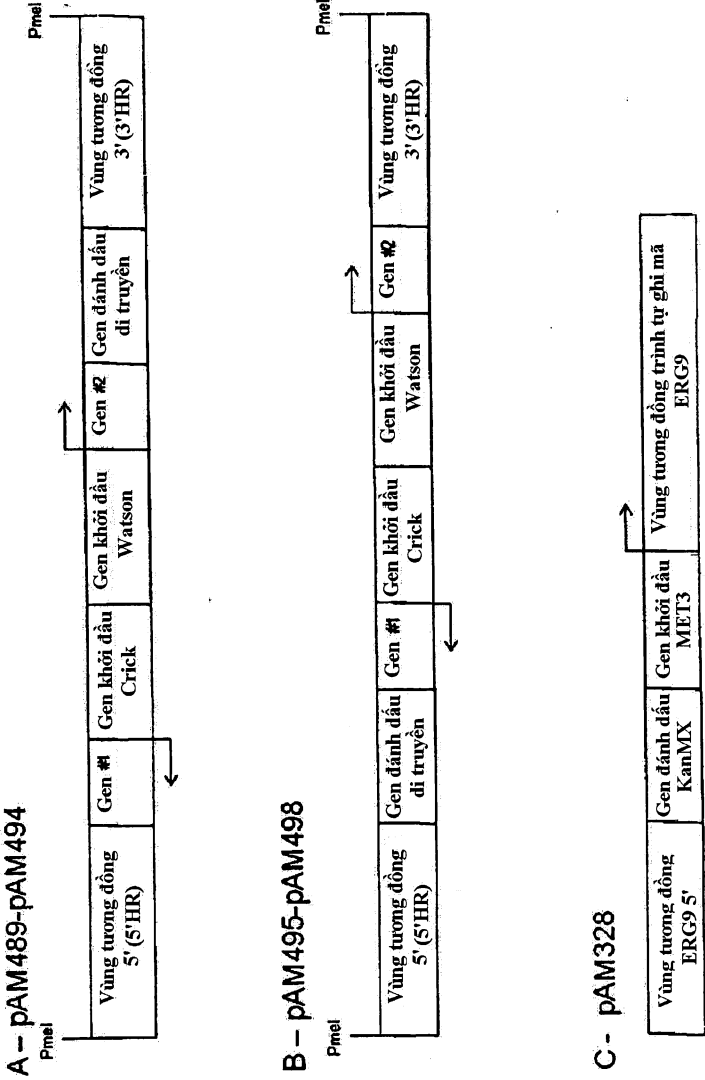


Fig. 8

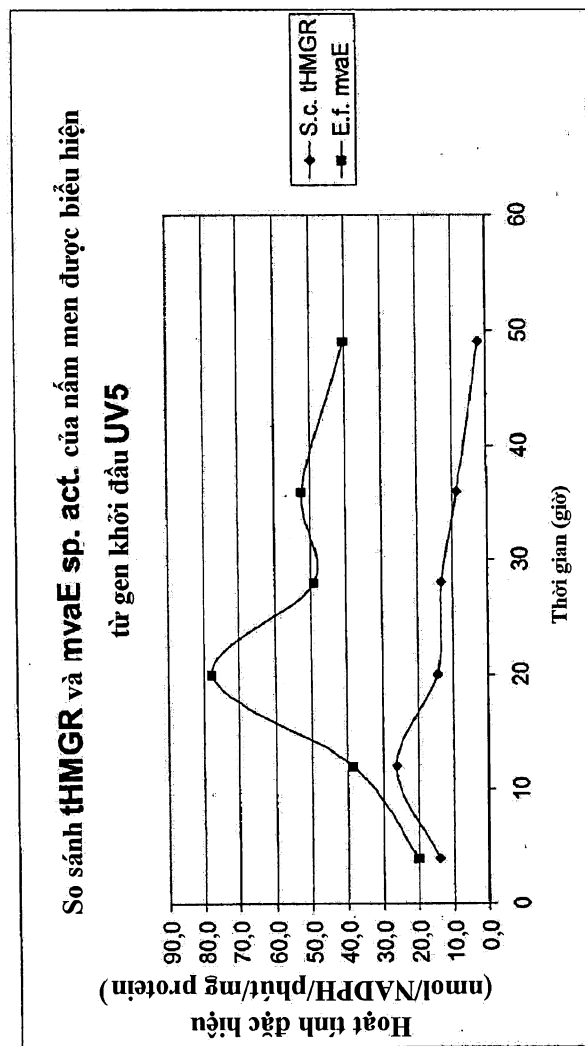
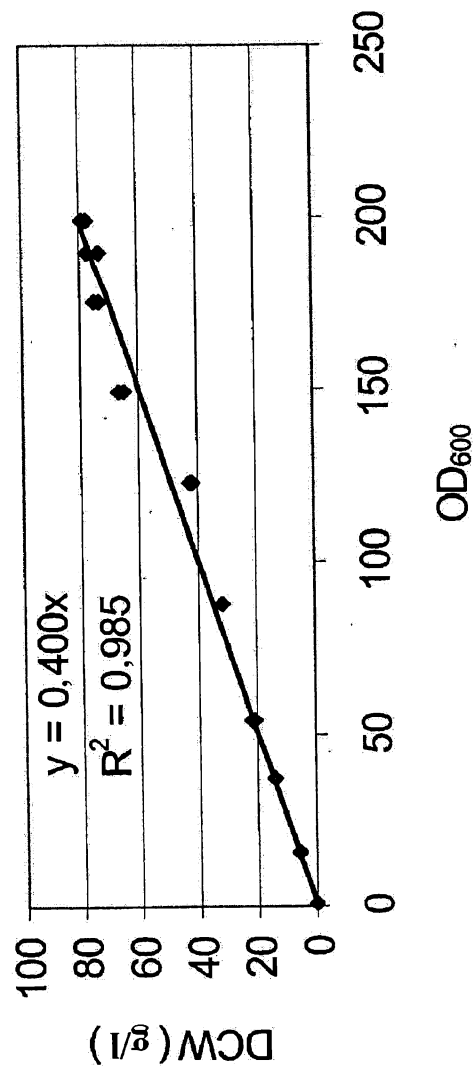


Fig. 9



11/20

Fig.10A-B

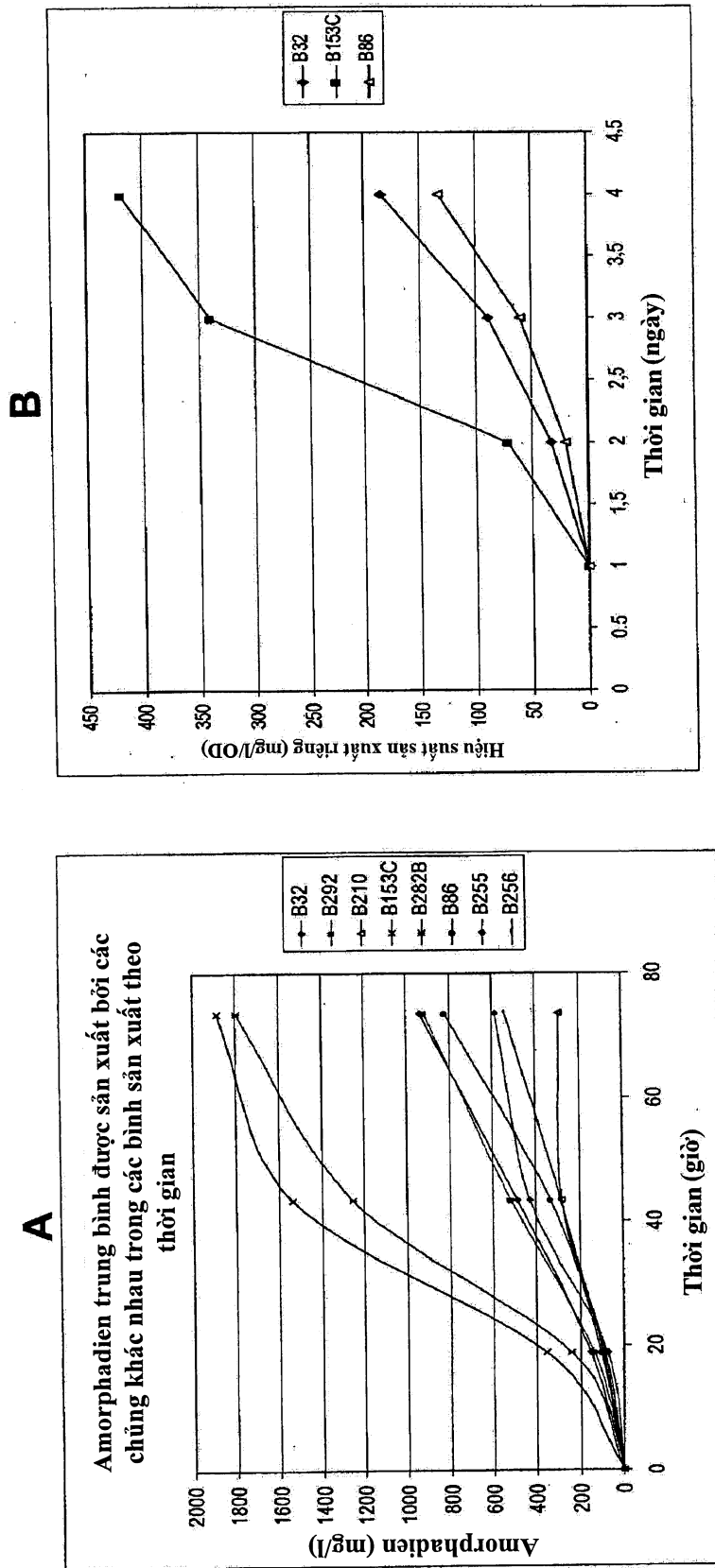


Fig.11A-B

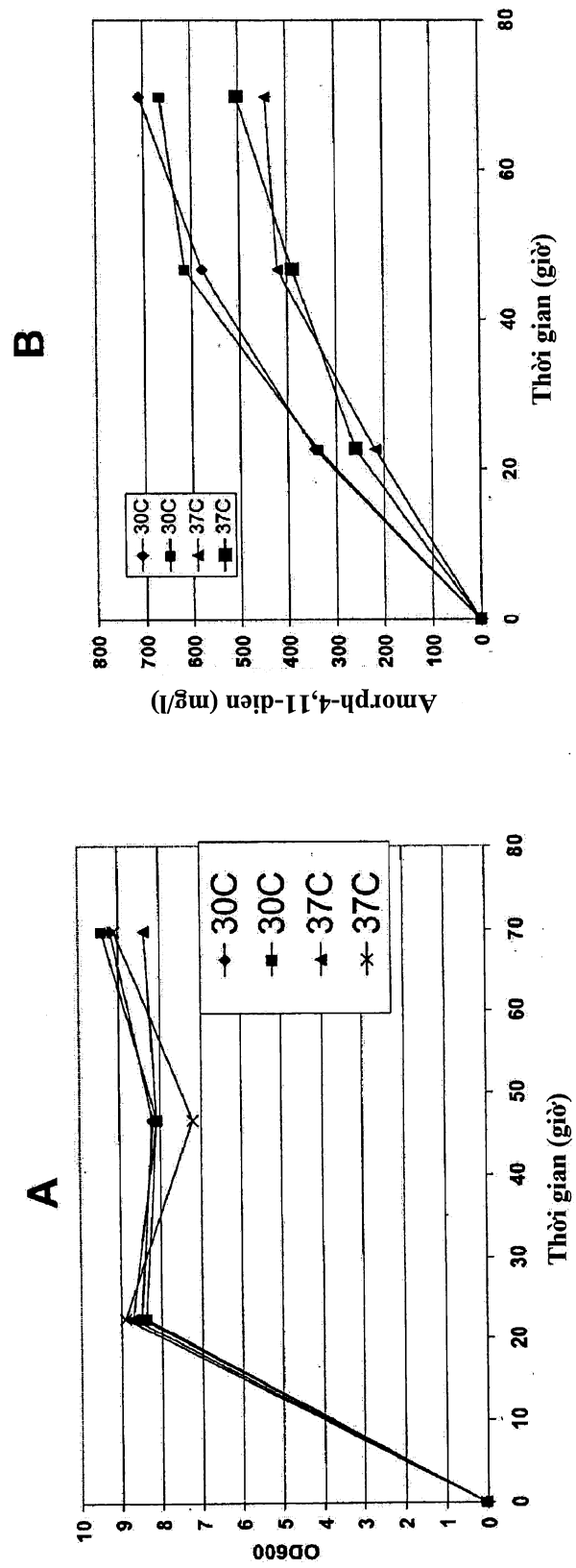


Fig.12A-B

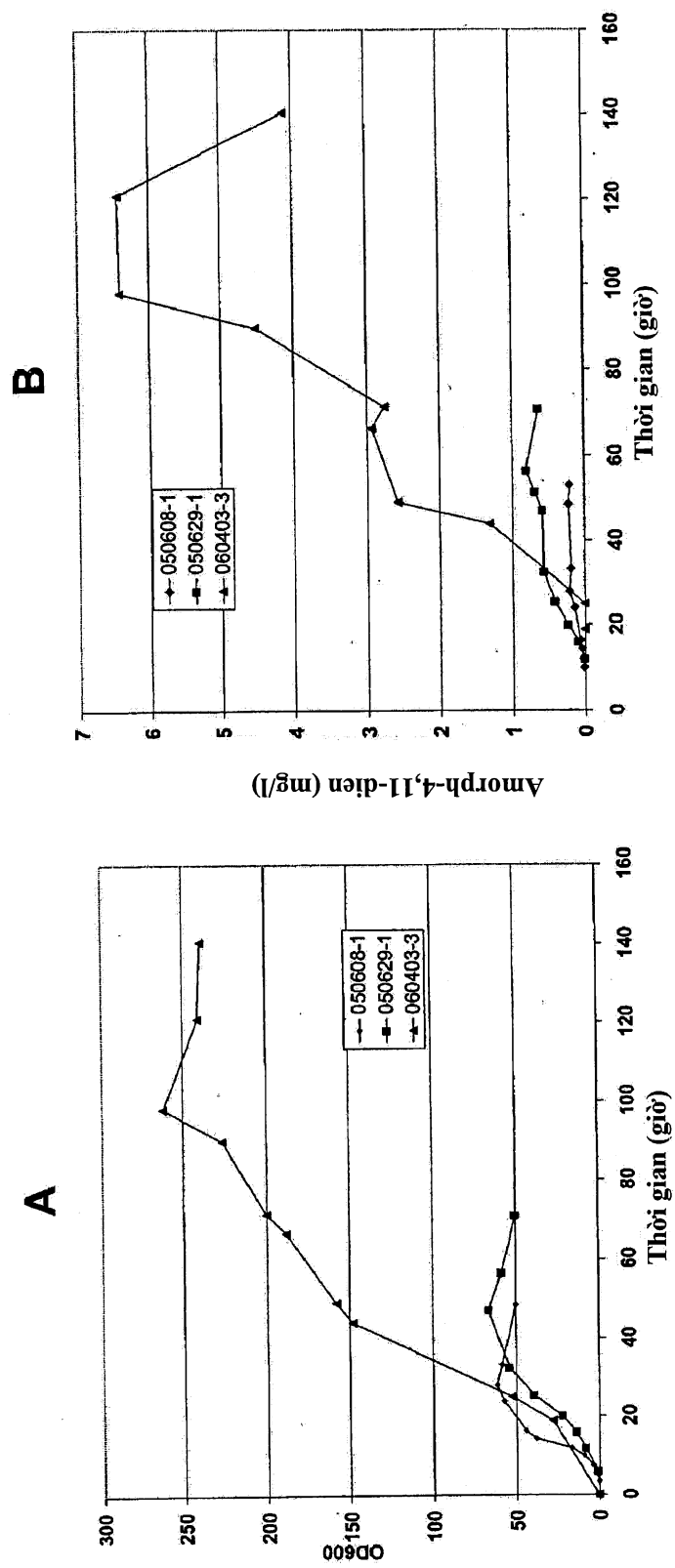


Fig.12C-D

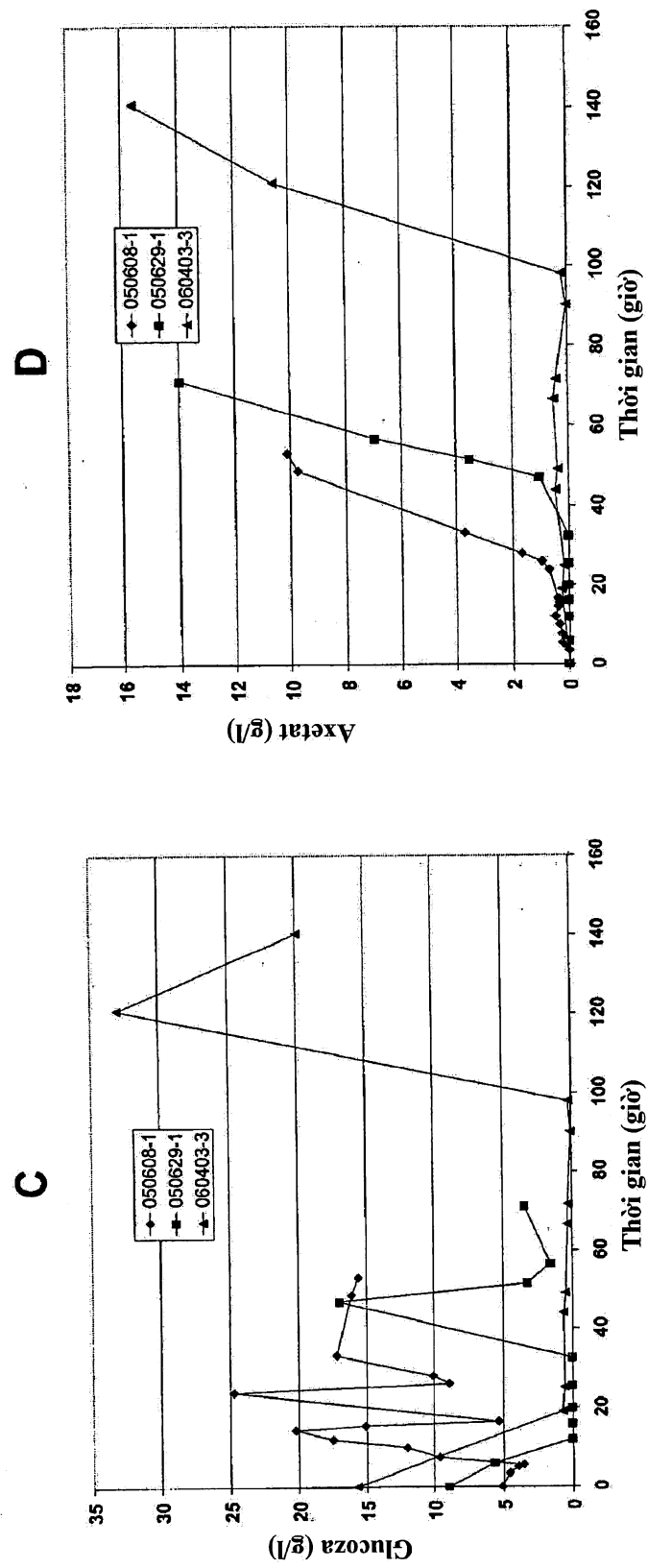


Fig.13A-B

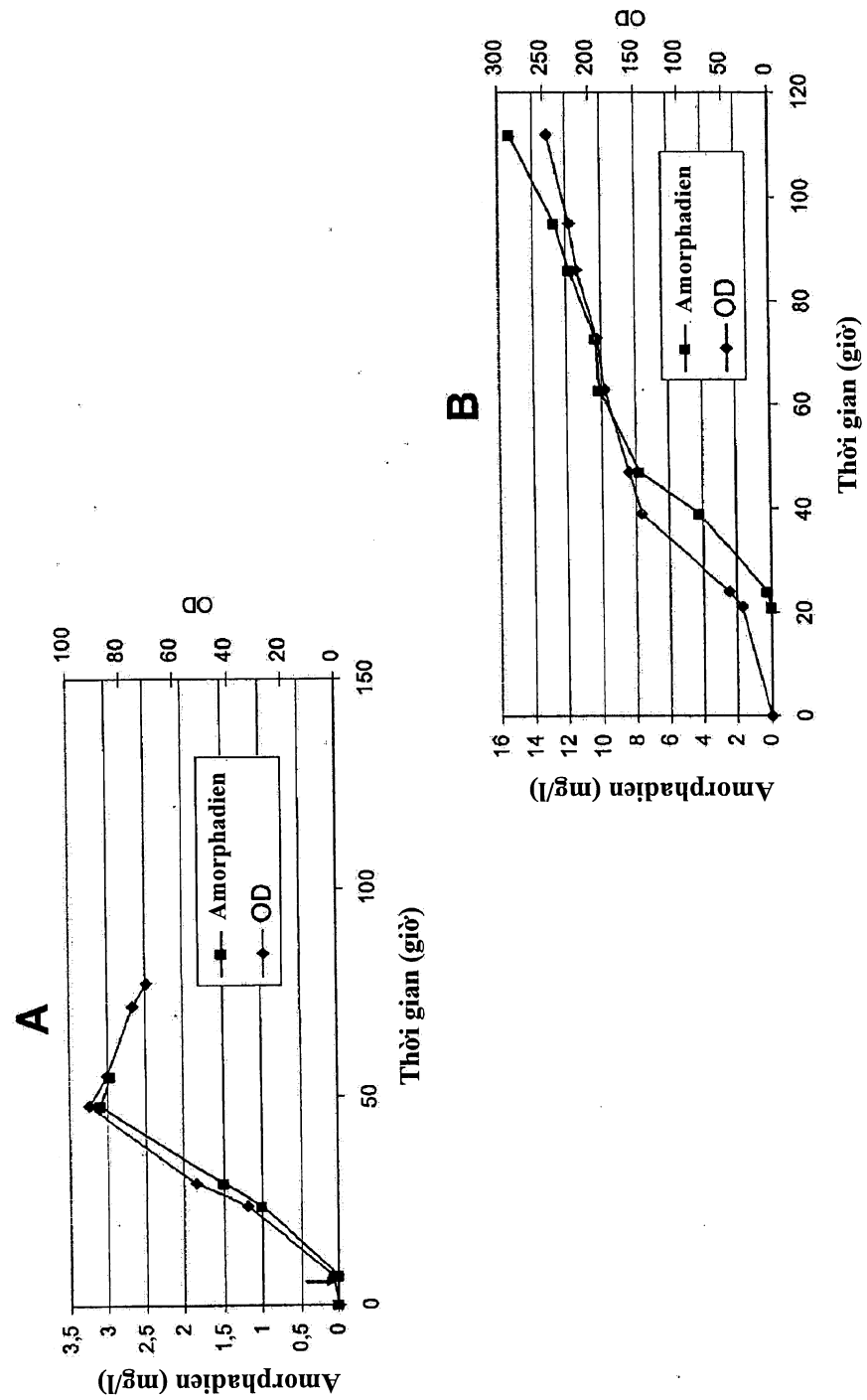


Fig.14A-B

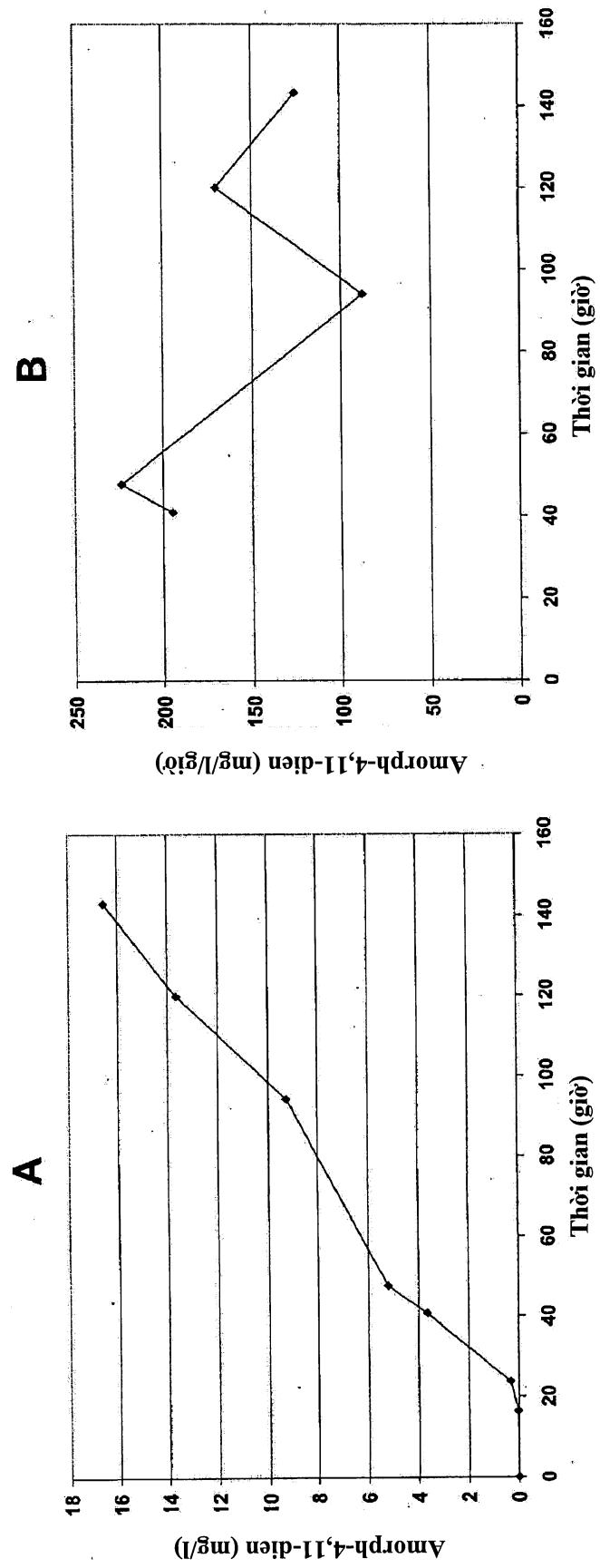


Fig.14C-E

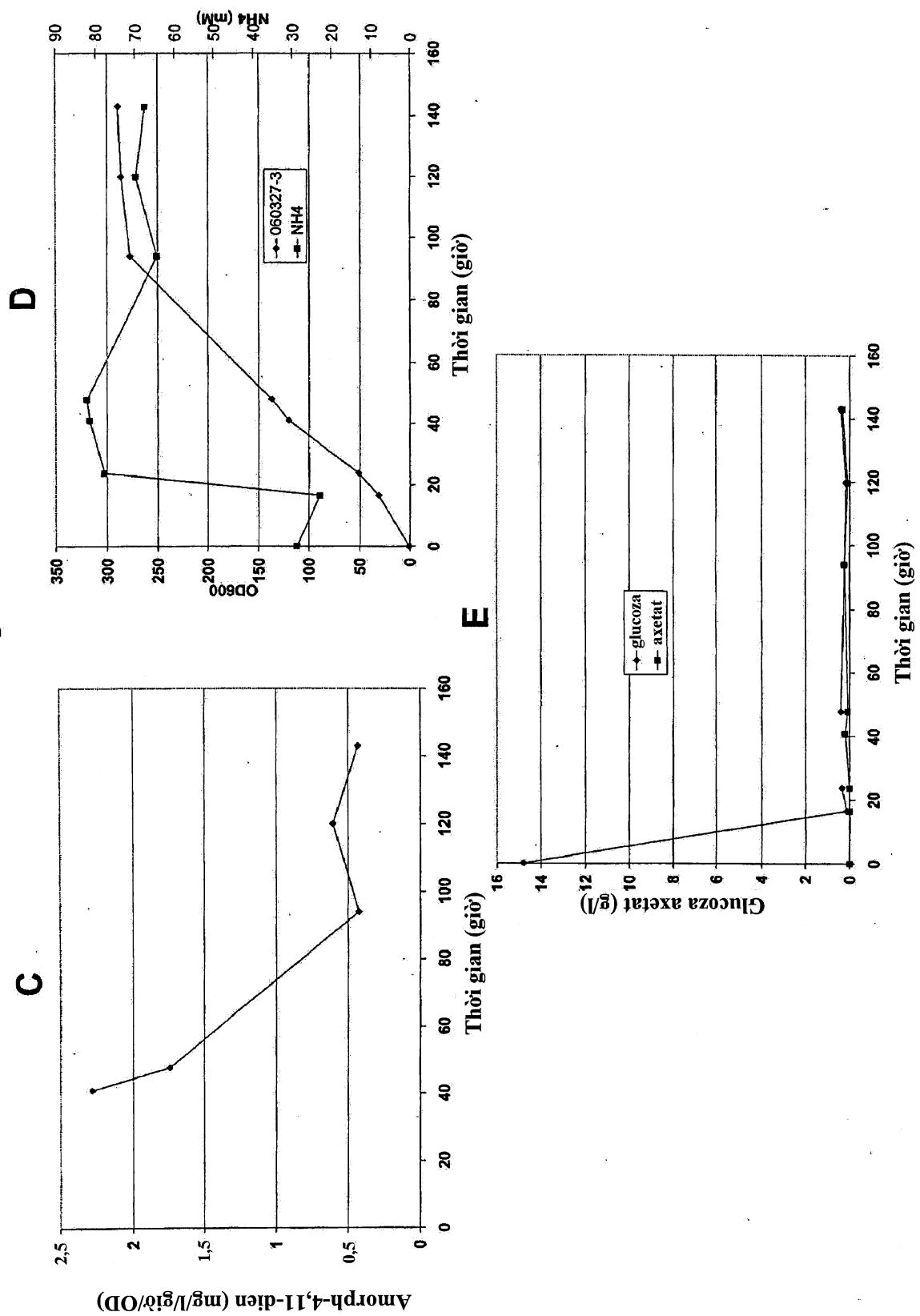


Fig.15A-B

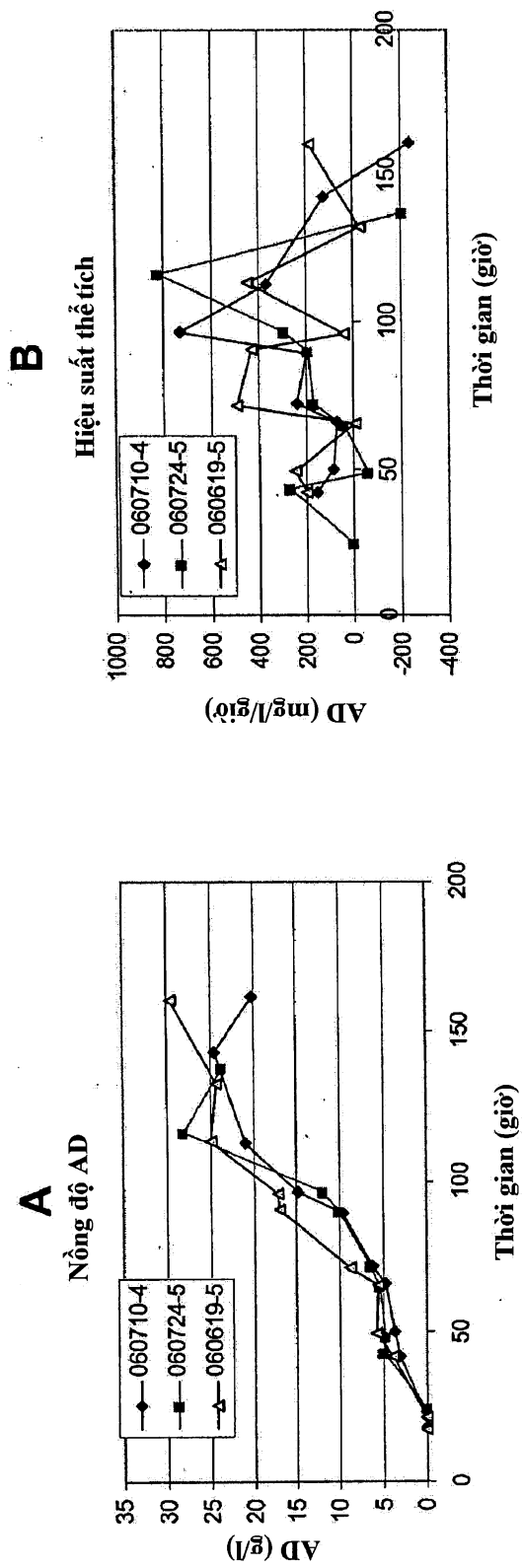


Fig.15C-E

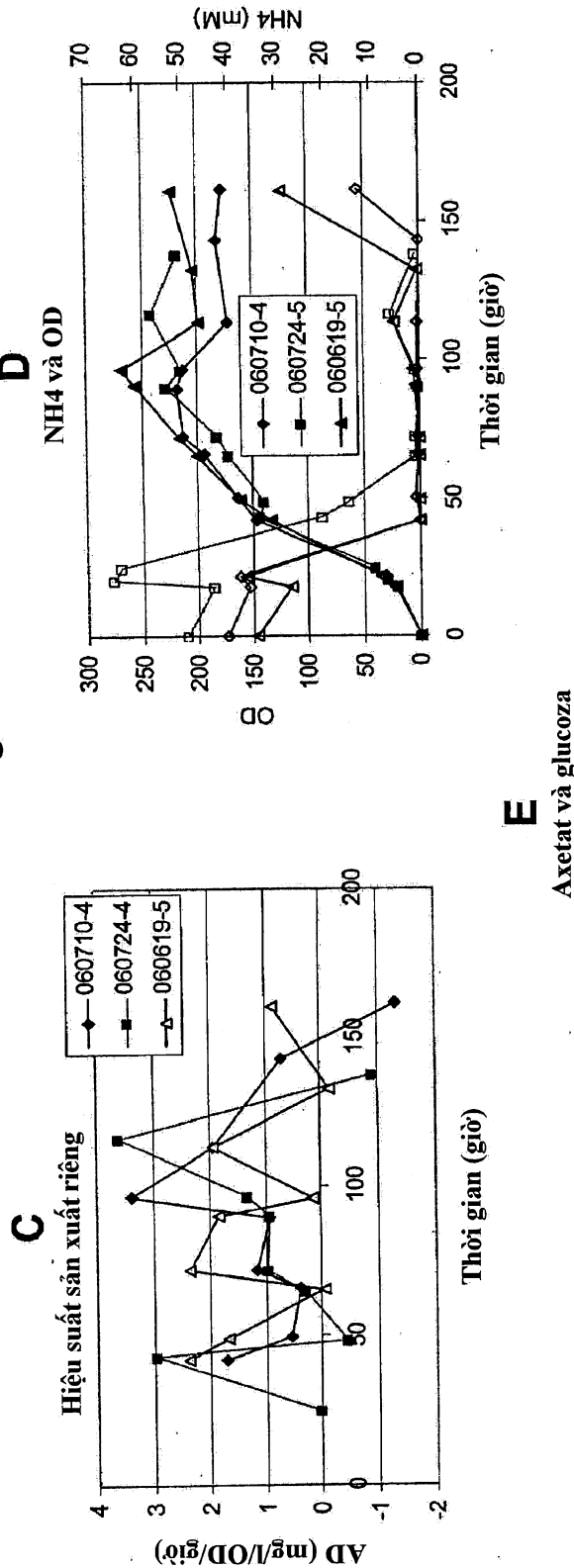
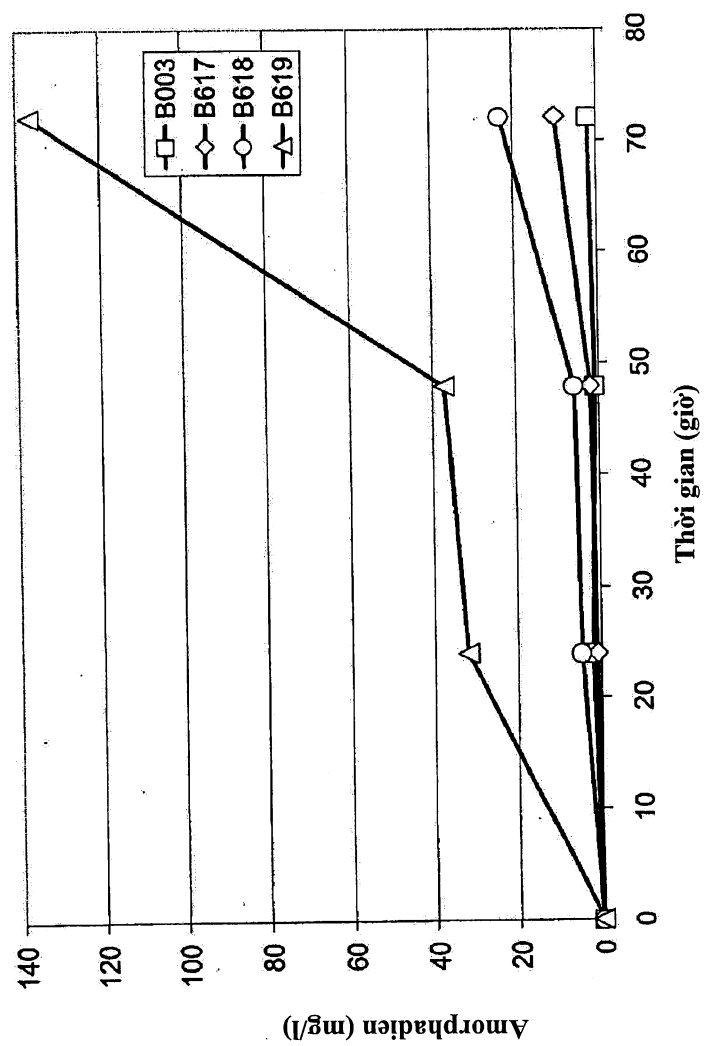


Fig.16



Danh mục trình tự

SEQ ID NO: 1

Operon MevT66

GAATTCAAAGGAGGAAAAATAAAATGAAGAACTGTGTGATTGTTTCTGCGGTCCGCACGGCGATCGGC
 AGCTTTAACGGCTCTTTAGCGAGCACCTCTGCAATCGATCTGGGTGCGACGGTCATTAAGGCCGCCATT
 GAACGCGCCAAAATCGACAGCCAGCACGTTGATGAGGTGATCATGGGCAATGTGTTACAAGCCGGCC
 TGGGTCAAAACCCAGCGCGTCAAGCACTGTTAAAATCTGGTCTGGCCGAGACCGTGTGTGGCTTCACC
 GTCAATAAGGTTTGCGGCTCTGGCCTGAAGAGCGTGGCCCTGGCAGCACAAGCGATTCAAGCCGGTCA
 GGCACAAAGCATCGTTGCGGGTGGCATGGAGAACATGTCTCTGGCGCCGTACTIONATTAGATGCCAAAG
 CCCGCAGCGGTTATCGCCTGGGCGATGGTCAGGTGTACGACGTCATCTTACGCGATGGCTTAATGTGC
 GCGACCCACGGTTACCACATGGGTATTACGGCCGAAAACGTGGCGAAAGAATACGGCATTACGCGCG
 AGATGCAGGATGAATTAGCACTGCACTCTCAGCGCAAAGCAGCAGCCGCGATCGAGTCTGGTGC GTTT
 ACGGCGGAAATCGTGCCAGTTAACGTGGTCACGCGCAAGAAGACGTTTCGTTTTTCAGCCAGGACGAGTT
 CCCGAAGGCAAACAGCACCGCGGAGGCCTTAGGTGCCTTACGCCCAGCCTTTGACAAAGCGGGCACG
 GTCACCGCCGGTAATGCGAGCGGCATCAATGATGGTGCAGCGGCACTGGTCATCATGGAAGAGAGCG
 CCGCATTAGCAGCGGGTCTGACCCCATTAGCGCGCATTAAATCTTATGCCAGCGGCGGCGTCCCACCA
 GCCCTGATGGGCATGGGTCCGGTCCCAGCCACGCAAAAAGCCCTGCAATTAGCGGGCCTGCAACTGGC
 CGACATTGATCTGATCGAGGCGAACGAGGCGTTTGACGCGCAGTTCCCTGGCGGTGGGTAAGAATCTGG
 GCTTCGACAGCGAGAAAGTCAATGTGAACGGTGGCGCGATTGCGTTAGGCCATCCGATTGGTGCAAGC
 GGCGCACGCATCTTAGTGACGTTACTGCACGCCATGCAGGCACGCGACAAGACCTTAGGCCTGGCGAC
 CTTATGTATTGGTGGCGGTCAAGGTATCGCCATGGTGATCGAACGCCTGAACTGAAGATCTAGGAGGA
 AAGCAAAATGAACTGAGCACCAAGCTGTGCTGGTGTGGCATCAAGGGTCGCCTGCGCCACAAAAG
 CAGCAACAGCTGCACAACACGAACCTGCAAATGACCGAGCTGAAAAAGCAGAAGACGGCCGAGCAA
 AAGACCCGCCCCGAGAACGTTGGCATCAAGGGCATCCAGATTTATATCCCGACGCAGTGTGTCAACCA
 ATCTGAGCTGGAGAAATTCGATGGCGTCAGCCAGGGTAAGTACACCATCGGCCTGGGCCAGACCAAC
 ATGAGCTTCGTGAACGACCGTGAGGACATCTATTCTATGAGCCTGACGGTGTGTCTAAGCTGATCAA
 GAGCTACAACATCGACACGAATAAGATCGGTCTGTGGAGGTGGGTACGGAGACGCTGATTGACAAG
 AGCAAAAGCGTGAAGTCTGTCTTAATGCAGCTGTTTCGGCGAGAACACGGATGTGAGGGTATCGACAC
 CCTGAACGCGTGTTACGGCGGCACCAACGCACTGTTCAATAGCCTGAACTGGATTGAGAGCAACGCCT
 GGGATGGCCGCGATGCGATCGTCTGTGTGCGGCGATATCGCCATCTATGACAAGGGTGGCGCACGTCCG
 ACCGGCGGTGCAGGCACCGTTGCGATGTGGATTGGCCCGGACGCACCAATTGTCTTCGATTCTGTCCG
 CGCGTCTTACATGGAGCACGCCTACGACTTTTACAAGCCGGACTTCACGAGCGAATACCCGTACGTGG
 ACGGCCACTTCTCTCTGACCTGCTATGTGAAGGCGCTGGACCAGGTTTATAAGTCTTATAGCAAAAAG
 GCGATTTCTAAGGGCCTGGTCAGCGACCCGGCAGGCAGCGACGCCCTGAACGTGCTGAAGTATTTTGA
 CTACAACGTGTTCCATGTCCCGACCTGCAAATTAGTGACCAAATCTTATGGCCGCCTGTTATATAATGA
 TTTCCGTGCCAACCCGACGCTGTTCCCGGAGGTTGACGCCGAGCTGGCGACGCGTGATTACGACGAGA
 GCCTGACCGACAAGAACATCGAGAAGACCTTCGTCAACGTGCGGAAGCCGTTCCACAAAGAGCGTGT
 GGCCCAAAGCCTGATCGTCCCGACCAACACGGGCAACATGTATACCGCGTCTGTCTACGCGGCATTTCG
 CGAGCCTGCTGAATTACGTCGGTTCTGACGACCTGCAGGGCAAGCGCGTTGGCCTGTTACGCTACGGT
 AGCGGCTTAGCGGCCAGCCTGTATAGCTGCAAAATTGTGCGCGACGTCCAGCACATCATCAAGGAGCT

GGACATCACCAACAAGCTGGCGAAGCGCATCACCGAGACGCCGAAAGATTACGAGGCAGCGATCGAG
 TTACGCGAGAATGCGCATCTGAAGAAGAACTTCAAGCCGCAAGGTAGCATCGAGCACCTGCAGAGCG
 GCGTCTACTACCTGACGAACATTGACGACAAGTTCCGCCGTTCTTATGACGTCAAAAAGTAACTAGTA
 GGAGGAAAACATCATGGTGCTGACGAACAAAACCGTCATTAGCGGCAGCAAGGTGAAGTCTCTGAGC
 AGCGCCCAAAGCTCTAGCAGCGGCCCGTCTAGCAGCAGCGAGGAGGACGACAGCCGTGACATTGAGT
 CTCTGGACAAGAAGATCCGCCCGCTGGAGGAGTTAGAGGGCCCTGCTGAGCAGCGGCAACACCAAGCA
 GCTGAAGAACAAGGAAGTTGCAGCGCTGGTGATCCACGGTAAGCTGCCACTGTATGCGCTGGAAAAG
 AAAGTGGGCGATACGACGCGTGCGGTGCGGTGCGTCGCAAAGCCTTAAGCATCTTAGCGGAGGCCCC
 GGTGTTAGCCAGCGACCGCCTGCCGTACAAGAACTACGACTACGACCGCGTGTTTGGCGCGTGCTGCG
 AGAATGTCATTGGCTACATGCCGTTACCGGTTGGTGATCGGCCCGCTGGTTCATTGATGGCACGAGC
 TATCACATTCCAATGGCGACCACGGAAGGTTGCTTAGTCGCCAGCGCCATGCGTGGCTGTAAGGCGAT
 TAACGCCGGCGGTGGCGCGACGACCGTGTTAACCAAGGATGGTATGACGCGCGGTCCGGTCGTCCGCT
 TCCCAACGCTGAAGCGCAGCGGCGCGTGTAAGATTTGGCTGGATTCTGAGGAGGGCCAAAACGCGAT
 CAAGAAAGCCTTCAACTCTACGAGCCGTTTCGCGCGTTTACAGCATATCCAGACCTGCCTGGCCGGCG
 ACCTGCTGTTTCATGCGCTCCGCAACCACCGGCGATGCGATGGGCATGAACATGATCAGCAAGGGC
 GTCGAATATAGCCTGAAACAAATGGTGGAAGAATATGGCTGGGAGGACATGGAGGTTGTCTCTGTGA
 GCGGCAACTATTGCACCGACAAGAAGCCGGCAGCCATTAAGTGGATTGAGGGTCGCGGCAAAAAGCGT
 CGTGGCAGAAGCGACCATCCCAGGCGACGTGGTCCGTAAGGTTCTGAAGAGCGACGTCAGCGCCCTG
 GTTGAGTTAAATATCGCGAAAAACCTGGTCGGCAGCGCGATGGCGGGCAGCGTGGGTGGCTTTAACGC
 ACATGCAGCGAATCTGGTTACGGCGGTTTTCTTAGCCTTAGGTCAGGACCCAGCCAAAATGTGCGAGA
 GCAGCAACTGCATTACCTTAATGAAAGAGGTTGACGGTGACCTGCGCATCAGCGTTTCTATGCCGTCT
 ATCGAGGTGCGCACGATCGGCGGCGGCACCGTTTTAGAACCGCAAGGTGCGATGCTGGATCTGCTGGG
 CGTGGCGGCGCCACATGCAACGGCCCCAGGCACCAATGCCCGCCAACTGGCCCGTATCGTGGCCTGCG
 CGGTTCTGGCGGGTGAGCTGAGCCTGTGCGCCGCATTAGCCGCGGGCCATTTAGTTCAATCTCACATG
 ACCCACAAACCGCAAGCCGGCAGAACCAACCAAGCCAAATAACCTGGACGCAACCGACATTAACCGTC
 TGAAGGATGGCAGCGTCACGTGCATTAAAAGCTGAGCATGCTACTAAGCTT

SEQ ID NO: 2

Gen khởi đầu 4-49 mvaA SpeI

5'-GCTACTAGTAGGAGGAAAACATCATGCAAAGTTTAGATAAGAATTTCCG-3'

SEQ ID NO: 3

Gen khởi đầu 4-49 mvaAR XbaI

5'-GCTTCTAGACTATTGTTGTCTAATTTCTTGTAATAATGCG-3'

SEQ ID NO: 4

Gen khởi đầu HMGS 5' Sa mvaS-S

5'-GAACTGAAGATCTAGGAGGAAAGCAAAATGACAATAGGTATCGACAAAATAAACT-3'

SEQ ID NO: 5

Gen khởi đầu HMGS 3' Sa mvaS-AS

5'-TTGCATGATGTTTTCTCCTACTAGTTACTCTGGTCTGTGATATTCGCGAAC-3'

SEQ ID NO: 6Gen khởi đầu 19-25 *atoB* SfiI-S

5'-GCTAGGCCATCCTGGCCATGAAGAACTGTGTGATTGTTTCTG-3'

SEQ ID NO: 7Mồi 19-25 *mvaA*-AsiSI-AS

5'-GCTTGC GATCGCCGGCGGATTTGTCCTACTCAG-3'

SEQ ID NO: 8

Gen khởi đầu 9-70C

5'-CCACCTCGAGATGTCATTACCGTTCTTAACCTTCTG-3'

SEQ ID NO: 9

Gen khởi đầu 26-39B

5'-TGGTGGAGCTCTTATTTAAGCTGGGTAAATGCAGATAATCG-3'

SEQ ID NO: 10

Gen khởi đầu 26-39A

5'-TTCTTGAGCTCTTATTCCTTTGGTAGACCAGTCTTTGCG-3'

SEQ ID NO: 11Gen khởi đầu 4-40 *mvaEF* BamHI

5' – TATGGATCCTAAGGAGGATATTTAGATGAAAACAGTAGTTATTATTGATGC – 3'

SEQ ID NO: 12Gen khởi đầu 4-40 *mvaER* HindIII

5'- AGCTAAGCTTTTATTGTTTTCTTAAATCATTTAAAATAGC – 3'

SEQ ID NO: 13Gen khởi đầu 4-40 *mvaSF* BglII

5' – TATAGATCTTAAGGAGGATATTTAGATGACAATTGGGATTGATAAAAATTAG – 3'

SEQ ID NO: 14Gen khởi đầu 4-39 *mvaSR* BamHI

5' – TTTGGATCCTTAGTTTCGATAAGAGCGAACGG – 3'

SEQ ID NO: 15Gen khởi đầu 67-1A-C để khuếch đại PCR trình tự ghi mã của gen *dxs*

5'- ACA CTC GAG GAG GAA TAA ATG AGT TTT GAT ATT GCC AAA TAC CCG -3'

SEQ ID NO: 16

Gen khởi đầu 67-1B-C để khuếch đại PCR trình tự ghi mã của gen dxs
5'- TGA TGG TAC CTT ATG CCA GCC AGG CCT TGA TTT TGG C -3'

SEQ ID NO: 17

Gen khởi đầu 67-1C-C để khuếch đại PCR trình tự ghi mã của gen dxr
5'- ACT AGG TAC CAG GAG GAA TAA ATG AAG CAA CTC ACC ATT CTG GGC -3'

SEQ ID NO: 18

Gen khởi đầu 67-1D-C để khuếch đại PCR trình tự ghi mã của gen dxr
5'- AAT TGA TGG GCC CTC AGC TTG CGA GAC GCA TCA CCT C -3'

SEQ ID NO: 19

Gen khởi đầu 67-1E-C để khuếch đại PCR trình tự ghi mã của gen ispD
5'- CAT AAA GGG CCC AGG AGG AAT AAA TGG CAA CCA CTC ATT TGG ATG -3'

SEQ ID NO: 20

Gen khởi đầu 67-1F-C để khuếch đại PCR trình tự ghi mã của gen ispD
5'- TAT TGT TCA TAT GTT ATG TAT TCT CCT GAT GGA TGG TTC G -3'

SEQ ID NO: 21

Gen khởi đầu 67-1G-C để khuếch đại PCR trình tự ghi mã của gen ispE
5'- AAC TAA CAC ATA TGA GGA GGA ATA AAT GCG GAC ACA GTG GCC CTC -3'

SEQ ID NO: 22

Gen khởi đầu 67-1H-C để khuếch đại PCR trình tự ghi mã của gen ispE
5'- TGT TAG TTA CGC GTT TAA AGC ATG GCT CTG TGC AAT GG -3'

SEQ ID NO: 23

Gen khởi đầu 67-2A-C để khuếch đại PCR trình tự ghi mã của gen ispF
5'- ACG GGA TCC AGG AGG AAT AAA TGC GAA TTG GAC ACG GTT TTG ACG -3'

SEQ ID NO: 24

Gen khởi đầu 67-2B-C để khuếch đại PCR trình tự ghi mã của gen ispF
5'- TTT AGT TGG GCC CTC ATT TTG TTG CCT TAA TGA GTA GCG CC -3'

SEQ ID NO: 25

Gen khởi đầu 67-2C-C để khuếch đại PCR trình tự ghi mã của gen ispG
5'- TAC TAA GGG CCC AGG AGG AAA TAA TGC ATA ACC AGG CTC CAA TTC AAC G -3'

SEQ ID NO: 26

Gen khởi đầu 67-2D-C để khuếch đại PCR trình tự ghi mã của gen ispG
5'- TCC GGG TAC CTT ATT TTT CAA CCT GCT GAA CGT CAA TTC G -3'

SEQ ID NO: 27

Gen khởi đầu 67-2E-C để khuếch đại PCR trình tự ghi mã của gen *ispH*

5'- AAC AGG TAC CAG GAG GAA ATA ATG CAG ATC CTG TTG GCC AAC C -3'

SEQ ID NO: 28

Gen khởi đầu 67-2F-C để khuếch đại PCR trình tự ghi mã của gen *ispH*

5'- TGG ATG AAG TCG ACT TAA TCG ACT TCA CGA ATA TCG ACA CGC AGC -3'

SEQ ID NO: 29

Gen khởi đầu 67-2G-C để khuếch đại PCR trình tự ghi mã của gen *idi*

5'- CAT CAA GTC GAC AGG AGG AAA TAA TGC AAA CGG AAC ACG TCA TTT TAT TG -3'

SEQ ID NO: 30

Gen khởi đầu 67-2H-C để khuếch đại PCR trình tự ghi mã của gen *idi*

5'- TAA TGC AAG CTT ATT TAA GCT GGG TAA ATG CAG ATA ATC G -3'

SEQ ID NO: 31

Gen khởi đầu 67-2I-C để khuếch đại PCR trình tự ghi mã của gen *ispA*

5'- CAG TAA AGC TTA GGA GGA AAT AAT GGA CTT TCC GCA GCA ACT CG -3'

SEQ ID NO: 32

Gen khởi đầu 67-2J-C để khuếch đại PCR trình tự ghi mã của gen *ispA*

5'- TAG TTC CAT GGT TAT TTA TTA CGC TGG ATG ATG TAG TCC GC -3'

SEQ ID NO: 33

Gen khởi đầu 9-156A để khuếch đại PCR của locus *RK2 par*

5'- ACATAGACGTCGGGAAAGCGAGGATCTAGGTAGGG -3'

SEQ ID NO: 34

Gen khởi đầu 9-156B để khuếch đại PCR của locus *RK2 par*

5'- TTCCCGCTCGAGGTGGCGGACCATATAGGCAGATCAG -3'

SEQ ID NO: 35

Gen khởi đầu 19-137 *cml-pAM37-AS*

5' – GACGTCGATATCTGGCGAAAATG – 3'

SEQ ID NO: 36

Gen khởi đầu 19-137 *cml-pAM37-S*

5' – TACTAGTGCTTGGATTCTCACC – 3'

SEQ ID NO: 37

Gen khởi đầu để khuếch đại PCR trình tự nucleotit mã hoá cho β -farnesen synthaza
 5'-CCATGGGACACTCTGCCGATCTCTCCGTAAGC-3'

SEQ ID NO: 38

Gen khởi đầu để khuếch đại PCR trình tự nucleotit mã hoá cho β -farnesen synthaza
 5'-GAGCTCTCATACGACCATAGGGTGTACG-3'

SEQ ID NO: 39

Gen khởi đầu để khuếch đại PCR trình tự nucleotit mã hoá cho α -farnesen synthaza
 5'-CCATGGACCTGGCAGTAGAAATTGC-3'

SEQ ID NO: 40

Gen khởi đầu để khuếch đại PCR trình tự nucleotit mã hoá cho α -farnesen synthaza
 5'-GAGCTCTTACATCGGTACCGGCTCCAG-3'

SEQ ID NO: 41

Operon *atoB*(opt):*HMGs*(opt):*mvaA*

ATGAAGAACTGTGTGATTGTTTCTGCGGTCCGCACGGCGATCGGCAGCTTTAACGGCTCTTTAGCGAG
 CACCTCTGCAATCGATCTGGGTGCGACGGTCATTAAGGCCGCCATTGAACGCGCCAAAATCGACAGCC
 AGCACGTTGATGAGGTGATCATGGGCAATGTGTTACAAGCCGGCCTGGGTCAAAACCCAGCGCGTCAA
 GCACTGTTAAAATCTGGTCTGGCCGAGACCGTGTGTGGCTTACCGTCAATAAGGTTTGCGGCTCTGGC
 CTGAAGAGCGTGGCCCTGGCAGCACAAAGCGATTCAAGCCGGTCAGGCACAAAGCATCGTTGCGGGTG
 GCATGGAGAACATGTCTCTGGCGCCGTAATTATTAGATGCCAAAGCCCGCAGCGGTTATCGCCTGGGC
 GATGGTCAGGTGTACGACGTCATCTTACGCGATGGCTTAATGTGCGCGACCCACGGTTACCACATGGG
 TATTACGGCCGAAAACGTGGCGAAAGAATACGGCATTACGCGCGAGATGCAGGATGAATTAGCACTG
 CACTCTCAGCGCAAAGCAGCAGCCGCGATCGAGTCTGGTGC GTTTACGGCGGAAATCGTGCCAGTTAA
 CGTGGTCACGCGCAAGAAGACGTTTCGTTTTACGCCAGGACGAGTTCCCGAAGGCAAACAGCACCGCG
 GAGGCCTTAGGTGCCTTACGCCCAGCCTTTGACAAAGCGGGCACGGTCACCGCCGGTAATGCGAGCGG
 CATCAATGATGGTGCAGCGGCACTGGTCATCATGGAAGAGAGCGCCGCATTAGCAGCGGGTCTGACCC
 CATTAGCGCGCATTAATCTTATGCCAGCGGCGGCGTCCACCAGCCCTGATGGGCATGGGTCCGGTC
 CCAGCCACGCAAAAAGCCCTGCAATTAGCGGGCCTGCAACTGGCCGACATTGATCTGATCGAGGCGA
 ACGAGGCGTTTGCAGCGCAGTTCCTGGCGGTGGGTAAGAATCTGGGCTTCGACAGCGAGAAAGTCAAT
 GTGAACGGTGGCGCGATTGCGTTAGGCCATCCGATTGGTGCAAGCGGCGCACGCATCTTAGTGACGTT
 ACTGCACGCCATGCAGGCACGCGACAAGACCTTAGGCCTGGCGACCTTATGTATTGGTGGCGGTCAAG
 GTATCGCCATGGTGATCGAACGCCTGAACTGAAGATCTAGGAGGAAAGCAAAATGAAACTGAGCACC
 AAGCTGTGCTGGTGTGGCATCAAGGGTGCCTGCGCCACAAAAGCAGCAACAGCTGCACAACACGA
 ACCTGCAAATGACCGAGCTGAAAAAGCAGAAGACGGCCGAGCAAAAGACCCGCCCGCAGAACGTTGG
 CATCAAGGGCATCCAGATTTATATCCCGACGCAGTGTGTCAACCAATCTGAGCTGGAGAAATTTCGATG
 GCGTCAGCCAGGGTAAGTACACCATCGGCCTGGGCCAGACCAACATGAGCTTCGTGAACGACCGTGA
 GGACATCTATTCTATGAGCCTGACGGTGTCTTAAGCTGATCAAGAGCTACAACATCGACACGAATA
 AGATCGGTCTGTGGAGGTGGGTACGGAGACGCTGATTGACAAGAGCAAAAGCGTGAAGTCTGTCTT

AATGCAGCTGTTTCGGCGAGAACACGGATGTCTGAGGGTATCGACACCCCTGAACGCGTGTACGGGCGG
 ACCAACGCACTGTTCAATAGCCTGAACTGGATTGAGAGCAACGCCTGGGATGGCCGCGATGCGATCGT
 CGTGTGCGGCGATATCGCCATCTATGACAAGGGTGCGGCACGTCCGACCGGCGGTGCAGGCACCGTTG
 CGATGTGGATTGGCCCGGACGCACCAATTGTCTTCGATTCTGTCCGCGCGTCTTACATGGAGCACGCCT
 ACGACTTTTACAAGCCGGACTTCACGAGCGAATACCCGTACGTGGACGGCCACTTCTCTCTGACCTGCT
 ATGTGAAGGCGCTGGACCAGGTTTATAAGTCTTATAGCAAAAAGGCGATTCTAAGGGCCTGGTCAGC
 GACCCGGCAGGCAGCGACGCCCTGAACGTGCTGAAGTATTTGACTACAACGTGTTCCATGTCCCGAC
 CTGCAAATTAGTGACCAAATCTTATGGCCGCCTGTTATATAATGATTTCCGTGCCAACCCGCAGCTGTT
 CCCGGAGGTTGACGCCGAGCTGGCGACGCGTGATTACGACGAGAGCCTGACCGACAAGAACATCGAG
 AAGACCTTCGTCAACGTGCGGAAGCCGTTCCACAAAAGAGCGTGTGGCCCAAAGCCTGATCGTCCCGAC
 CAACACGGGCAACATGTATACCGCGTCTGTCTACGCGGCATTTCGCGAGCCTGCTGAATTACGTCGGTT
 CTGACGACCTGCAGGGCAAGCGCGTTGGCCTGTTACGCTACGGTAGCGGCTTAGCGGCCAGCCTGTAT
 AGCTGCAAAATTGTGCGGCGACGTCCAGCACATCATCAAGGAGCTGGACATCACCAACAAGCTGGCGA
 AGCGCATCACCGAGACGCCGAAAGATTACGAGGCAGCGATCGAGTTACGCGAGAATGCGCATCTGAA
 GAAGAACTTCAAGCCGCAAGGTAGCATCGAGCACCTGCAGAGCGGCGTCTACTACCTGACGAACATT
 GACGACAAGTTCGCGCGTCTTATGACGTCAAAAAGTAACTAGTAGGAGGAAAACATCATGCAAAGTT
 TAGATAAGAATTTCCGACATTTATCTCGTCAACAAAAGTTACAACAATTGGTAGATAAGCAATGGTTA
 TCAGAAGATCAATTCGACATTTTATTGAATCATCCATTAATTGATGAGGAAGTAGCAAATAGTTTAATT
 GAAAATGTCATCGCGCAAGGTGCATTACCCGTTGGATTATTACCGAATATCATTTGTGGACGATAAGGC
 ATATGTTGTACCTATGATGGTGGAAGAGCCTTCAGTTGTCGCTGCAGCTAGTTATGGTGCAAAGCTAGT
 GAATCAGACTGGCGGATTTAAAAACGGTATCTTCTGAACGTATTATGATAGGTCAAATCGTCTTTGATG
 GCGTTGACGATACTGAAAAATTATCAGCAGACATTAAAGCTTTAGAAAAGCAAATTCATAAAATTGCG
 GATGAGGCATATCCTTCTATTAAAGCGCGTGGTGGTGGTTACCAACGTATAGCTATTGATACATTTCCCT
 GAGCAACAGTTACTATCTTTAAAGTATTTGTTGATACGAAAGATGCTATGGGCGCTAATATGCTTAAT
 ACGATTTTAGAGGCCATAACTGCATTTTAAAAAATGAATCTCCACAAAGCGACATTTTAATGAGTATT
 TTATCCAATCATGCAACAGCGTCCGTTGTTAAAGTTCAAGGCGAAATTGACGTTAAAGATTTAGCAAG
 GGGCGAGAGAACTGGAGAAGAGGTTGCCAAACGAATGGAACGTGCTTCTGTATTGGCACAAGTTGAT
 ATTCATCGTGCTGCAACACATAATAAAGGTGTTATGAATGGCATAACATGCCGTTGTTTTAGCAACAGG
 AAATGATACGCGTGGTGCAGAAGCAAGTGCGCATGCATACGCGAGTCGTGACGGACAGTATCGTGGT
 ATTGCAACATGGAGATACGATCAAAAACGTCAACGTTTAATTGGTACAATAGAAGTGCCTATGACATT
 GGCAATCGTTGGCGGTGGTACAAAAGTATTACCAATTGCTAAAGCTTCTTTAGAATTGCTAAATGTAG
 ATTCAGCACAAGAATTAGGTCATGTAGTTGCTGCCGTTGGTTTAGCACAGAAGTTTGACGCATGTGCG
 GCGCTCGTTTCCGAAGGTATCCAGCAAGGCCATATGAGCTTGCAATATAAATCTTTAGCTATTGTTGTA
 GGTGCAAAAGGTGATGAAATTGCGCAAGTAGCTGAAGCATTGAAGCAAGAACCCCGTGCGAATACAC
 AAGTAGCTGAACGCATTTTACAAGAAATTAGACAACAATAG

SEQ ID NO: 42

Operon *atoB*(opt):*mvaS*(opt):*mvaA*

ATGAAGAACTGTGTGATTGTTTCTGCGGTCCGCACGGCGATCGGCAGCTTTAACGGCTCTTTAGCGAG
 CACCTCTGCAATCGATCTGGGTGCGACGGTCATTAAGGCCGCCATTGAACGCGCCAAAATCGACAGCC
 AGCACGTTGATGAGGTGATCATGGGCAATGTGTTACAAGCCGGCCTGGGTCAAAACCCAGCGCGTCAA
 GCACTGTTAAATCTGGTCTGGCCGAGACCGTGTGTGGCTTCACCGTCAATAAGGTTTGCGGCTCTGGC

CTGAAGAGCGTGGCCCTGGCAGCACAAAGCGATTCAAGCCGGTCAGGCACAAAGCATCGTTGCGGGTIG
 GCATGGAGAACATGTCTCTGGCGCCGTACTTATTAGATGCCAAAGCCCGCAGCGGTTATCGCCTGGGC
 GATGGTCAGGTGTACGACGTCATCTTACGCGATGGCTTAATGTGCGCGACCCACGGTTACCACATGGG
 TATTACGGCCGAAAACGTGGCGAAAAGAATACGGCATTACGCGCGAGATGCAGGATGAATTAGCACTG
 CACTCTCAGCGCAAAGCAGCAGCCGCGATCGAGTCTGGTGCGTTTACGGCGGAAATCGTGCCAGTTAA
 CGTGGTCACGCGCAAGAAGACGTTTCGTTTTTCAGCCAGGACGAGTTCCCGAAGGCAAAACAGCACCGCG
 GAGGCCTTAGGTGCCTTACGCCAGCCTTTGACAAAGCGGGCACGGTCACCGCCGGTAATGCGAGCGG
 CATCAATGATGGTGCAGCGGCACTGGTCATCATGGAAGAGAGCGCCGCATTAGCAGCGGGTCTGACCC
 CATTAGCGCGCATTAAATCTTATGCCAGCGGCGGCGTCCCACCAGCCCTGATGGGCATGGGTCCGGTC
 CCAGCCACGCAAAAAGCCCTGCAATTAGCGGGCCTGCAACTGGCCGACATTGATCTGATCGAGGCGA
 ACGAGGCGTTTGCAGCGCAGTTCCTGGCGGTGGGTAAGAATCTGGGCTTCGACAGCGAGAAAAGTCAAT
 GTGAACGGTGGCGCGATTGCGTTAGGCCATCCGATTGGTGCAAGCGGCGCACGCATCTTAGTGACGTT
 ACTGCACGCCATGCAGGCACGCGACAAGACCTTAGGCCTGGCGACCTTATGTATTGGTGGCGGTCAAG
 GTATCGCCATGGTGATCGAACGCCTGAACTGAAGATCTAGGAGGAAAAGCAAAATGACAATAGGTATC
 GACAAAATAAACTTTTACGTTCCAAAGTACTATGTAGACATGGCTAAATTAGCAGAAGCACGCCAAGT
 AGACCCAAACAAATTTTAAATTGGAATTGGTCAAACCTGAAATGGCTGTTAGTCTGTAAACCAAGACA
 TCGTTTCAATGGGCGCTAACGCTGCTAAGGACATTATAACAGACGAAGATAAAAAGAAAATTGGTATG
 GTAATTGTGGCAACTGAATCAGCAGTTGATGCTGCTAAAGCAGCCGCTGTTCAAATTCACAACTTATT
 AGGTATTCAACCTTTTGCACGTTGCTTTGAAATGAAAGAAGCTTGTTATGCTGCAACACCAGCAATTCA
 ATTAGCTAAAGATTATTTAGCAACTAGACCGAATGAAAAAGTATTAGTTATTGCTACAGATACAGCAC
 GTTATGGATTGAATTCAGGCGGCGAGCCAACACAAGGTGCTGGCGCAGTTGCGATGGTTATTGCACAT
 AATCCAAGCATTTTGGCATTAAATGAAGATGCTGTTGCTTACACTGAAGACGTTTATGATTTCTGGCGT
 CCAACTGGACATAAATATCCATTAGTTGATGGTGCATTATCTAAAGATGCTTATATCCGCTCATTCCAA
 CAAAGCTGGAATGAATACGCAAAACGTCAAGGTAAGTCGCTAGCTGACTTCGCATCTCTATGCTTCCA
 TGTTCCATTTACAAAAATGGGTAAAAAGGCATTAGAGTCAATCATTGATAACGCTGATGAAACAACCTC
 AAGAGCGTTTACGTTTCAGGATATGAAGATGCTGTAGATTATAACCGTTATGTCGGTAATATTTATACTG
 GATCATTATATTTAAGCCTAATATCATTACTTGAAAATCGTGATTTACAAGCTGGTGAAACAATCGGTT
 TATTCAGTTATGGCTCAGGTTGAGTTGGTGAATTTTATAGTGCGACATTAGTTGAAGGCTACAAAGATC
 ATTTAGATCAAGCTGCACATAAAGCATTATTAATAAACCGTACTGAAGTATCTGTTGATGCATATGAA
 ACATTCTTCAAACGTTTTGATGACGTTGAATTTGACGAAGAACAAGATGCTGTTTATGAAGATCGTCAT
 ATTTTCTACTTATCAAATATTGAAAATAACGTTTCGCGAATATCACAGACCAGAGTAACTAGTAGGAGG
 AAAACATCATGCAAAGTTTATGATAAGAATTTCCGACATTTATCTCGTCAACAAAAGTTACAACAATTG
 GTAGATAAGCAATGGTTATCAGAAGATCAATTCGACATTTTATTGAATCATCCATTAATTGATGAGGA
 AGTAGCAAATAGTTTAATTGAAAATGTCATCGCGCAAGGTGCATTACCCGTTGGATTATTACCGAATA
 TCATTGTGGACGATAAAGGCATATGTTGTACCTATGATGGTGGAAGAGCCTTCAGTTGTCGCTGCAGCT
 AGTTATGGTGCAAAGCTAGTGAATCAGACTGGCGGATTTAAAACGGTATCTTCTGAACGTATTATGAT
 AGGTCAAATCGTCTTTGATGGCGTTGACGATACTGAAAAATTATCAGCAGACATTAAAGCTTTAGAAA
 AGCAAATTCATAAAATTGCGGATGAGGCATATCCTTCTATTAAAGCGCGTGGTGGTGGTTACCAACGT
 ATAGCTATTGATACATTTCTGAGCAACAGTTACTATCTTTAAAAGTATTTGTTGATACGAAAGATGCT
 ATGGGCGCTAATATGCTTAATACGATTTTATAGAGGCCATAACTGCATTTTTTAAAAAATGAATCTCCACA
 AAGCGACATTTTAAATGAGTATTTTATCCAATCATGCAACAGCGTCCGTTGTTAAAGTTCAAGGCGAAA
 TTGACGTTAAAGATTTAGCAAGGGGCGAGAGAACTGGAGAAGAGGTTGCCAAACGAATGGAACGTGC

[illegible]

SEQ ID NO: 43

pAM328 – ERG9-KANMX-gen khởi đầu MET3-ERG9 (không kể khung vector)

[illegible]

TGTGATCTCTAGCTTGGGTCTCTCTCTGTCGTAACAGTTGTGATATCGGCTGCTTCATCTCGACCGGAT
GCAATGCCAATTGTAATAGCTTTCCCATGTTAATTATACTTTATTCTT

SEQ ID NO: 44

GAGTGAACCTGCTGCCTGGCGTGCTCTGACTCAGTACATTTCATAGTGGATGGCGGCGTTAGTATC

SEQ ID NO: 45

CGTGTATACGTTTTCCGCTTCTGCTCTTCGTCTTTTCTCTTCTTCCGATATCACAACCTGTTACGA

SEQ ID NO: 46

GGTAAGACGGTTGGGTTTTATCTTTTGCAGTTGGTACTATTAAGAACAATCACAGGAAACAGCTATGA
CC

SEQ ID NO: 47

TTGCGTTTTGTACTTTGGTTCGCTCAATTTTGCAGGTAGATAATCGAAAAGTTGTAAAACGACGGCCAG
T

SEQ ID NO: 48

Trình tự pAM491 (không kể khung vector)

GTTTAAACTTGCTAAATTCGAGTGAAACACAGGAAGACCAGAAAATCCTCATTTTCATCCATATTAACA
ATAATTTCAAATGTTTATTTGCATTATTTGAACTAGGGAAGACAAGCAACGAAACGTTTTGAAAATTT
TGAGTATTTTCAATAAATTTGTAGAGGACTCAGATATTGAAAAAAGCTACAGCAATTAATACTTGAT
AAGAAGAGTATTGAGAAGGGCAACGGTTCATCATCTCATGGATCTGCACATGAACAAACACCAGAGT
CAAACGACGTTGAAATTGAGGCTACTGCGCCAATTGATGACAATACAGACGATGATAACAAACCGAA
GTTATCTGATGTAGAAAAGGATTAAAGATGCTAAGAGATAGTGATGATATTTTATAAATAATGTAATT
CTATATATGTTAATTACCTTTTTTTCGAGGCATATTTATGGTGAAGGATAAGTTTTGACCATCAAAGAA
GGTTAATGTGGCTGTGGTTTCAGGGTCCATACCCGGGAGTTATGACAATTACAACAACAGAATCTTTTC
TATATATGCACGAACCTTGTAATATGGAAGAAATTATGACGTACAACTATAAAGTAAATATTTTACGT
AACACATGGTGCTGTTGTGCTTCTTTTTCAAGAGAATACCAATGACGTATGACTAAGTTTAGGATTTAA
TGCAGGTGACGGACCCATCTTTCAAACGATTTATATCAGTGGCGTCCAAATTGTTAGGTTTTGTGGTT
CAGCAGGTTTCCTGTTGTGGGTCATATGACTTTGAACCAAATGGCCGGCTGCTAGGGCAGCACATAAG
GATAATTCACCTGCCAAGACGGCACAGGCAACTATTCTTGCTAATTGACGTGCGTTGGTACCAGGAGC
GGTAGCATGTGGGCCTCTTACACCTAATAAGTCCAACATGGCACCTTGTTGGTTCTAGAACAGTACCAC
CACCGATGGTACCTACTTCGATGGATGGCATGGATACGGAAATTCTCAAATCACCGTCCACTTCTTTCA
TCAATGTTATACAGTTGGAACCTTTTCGACATTTTGTGCAGGATCTTGTCTAATGCCAAGAAAAACAGCTG
TCACTAAATTAGCTGCATGTGCGTTAAATCCACCAACAGACCCAGCCATTGCAGATCCAACCAAATTC
TTAGCAATGTTCAACTCAACCAATGCGGAAACATCACTTTTTAACACTTTTCTGACAACATCACCAGGA
ATAGTAGCTTCTGCGACGACACTCTTACCACGACCTTCGATCCAGTTGATGGCAGCTGGTTTTTTGTGCG
GTACAGTAGTTACCAGAAACGGAGACAACCTCCATATCTTCCAGCCATACTCTTCTACCATTTGCTTT
AATGAGTATTCGACACCCTTAGAAATCATATTCATACCCATTGCGTCACCAGTAGTTGTTCTAAATCTC
ATGAAGAGTAAATCTCCTGCTAGACAAGTTTGAATATGTTGCAGACGTGCAAATCTTGATGTAGAGTT
AAAAGCTTTTTTAATTGCGTTTTGTCCCTCTTCTGAGTCTAACCATATCTTACAGGCACCAGATCTTTTC

AAAGTTGGGAAACGGACTACTGGGCTCTTGTGCATACCATCCTTAGTTAAAACAGTTGTTGCACCAUC
 GCCAGCATTGATTGCCTTACAGCCACGCATGGCAGAAGCTACCAAACAACCCTCTGTAGTTGCCATTG
 GTATATGATAAGATGTACCATCGATAACCAAGGGGCCTATAACACCAACGGGCAAAGGCATGTAACC
 TATAACATTTTACAACAAGCGCCAAATACGCGGTCTGTAGTCATAATTTTATATGGTAAACGATCAG
 ATGCTAATACAGGAGCTTCTGCCAAAATTGAAAGAGCCTTCTACGTACCGCAACCGCTCTCGTAGTA
 TCACCTAATTTTTTCTCCAAAGCGTACAAAGGTAACCTTACCGTGAATAACCAAGGCAGCGACCTCTTTG
 TTCTTCAATTGTTTTGTATTTCCACTACTTAATAATGCTTCTAATTCTTCTAAAGGACGTATTTTTCTTATC
 CAAGCTTTCAATATCGCGGGAATCATCTTCTCACTAGATGATGAAGGTCCTGATGAGCTCGATTGCGC
 AGATGATAAACTTTTGACTTTTCGATCCAGAAATGACTGTTTTATTGGTTAAAACTGGTGTAGAAGCCTT
 TTGTACAGGAGCAGTAAAAGACTTCTTGGTGACTTCAGTCTTACCAATTGGTCTGCAGCCATTATAGT
 TTTTCTCCTTGACGTTAAAGTATAGAGGTATATTAACAATTTTTTGTGATACTTTTATGACATTTGAA
 TAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGCTTTTGCATTTAT
 ATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCAGAAATAAGGCTAAAAAACT
 AATCGCATTATTATCCTATGGTTGTTAATTTGATTCTGTTGATTTGAAGGTTTGTGGGGCCAGGTTACTG
 CCAATTTTTCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTTCGGAGCAGTGCGGCG
 CGAGGCACATCTGCGTTTCAGGAACGCGACCGGTGAAGACCAGGACGCACGGAGGAGAGTCTTCCGT
 CGGAGGGCTGTGCCCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTA
 TTACTGAAAGTTCCAAAGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACAT
 ATAAGTAAGATTAGATATGGATATGTATATGGTGGTATTGCCATGTAATATGATTATTAAACTTCTTTG
 CGTCCATCCAAAAAAAAGTAAGAATTTTTGAAAATTCAATATAAATGAACTCTCAACTAAACTTTTG
 TTGGTGTGGTATTAAAGGAAGACTTAGGCCGCAAAAGCAACAACATTACACAATACAAACTTGCAA
 ATGACTGAACTAAAAAAACAAAAGACCGCTGAACAAAAAACAGACCTCAAATGTCTGGTATTAAAG
 GTATCCAAATTTACATCCCAACTCAATGTGTCAACCAATCTGAGCTAGAGAAATTTGATGGCGTTTCTC
 AAGGTAAATACACAATTGGTCTGGGCCAAACCAACATGTCTTTTGTCAATGACAGAGAAGATATCTAC
 TCGATGTCCCTAACTGTTTTGTCTAAGTTGATCAAGAGTTACAACATCGACACCAACAAAATTGGTAG
 ATTAGAAGTCGGTACTGAACTCTGATTGACAAGTCCAAGTCTGTCAAGTCTGTCTTGATGCAATTGTT
 TGGTGAACAACTGACGTCAAGGTATTGACACGCTTAATGCCTGTTACGGTGGTACCAACGCGTTGT
 TCAACTCTTTGAACTGGATTGAATCTAACGCATGGGATGGTAGAGACGCCATTGTAGTTTTCGGTGAT
 ATTGCCATCTACGATAAGGGTGCCGCAAGACCAACCGGTGGTGCCGGTACTGTTGCTATGTGGATCGG
 TCCTGATGCTCCAATTGTATTTGACTCTGTAAGAGCTTCTTACATGGAACACGCCTACGATTTTTACAA
 GCCAGATTTACCAGCGAATATCCTTACGTGATGGTCAATTTTTCATTAACCTGTTACGTCAAGGCTCT
 TGATCAAGTTTACAAGAGTTATTCCAAGAAGGCTATTTCTAAAGGGTTGGTTAGCGATCCCGCTGGTTC
 GGATGCTTTGAACGTTTTGAAATATTTGACTACAACGTTTTCCATGTTCCAACCTGTAAATTGGTCAC
 AAAATCATACGGTAGATTACTATATAACGATTTTACAGAGCCAATCCTCAATTGTTCCCAGAAGTTGACG
 CCGAATTAGCTACTCGCGATTATGACGAATCTTTAACCGATAAGAACATTGAAAAAACTTTTGTTAAT
 GTTGCTAAGCCATTCCACAAAGAGAGAGTTGCCCAATCTTTGATTGTTCCAACAAACACAGGTAACAT
 GTACACCGCATCTGTTTATGCCGCTTTGCATCTCTATTAACCTATGTTGGATCTGACGACTTACAAGG
 CAAGCGTGTTGGTTTATTTCTTACGGTTCCGGTTTAGCTGCATCTCTATATTCTTGCAAATTTGTTGGT
 GACGTCCAACATATTATCAAGGAATTAGATATTACTAACAAATTAGCCAAGAGAATCACCGAAACTCC
 AAAGGATTACGAAGCTGCCATCGAATTGAGAGAAAATGCCCATTTGAAGAAGAACTTCAAACCTCAA
 GGTTCCATTGAGCATTGCAAAGTGGTGTTTACTACTTGACCAACATCGATGACAAATTTAGAAGATCT
 TACGATGTTAAAAAATAATCTTCCCCCATCGATTGCATCTTGCTGAACCCCTTCATAAATGCTTTATTT

TTTTGGCAGCCTGCTTTTTTTAGCTCTCATTTAATAGAGTAGTTTTTTAATCTATATACTAGGAAAACTC
 TTTATTTAATAACAATGATATATATATACCCGGGAAGCTTTTCAATTCATCTTTTTTTTTTTTGTCTTTT
 TTTTGATTCCGGTTTTCTTTGAAATTTTTTTGATTCCGTAATCTCCGAGCAGAAGGAAGAACGAAGGAAG
 GAGCACAGACTTAGATTGGTATATATACGCATATGTGGTGTGAAGAAACATGAAATTGCCAGTATT
 CTTAACCCAACTGCACAGAACAAAAACCTGCAGGAAACGAAGATAAATCATGTCGAAAAGCTACATAT
 AAGGAACGTGCTGCTACTCATCCTAGTCCTGTTGCTGCCAAGCTATTTAATATCATGCACGAAAAGCA
 AACAAACTTGTGTGCTTCATTGGATGTTTCGTACCACCAAGGAATTACTGGAGTTAGTTGAAGCATTAG
 GTCCCAAAATTTGTTTACTAAAAACACATGTGGATATCTTGACTGATTTTTCCATGGAGGGCACAGTTA
 AGCCGCTAAAGGCATTATCCGCCAAGTACAATTTTTTACTCTTCGAAGACAGAAAATTTGCTGACATTG
 GTAATACAGTCAAATTGCAGTACTCTGCGGGTGTATACAGAATAGCAGAATGGGCAGACATTACGAAT
 GCACACGGTGTGGTGGGCCAGGTATTGTTAGCGGTTTGAAGCAGGCGGCGGAAGAAGTAACAAAGG
 AACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCATGCAAGGGCTCCCTAGCTACTGGAGAATATACT
 AAGGGTACTGTTGACATTGCGAAGAGCGACAAAGATTTTGTATCGGCTTTATTGCTCAAAGAGACAT
 GGGTGAAGAGATGAAGGTTACGATTGGTTGATTATGACACCCGGTGTGGGTTTAGATGACAAGGGA
 GACGCATTGGGTCAACAGTATAGAACCGTGGATGATGTGGTCTCTACAGGATCTGACATTATTATTGTT
 GGGTTTAAAC

SEQ ID NO: 49

Trình tự pAM492 (không kể khung vectơ)

GTTTAAACTTGCTAAATTCGAGTGAAACACAGGAAGACCAGAAAATCCTCATTTTCATCCATATTAACA
 ATAATTTCAAATGTTTATTTGCATTATTTGAAACTAGGGAAGACAAGCAACGAAACGTTTTTGAAAATT
 TTGAGTATTTTCAATAAATTTGTAGAGGACTCAGATATTGAAAAAAGCTACAGCAATTAATACTTGA
 TAAGAAGAGTATTGAGAAGGGCAACGGTTCATCATCTCATGGATCTGCACATGAACAAACACCAGAG
 TCAAACGACGTTGAAATTGAGGCTACTGCGCCAATTGATGACAATACAGACGATGATAACAAACCGA
 AGTTATCTGATGTAGAAAAGGATTAAAGATGCTAAGAGATAGTGATGATATTTTATAAATAATGTAAT
 TCTATATATGTTAATTACCTTTTTTTCGAGGCATATTTATGGTGAAGGATAAGTTTTGACCATCAAAGA
 AGGTAAATGTGGCTGTGGTTTCAGGGTCCATACCCGGGTATATATATATCATTGTTATTAATAAAGAG
 TTTTCCTAGTATATAGATTAAAAAACTACTCTATTAAATGAGAGCTAAAAAAGCAGGCTGCCAAAAA
 AATAAAGCATTATGAAGGGGGTTCAGCAAGATGCAATCGATGGGGGAAGATTATTTTTTAACATCGT
 AAGATCTTCTAAATTTGTCTCGATGTTGGTCAAGTAGTAAACACCACTTTGCAATGCTCAATGGAAC
 CTTGAGGTTTGAAGTTCTTCTTCAAATGGGCATTTTCTCTCAATTCGATGGCAGCTTCGTAATCCTTTGG
 AGTTTCGGTGATTCTCTTGGCTAATTTGTTAGTAATATCTAATTCCTTGATAATATGTTGGACGTCACCA
 ACAATTTTGCAAGAATATAGAGATGCAGCTAAACCGGAACCGTAAGAAAATAAACCAACACGCTTGC
 CTTGTAAGTCGTCAGATCCAACATAGTTTAATAGAGATGCAAAGGCGGCATAAACAGATGCGGTGTAC
 ATGTTACCTGTGTTTGTGGAACAATCAAAGATTGGGCAACTCTCTCTTTGTGGAATGGCTTAGCAACA
 TTAACAAAAGTTTTTTCAATGTTCTTATCGGTTAAAGATTCGTCATAATCGCGAGTAGCTAATTCGGCG
 TCAACTTCTGGGAACAATTGAGGATTGGCTCTGAAATCGTTATATAGTAATCTACCGTATGATTTTGTG
 ACCAATTTACAGGTTGGAACATGGAAAACGTTGTAGTCGAAATATTTCAAACGTTCAAAGCATCCGA
 ACCAGCGGGATCGCTAACCAACCCCTTTAGAAATAGCCTTCTTGAATAACTCTTGTAAACTTGATCAA
 GAGCCTTGACGTAACAAGTTAATGAAAAATGACCATCGACGTAAGGATATTCGCTGGTGAAATCTGGC
 TTGTAAAAATCGTAGGCGTGTTCCATGTAAGAAGCTCTTACAGAGTCAAATACAATTGGAGCATCAGG
 ACCGATCCACATAGCAACAGTACCGGCACCACCGGTTGGTCTTGCGGCACCCTTATCGTAGATGGCAA

TATCACCGCAAACACTACAATGGCGTCTCTACCATCCCATGCGTTAGATTCAATCCAGTTCAAAGAGTTGA
 ACAACGCGTTGGTACCACCGTAACAGGCATTAAGCGTGTCAATACCTTCGACGTCAGTGTTTTACCA
 AACAATTGCATCAAGACAGACTTGACAGACTTGGACTTGTCAATCAGAGTTTCAGTACCGACTTCTAA
 TCTACCAATTTTGTGGTGTGATGTTGTAACCTTGTGATCAACTTAGACAAAACAGTTAGGGACATCGA
 GTAGATATCTTCTCTGTCATTGACAAAAGACATGTTGGTTTGGCCCAGACCAATTGTGTATTTACCTTG
 AGAAACGCCATCAAATTTCTCTAGCTCAGATTGGTTGACACATTGAGTTGGGATGTAAATTTGGATAC
 CTTTAATACCGACATTTTGGAGTCTGGTTTTTTTGTTCAGCGGTCTTTTGTTTTTTTAGTTTCAGTCATTTGC
 AAGTTTGTATTGTGTAATTGTTGTTGCTTTTGC GGCCTAAGTCTTCCTTTAATACCACACCAACAAAGTT
 TAGTTGAGAGTTTCATTTTATGTGATGATTGATTGATTGATTGTACAGTTTGTTTTTCTTAATATCTATT
 TCGATGACTTCTATATGATATTGCACTAACAAGAAGATATTATAATGCAATTGATACAAGACAAGGAG
 TTATTTGCTTCTCTTTTATATGATTCTGACAATCCATATTGCGTTGGTAGTCTTTTTTGTCTGGAACGGTT
 CAGCGGAAAAGACGCATCGCTCTTTTTGCTTCTAGAAGAAATGCCAGCAAAAAGAACTCTTTGACAGTG
 ACTGACAGCAAAAATGTCTTTTTCTAACTAGTAACAAGGCTAAGATATCAGCCTGAAATAAAGGGTGG
 TGAAGTAATAATTAAATCATCCGTATAAACCTATACACATATATGAGGAAAAATAATACAAAAGTGTT
 TTAAATACAGATACATACATGAACATATGCACGTATAGCGCCCAAATGTCGGTAATGGGATCGGCCTA
 CTAATTATAAAATGCATCATAGAAATCGTTGAAGTTGACGCAGCGACTCGAGATCCATAGGAGCAACT
 CATGTCTGAACTTCAACGATTTCTATGATGCATTTTATAATTAGTAAGCCGATCCCATTACCGACATTT
 GGGCGCTATACGTGCATATGTTTCATGTATGTATCTGTATTTAAACACTTTTGTATTATTTTTCCTCATA
 TATGTGTATAGGTTTATACGGATGATTTAATTATTACTTCACCACCCTTTATTTTCAGGCTGATATCTTAG
 CCTTGTTACTAGTTAGAAAAAGACATTTTGTCTGTCAGTCACTGTCAAGAGATTCTTTTGTCTGGCATT
 CTTCTAGAAGCAAAAAGAGCGATGCGTCTTTTCCGCTGAACCGTTCCAGCAAAAAGACTACCAACGC
 AATATGGATTGTCAGAATCATATAAAAGAGAAGCAAATAACTCCTTGTCTTGTATCAATTGCATTATA
 ATATCTTCTTGTAGTGCAATATCATATAGAAGTCATCGAAATAGATATTAAGAAAAACAACTGTAC
 AATCAATCAATCAATCATACATAAAATGGCTGCAGACCAATTGGTGAAGACTGAAGTCACCAAGAA
 GTCTTTTACTGCTCCTGTACAAAAGGCTTCTACACCAGTTTTTAACCAATAAAACAGTCATTTCTGGATC
 GAAAGTCAAAAAGTTTATCATCTGCGCAATCGAGCTCATCAGGACCTTCATCATCTAGTGAGGAAGATG
 ATCCCGCGATATTGAAAGCTTGGATAAGAAAATACGTCCTTTAGAAGAATTAGAAGCATTATTAAGT
 AGTGGAATAACAAAACAATTGAAGAACAAGAGGTCGCTGCCTTGGTTATTCACGGTAAGTTACCTTT
 GTACGCTTTGGAGAAAAAATTAGGTGATACTACGAGAGCGGTTGCGGTACGTAGGAAGGCTCTTTCAA
 TTTTGGCAGAAGCTCCTGTATTAGCATCTGATCGTTTACCATATAAAAATTATGACTACGACCGCGTAT
 TTGGCGCTTGTGTGAAAATGTTATAGGTTACATGCCTTTGCCCGTTGGTGTTATAGGCCCTTGGTTAT
 CGATGGTACATCTTATCATATACCAATGGCAACTACAGAGGGTGTGTTTGGTAGCTTCTGCCATGCGTGG
 CTGTAAGGCAATCAATGCTGGCGGTGGTGCAACAACCTGTTTTAACTAAGGATGGTATGACAAGAGGCC
 CAGTAGTCCGTTTCCCAACTTTGAAAAGATCTGGTGCCTGTAAGATATGGTTAGACTCAGAAGAGGGA
 CAAAACGCAATTAAAAAAGCTTTTAACTCTACATCAAGATTTGCACGTCTGCAACATATTCAAACTTGT
 CTAGCAGGAGATTTACTCTTCATGAGATTTAGAACAACACTACTGGTGACGCAATGGGTATGAATATGAT
 TTCTAAGGGTGTGCAATACTCATTAAAGCAAATGGTAGAAGAGTATGGCTGGGAAGATATGGAGGTTG
 TCTCCGTTTCTGGTAACTACTGTACCGACAAAAAACAGCTGCCATCAACTGGATCGAAGGTCGTGGT
 AAGAGTGTGTCGTCAGAAAGCTACTATTCCTGGTGATGTTGTCAGAAAAGTGTTAAAAAGTGATGTTTC
 CGCATTGGTTGAGTTGAACATTGCTAAGAATTTGGTTGGATCTGCAATGGCTGGGTCTGTTGGTGGATT
 TAACGCACATGCAGCTAATTTAGTGACAGCTGTTTTCTTGGCATTAGGACAAGATCCTGCACAAAATG
 TCGAAAGTTCCAACGTATACATTGATGAAAGAAGTGGACGGTGATTTGAGAATTTCCGTATCCATG

CCATCCATCGAAGTAGGTACCATCGGTGGTGGTACTGTTCTAGAACACAAAGGTGCCATGTTGGACTT
 ATTAGGTGTAAGAGGCCACATGCTACCGCTCCTGGTACCAACGCACGTCAATTAGCAAGAATAGTTG
 CCTGTGCCGTCTTGGCAGGTGAATTATCCTTATGTGCTGCCCTAGCAGCCGGCCATTTGGTTCAAAGTC
 ATATGACCCACAACAGGAAACCTGCTGAACCAACAAAACCTAACAATTTGGACGCCACTGATATAAAT
 CGTTTGAAAGATGGGTCCGTCACCTGCATTAAATCCTAAACTTAGTCATACGTCATTGGTATTCTCTTG
 AAAAGAAGCACAAACAGCACCATGTGTTACGTAAAATATTTACTTTATAGTTTGTACGTCATAATTTCT
 TCCATATTACAAGTTCGTGCATATATAGAAAGAATTCTGTTGTTGTAATTGTCATAACTCCCGGAAGC
 TTTTCAATTCATCTTTTTTTTTTTTGTCTTTTTTTTGATTCCGGTTTCTTTGAAATTTTTTTGATTCCGGTA
 ATCTCCGAGCAGAAGGAAGAACGAAGGAAGGAGCACAGACTTAGATTGGTATATATACGCATATGTG
 GTGTTGAAGAAACATGAAATTGCCAGTATTCTTAACCCAACTGCACAGAACAAAAACCTGCAGGAA
 ACGAAGATAAATCATGTCGAAAGCTACATATAAGGAACGTGCTGCTACTCATCCTAGTCCTGTTGCTG
 CCAAGCTATTTAATATCATGCACGAAAAGCAAAACAACTTGTGTGCTTCATTGGATGTTTCGTACCACC
 AAGGAATTACTGGAGTTAGTTGAAGCATTAGGTCCCAAAATTTGTTTACTAAAAACACATGTGGATAT
 CTTGACTGATTTTCCATGGAGGGCACAGTTAAGCCGCTAAAGGCATTATCCGCCAAGTACAATTTTTT
 ACTCTTCGAAGACAGAAAATTTGCTGACATTGGTAATACAGTCAAATTGCAGTACTCTGCGGGTGTAT
 ACAGAATAGCAGAATGGGCAGACATTACGAATGCACACGGTGTGGTGGGCCAGGTATTGTTAGCGG
 TTTGAAGCAGGCGGCGGAAGAAGTAACAAAGGAACCTAGAGGCCTTTTGATGTTAGCAGAATTGTCAT
 GCAAGGGCTCCCTAGCTACTGGAGAATATACTAAGGGTACTGTTGACATTGCCAAGAGCGACAAAGA
 TTTTGTATCGGCTTTATTGCTCAAAGAGACATGGGTGGAAGAGATGAAGGTTACGATTGGTTGATTAT
 GACACCCGGTGTGGGTTTAGATGACAAGGGAGACGCATTGGGTCAACAGTATAGAACCGTGGATGAT
 GTGGTCTCTACAGGATCTGACATTATTATTGTTGGGTTTAAAC

SEQ ID NO: 50

Trình tự pAM489 (không kể khung vectơ)

GTTTAAACTACTATTAGCTGAATTGCCACTGCTATCGTTGTTAGTGGCGTTAGTGCTTGCATTCAAAGA
 CATGGAGGGCGTTATTACGCCGGAGCTCCTCGACAGCAGATCTGATGACTGGTCAATATATTTTTGCAT
 TGAGGCTCTGTTTGAATTATATTTTGAGATGACCCATCTAATGTACTGGTATCACCAGATTTTCATGTC
 GTTTTTTAAAGCGGCTGCTTGAGTCTTAGCAATAGCGTCACCATCTGGTGAATCCTTTGAAGGAACAC
 TGACGAAGGTTTGGACAGTGACGAAGAGGATCTTTCTGCTTTGAATTAGTCGCGCTGGGAGCAGATG
 ACGAGTTGGTGGAGCTGGGGGCAGGATTGCTGGCCGTCGTGGGTCTGAATGGGTCCTTGGCTGGTCC
 ATCTCTATTCTGAAAACGGAAGAGGAGTAGGGAATATTACTGGCTGAAAATAAGTCTTGAATGAACGT
 ATACGCGTATATTTCTACCAATCTCTCAACACTGAGTAATGGTAGTTATAAGAAAGAGACCGAGTTAG
 GGACAGTTAGAGGCGGTGGAGATATTCCTTATGGCATGTCTGGCGATGATAAACTTTTCAAACGGCA
 GCCCCGATCTAAAAGAGCTGACACCCGGGAGTTATGACAATTACAACAACAGAATTCCTTTCTATATAT
 GCACGAACTTGTAATATGGAAGAAATTATGACGTACAACTATAAAGTAAATATTTTACGTAACACAT
 GGTGCTGTTGTGCTTCTTTTTCAAGAGAATACCAATGACGTATGACTAAGTTTAGGATTTAATGCAGGT
 GACGGACCCATCTTTCAAACGATTTATATCAGTGGCGTCCAAATTGTTAGGTTTTGTTGGTTTACGAGG
 TTTCTGTTGTGGGTCATATGACTTTGAACCAAATGGCCGGCTGCTAGGGCAGCACATAAGGATAATT
 CACCTGCCAAGACGGCACAGGCAACTATTCTTGCTAATTGACGTGCGTTGGTACCAGGAGCGGTAGCA
 TGTGGGCCTCTTACACCTAATAAGTCCAACATGGCACCTTGTGGTTCTAGAACAGTACCACCACCGAT
 GGTACCTACTTCGATGGATGGCATGGATACGGAAATCTCAAATCACCGTCCACTTCTTTCATCAATGT
 TATACAGTTGGAACCTTCGACATTTTGTGCAGGATCTTGTCTTAATGCCAAGAAAACAGCTGTCACTAA

ATTAGCTGCATGTGCGTTAAATCCACCAACAGACCCAGCCATTGCAGATCCAACCAAATTCCTAGCAA
 TGTTCAACTCAACCAATGCGGAAACATCACTTTTTAACACTTTTCTGACAACATCACCAGGAATAGTAG
 CTTCTGCGACGACACTCTTACCACGACCTTCGATCCAGTTGATGGCAGCTGGTTTTTTGTCGGTACAGT
 AGTTACCAGAAAACGGAGACAACCTCCATATCTTCCAGCCATACTCTTCTACCATTTGCTTTAATGAGT
 ATTCGACACCCTTAGAAAATCATATTCATACCCATTGCGTCACCAGTAGTTGTTCTAAATCTCATGAAGA
 GTAAATCTCCTGCTAGACAAGTTTGAATATGTTGCAGACGTGCAAATCTTGATGTAGAGTTAAAAGCT
 TTTTAAATTGCGTTTTGTCCTCTTCTGAGTCTAACCATATCTTACAGGCACCAGATCTTTTCAAAGTTG
 GGAAACGGACTACTGGGCCTCTTGTCTATACCATCCTTAGTTAAAACAGTTGTTGCACCACCGCCAGCA
 TTGATTGCCTTACAGCCACGCATGGCAGAAGCTACCAAACAACCCTCTGTAGTTGCCATTGGTATATG
 ATAAGATGTACCATCGATAACCAAGGGGCTATAACACCAACGGGCAAAGGCATGTAACCTATAACA
 TTTTCACAACAAGCGCCAAATACGCGGTCGTAGTCATAATTTTTATATGGTAAACGATCAGATGCTAAT
 ACAGGAGCTTCTGCCAAAATTGAAAGAGCCTTCCTACGTACCGCAACCGCTCTCGTAGTATCACCTAA
 TTTTTCTCCAAAGCGTACAAAGGTAACCTACCGTGAATAACCAAGGCAGCGACCTCTTTGTTCTTCAA
 TTGTTTTGTATTTCCACTACTTAATAATGCTTCTAATCTTCTAAAGGACGTATTTTCTTATCCAAGCTTT
 CAATATCGCGGGAATCATCTTCTCTACTAGATGATGAAGGTCCTGATGAGCTCGATTGCGCAGATGAT
 AAACCTTTTGACTTTTCGATCCAGAAATGACTGTTTTATTGGTTAAAACCTGGTGTAGAAGCCTTTTGTACA
 GGAGCAGTAAAAGACTTCTTGGTGACTTCAGTCTTCACCAATTGGTCTGCAGCCATTATAGTTTTTCT
 CCTTGACGTAAAGTATAGAGGTATATTAACAATTTTTGTTGATACTTTTATGACATTTGAATAAGAA
 GTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGCTTTTGCATTTATATATCTG
 TTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCAGAAATAAGGCTAAAAAACTAATCGCA
 TTATTATCCTATGGTTGTTAATTTGATTGCTTGATTTGAAGGTTTGTGGGGCCAGGTTACTGCCAATTTT
 TCCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTTCGGAGCAGTGCGGCGCGAGGCAC
 ATCTGCGTTTCAGGAACGCGACCGGTGAAGACCAGGACGCACGGAGGAGAGTCTTCCGTCGGAGGGC
 TGTCGCCCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTATTACTGAAA
 GTTCCAAAGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAACATATAAGTAA
 GATTAGATATGGATATGTATATGGTGGTATTGCCATGTAATATGATTATTAAACTTCTTTGCGTCCATC
 CAAAAAAAAGTAAGAATTTTTGAAAATTCAATATAAATGGCTTCAGAAAAAGAAATTAGGAGAGAG
 AGATTCTTGAACGTTTTCCCTAAATTAGTAGAGGAATTGAACGCATCGCTTTTGGCTTACGGTATGCCT
 AAGGAAGCATGTGACTGGTATGCCCACTCATTGAACTACAACACTCCAGGCGGTAAGCTAAATAGAG
 GTTTGTCCGTTGTGGACACGTATGCTATTCTCTCCAACAAGACCGTTGAACAATTGGGGCAAGAAGAA
 TACGAAAAGGTTGCCATTCTAGGTTGGTGCATTGAGTTGTTGCAGGCTTACTTCTTGGTCGCCGATGAT
 ATGATGGACAAGTCCATTACCAGAAGAGGCCAACCATGTTGGTACAAGGTTCTGAAAGTTGGGGAAAT
 TGCCATCAATGACGCATTATGTTAGAGGCTGCTATCTACAAGCTTTTGAAATCTCACTTCAGAAACGA
 AAAATACTACATAGATATCACCGAATTGTTCCATGAGGTCACCTTCCAACCGAATTGGGCCAATTGA
 TGGACTTAATCACTGCACCTGAAGACAAAGTCGACTTGAGTAAGTTCTCCCTAAAGAAGCACTCCTTC
 ATAGTTACTTTCAAGACTGCTTACTATTCTTCTACTTGCCTGTCGCATTGGCCATGTACGTTGCCGGTA
 TCACGGATGAAAAGGATTTGAAACAAGCCAGAGATGTCTTGATTCCATTGGGTGAATACTTCCAAATT
 CAAGATGACTACTTAGACTGCTTCGGTACCCAGAACAGATCGGTAAAGATCGGTACAGATATCCAAGA
 TAACAAATGTTCTTGGGTAATCAACAAGGCATTGGAACTTGCTTCCGCAGAACAAAGAAAGACTTTAG
 ACGAAAATTACGGTAAGAAGGACTCAGTCGCAGAAGCCAAATGCAAAAAGATTTTCAATGACTTGAA
 AATTGAACAGCTATACCACGAATATGAAGAGTCTATTGCCAAGGATTTGAAGGCCAAAATTTCTCAGG
 TCGATGAGTCTCGTGGCTTCAAAGCTGATGTCTTAACTGCGTTCTTGAACAAAGTTTACAAGAGAAGC

AAATAGAACTAACGCTAATCGATAAAAACATTAGATTTCAAACCTAGATAAGGACCATGTATAAGAACTA
TATACTTCCAATATAATATAGTATAAGCTTTAAGATAGTATCTCTCGATCTACCGTTCCACGTGACTAG
TCCAAGGATTTTTTTTAACCCGGGATATATGTGTACTTTGCAGTTATGACGCCAGATGGCAGTAGTGGA
AGATATTCTTTATTGAAAAATAGCTTGTACCTTACGTACAATCTTGATCCGGAGCTTTTCTTTTTTGC
CGATTAAGAATTCCGGTCGAAAAAGAAAAGGAGAGGGCCAAGAGGGAGGGCATTGGTGACTATTGAG
CACGTGAGTATACGTGATTAAAGCACACAAAGGCAGCTTGGAGTATGTCTGTTATTAATTTACAGGTA
GTTCTGGTCCATTGGTGAAAGTTTGC GGCTTGCAGAGCACAGAGGCCGCGAGAATGTGCTCTAGATTCC
GATGCTGACTTGCTGGGTATTATATGTGTGCCCAATAGAAAGAGAACAATTGACCCGGTTATTGCAAG
GAAAATTTCAAGTCTTGTAAGCATATAAAAATAGTTTCAGGCACTCCGAAATACTTGGTTGGCGTGT
TTCGTAATCAACCTAAGGAGGATGTTTTGGCTCTGGTCAATGATTACGGCATTGATATCGTCCAACCTGC
ATGGAGATGAGTCGTGGCAAGAATACCAAGAGTTTCTCGGTTTGCCAGTTATTAAGAACTCGTATTT
CCAAAAGACTGCAACATACTACTCAGTGACGCTTCACAGAAACCTCATTTCGTTTATCCCTTGTTGAT
TCAGAAGCAGGTGGGACAGGTGAACTTTTGGATTGGAACCTCGATTTCTGACTGGGTGGAAGGCAAGA
GAGCCCCGAAAGCTTACATTTTATGTTAGCTGGTGGACTGACGCCGTTTAAAC

SEQ ID NO: 51

Trình tự pAM497 (không kể khung vecto)

GTTTAAACTTTTCCAATAGGTGGTTAGCAATCGTCTTACTTTCTAACTTTTCTTACCTTTTACATTTTCAG
CAATATATATATATATATTTCAAGGATATACCATTCTAATGTCTGCCCCTAAGAAGATCGTCGTTTTGC
CAGGTGACCACGTTGGTCAAGAAATCACAGCCGAAGCCATTAAGGTTCTTAAAGCTATTTCTGATGTT
CGTTCCAATGTCAAGTTTCGATTTGAAAAATCATTTAATTGGTGGTGCTGCTATCGATGCTACAGGTGTT
CCACTTCCAGATGAGGCGCTGGAAGCCTCCAAGAAGGCTGATGCCGTTTTGTTAGGTGCTGTGGGTGG
TCCTAAATGGGGTACCGGTAGTGTTAGACCTGAACAAGGTTTACTAAAAATCCGTAAAGAACTTCAAT
TGTAACGCCAACTTAAGACCATGTAACCTTTGCATCCGACTCTCTTTTAGACTTATCTCCAATCAAGCCAC
AATTTGCTAAAGGTACTGACTTCGTTGTTGTCAGAGAATTAGTGGGAGGTATTTACTTTGGTAAGAGA
AAGGAAGACGTTTAGCTTGCCTCGTCCCCGCCGGGTACCCGCCAGCGACATGGAGGCCGAGAATAC
CCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTGCCCCGTACATTTAGCCCAT
ACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACG
GCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCC
GCGCCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTTCATATACTTCCTTTT
AAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATGGCAGAACCAGCCCCAA
AAAAAGCAAAAACAACTGTTTCAGGAGCGCAAGGCGTTTATCTCCCGTATCACTAATGAAACTAAAAT
TCAAATCGCTATTTGCTGAATGGTGGTTATATTCAAATAAAAGATTTCGATTCTTCCTGCAAAGAAGGA
TGACGATGTAGCTTCCCAAGCTACTCAGTCACAGGTCATCGATATTCACACAGGTGTTGGCTTTTTGGA
TCATATGATCCATGCGTTGGCAAAACACTCTGGTTGGTCTCTTATTGTTGAATGTATTGGTGACCTGCA
CATTGACGATCACCATACTACCGAAGATTGCGGTATCGCATTAGGGCAAGCGTTCAAAGAAGCAATGG
GTGCTGTCCGTGGTGTAAGAAAGATTTCGGTACTGGGTTCGCACCATTTGGATGAGGCGCTATCACGTGCC
GTAGTCGATTTATCTAGTAGACATTTGCTGTAATCGACCTTGGATTGAAGAGAGAGATGATTGGTGA
TTTATCCACTGAAATGATTCCACACTTTTTGGAAAGTTTCGCGGAGGCGGCCAGAATTACTTTGCATGT
TGATTGTCTGAGAGGTTTCAACGATCACCACAGAAGTGAGAGTGCGTTCAAGGCTTTGGCTGTTGCCA
TAAGAGAAGCTATTTCTAGCAATGGCACCAATGACGTTCCTCAACCAAGGTGTTTTGATGTGAAGT
ACTGACAATAAAAAGATTCTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTT

TATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCAGATGCGAAGTT
AAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATT
GATACTAACGCCGCCATCCACCCGGGTTTCTCATTCAAGTGGTAACTGCTGTTAAAATTAAGATATTTA
TAAATTGAAGCTTGGTCGTTCCGACCAATACCGTAGGGAAACGTAAATTAGCTATTGTAAAAAAGGA
AAAGAAAAGAAAAGAAAAATGTTACATATCGAATTGATCTTATTCCTTTGGTAGACCAGTCTTTGCGT
CAATCAAAGATTTCGTTTGTTCCTGTGGGCTGAACCGACTTGAGTTAAAATCACTCTGGCAACATCCT
TTTGCAACTCAAGATCCAATTCACGTGCAGTAAAGTTAGATGATTCAAATTGATGGTTGAAAGCCTCA
AGCTGCTCAGTAGTAAATTTCTTGTCCCATCCAGGAACAGAGCCAAACAATTTATAGATAAATGCAAA
GAGTTTCGACTCATTTTCAGCTAAGTAGTACAACACAGCATTGGACCTGCATCAAACGTGTATGCAA
CGATTGTTTCTCCGTAAAACGTATTAATGGTGTGGCACCAACTGATGATACGCTTGGAAGTGTCAATTCA
TGTAGAATATTGGAGGGAAAGAGTCCAAACATGTGGCATGGAAAGAGTTGGAATCCATCATTGTTTCC
TTTGCAAAGGTGGCGAAATCTTTTCAACAATGGCTTTACGCATGACTTCAAATCTCTTTGGTACGACA
TGTTCAATTCTTTCTTTAAATAGTTTCGGAGGTTGCCACGGTCAATTGCATACCCTGAGTGGAACCTACA
TCCTTTTTAATATCGCTGACAACCTAGGACACAAGCTTTTCATCTGAGGCCAGTCAGAGCTGTCTGCGATT
TGTAATGCCATGGAATCATGACCATCTTCAGCTTTTCCCATTTCCCAGGCCACGTATCCGCCAAACAAC
GATCTACAAGCTGAACCAGACCCCTTTCTTGCTATTCTAGATATTTCTGAAGTTGACTGTGGTAATTGG
TATAACTTAGCAATTGCAGAGACCAATGCAGCAAAGCCAGCAGCGGAGGAAGCTAAACCAGCTGCTG
TAGGAAAGTTATTTTCGGAGACAATGTGGAGTTTCCATTGAGATAATGTGGGCAATGAGGCGTCCTTC
GATTCCATTTCTTTCTTAATTGGCGTAGGTGCGCGAGACAATTTTGAGTTCTTTCATTGTGCGATGCTGT
GTGGTTCTCCATTAAACCACAAAGTGTGCGGTTCAAACCTCAGGTGCAGTAGCCGCAGAGGTCAACGTT
CTGAGGTCATCTTGCGATAAAGTCACTGATATGGACGAATTGGTGGGCAGATTCAACTTCGTGTCCCTT
TTCCCCCAATACTTAAGGGTTGCGATGTTGACGGGTGCGGTAACGGATGCTGTGTAAACGGTCATTAT
AGTTTTTTCTCCTTGACGTTAAAGTATAGAGGTATATTAACAATTTTTTGTGATACTTTTATGACATTT
GAATAAGAAGTAATACAAACCGAAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGCTTTTGCATT
TATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCAGAAATAAGGCTAAAAA
ACTAATCGCATTATTATCCTATGGTTGTTAATTTGATTTCGTTGATTGAAGGTTTGTGGGGCCAGGTTA
CTGCCAATTTTTCCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTCCGAGCAGTGCG
GCGCGAGGCACATCTGCGTTTCAGGAACGCGACCGGTGAAGACCAGGACGCACGGAGGAGAGTCTTC
CGTCGGAGGGGCTGTGCCCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGC
GTATTACTGAAAGTTCCAAAGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTTACATTTCCACAA
CATATAAGTAAGATTAGATATGGATATGTATATGGTGGTATTGCCATGTAATATGATTATTAACTTCT
TTGCGTCCATCCAAAAAAGTAAGAATTTTTGAAAATTCAATATAAATGTCAGAGTTGAGAGCCTT
CAGTGCCCCAGGGAAAGCGTTACTAGCTGGTGGATATTTAGTTTTAGATCCGAAATATGAAGCATTG
TAGTCGGATTATCGGCAAGAATGCATGCTGTAGCCCATCCTTACGGTTCATTGCAAGAGTCTGATAAG
TTTGAAGTGCGTGTGAAAAGTAACAATTTAAAGATGGGGAGTGGCTGTACCATATAAGTCCTAAAAC
TGGCTTCATTCCTGTTTCGATAGGCGGATCTAAGAACCCTTTCATTGAAAAAGTTATCGCTAACGTATT
TAGCTACTTTAAGCCTAACATGGACGACTACTGCAATAGAACTTGTTTCGTTATTGATATTTTCTCTGA
TGATGCCCTACCATTCTCAGGAGGACAGCGTTACCGAACATCGTGGCAACAGAAGATTGAGTTTTTCAAT
CGCACAGAATTGAAGAAGTTCCCAAAACAGGGCTGGGCTCCTCGGCAGGTTTAGTCACAGTTTTAACT
ACAGCTTTGGCCTCCTTTTTTGTATCGGACCTGGAAAAATAATGTAGACAAATATAGAGAAGTTATTCAT
AATTTATCACAAGTTGCTCATTGTCAAGCTCAGGGTAAAATTGGAAGCGGGTTTGATGTAGCGGCGGC
AGCATATGGATCTATCAGATATAGAAGATTCCCACCCGCATTAATCTCTAATTGCCAGATATTGGAA

GTGCTACTTACGGCAGTAAACTGGGCGCATTTGGTTAATGAAGAAGACTGGAATATAACGATTAAAAAGT
 AACCATTTACCTTCGGGATTAACCTTTATGGATGGGCGATATTAAGAATGGTTCAGAAACAGTAAAACT
 GGTCCAGAAGGTAAAAAATTGGTATGATTTCGCATATGCCGGAAGCTTGAAAATATATACAGAACTCG
 ATCATGCAAATTCTAGATTTATGGATGGACTATCTAACTAGATCGCTTACACGAGACTCATGACGATT
 ACAGCGATCAGATATTTGAGTCTCTTGAGAGGAATGACTGTACCTGTCAAAAGTATCCTGAGATCACA
 GAAGTTAGAGATGCAGTTGCCACAATTAGACGTTCCCTTAGAAAAATAACTAAAGAATCTGGTGCCGA
 TATCGAACCTCCCGTACAACTAGCTTATTGGATGATTGCCAGACCTTAAAAGGAGTTCCTACTTGCTT
 AATACCTGGTGCTGGTGGTTATGACGCCATTGCAGTGATTGCTAAGCAAGATGTTGATCTTAGGGCTC
 AAACCGCTGATGACAAAAGATTTTCTAAGGTTCAATGGCTGGATGTAACCTCAGGCTGACTGGGGTGT
 AGGAAAGAAAAAGATCCGGAACCTTATCTTGATAAATAACTTAAGGTAGATAATAGTGGTCCATGTG
 ACATCTTTATAAATGTGAAGTTTGAAGTGACCGCGCTTAACATCTAACCATTCTCTTCCGATAGTACT
 TGAAATTGTTCCCTTCGGCGGCATGATAAAATTCTTTAATGGGTACAAGCTACCCGGGAAAGATTCTC
 TTTTTTTATGATATTTGTACATAAACTTTATAAATGAAATTCATAATAGAAACGACACGAAATTACAAA
 ATGGAATATGTTTCATAGGGTAGACGAACTATATACGCAATCTACATACATTTATCAAGAAGGAGAAA
 AAGGAGGATGTAAAGGAATACAGGTAAGCAAATTGATACTAATGGCTCAACGTGATAAGGAAAAAGA
 ATTGCACTTTAACATTAATATTGACAAGGAGGAGGGCACCACACAAAAAGTTAGGTGTAACAGAAAA
 TCATGAAACTATGATTCCTAATTTATATATTGGAGGATTTTCTCTAAAAAATAACAACAAAT
 AAAAAACACTCAATGACCTGACCATTTGATGGAGTTTAAAGTCAATACCTTCTTGAACCATTTCCCATAA
 TGGTGAAAGTTCCTCAAGAATTTTACTCTGTGCAAAACGGCCTTAACGACGTAGTCGACCTCCTCTTC
 AGTACTAAATCTACCAATACCAAATCTGATGGAAGAATGGGCTAATGCATCATCCTTACCCAGCGCAT
 GTAAACATAAGAAGGTTCTAGGGAAGCAGATGTACAGGCTGAACCCGAGGATAATGCGATATCCCT
 TAGTGCCATCAATAAAGATTCTCCTTCCACGTAGGCGAAAGAAACGTTAACACGTTTAAAC

SEQ ID NO: 52

Trình tự pAM493 (không kể khung vectơ)

GTTTAACTACTCAGTATATTAAGTTTCGAATTGAAGGGCGAACTCTTATTCGAAGTCGGAGTCACCAC
 AACACTTCCGCCATACTCTCCGAATCCTCGTTTCCTAAAGTAAGTTTACTTCCACTTGTAGGCCTATTA
 TTAATGATATCTGAATAATCCTCTATTAGGGTTGGATCATTAGTAGCGCGTGCGATTGAAAGGAGTCC
 ATGCCCCGACGTGACGTGATTAGCGAAGGCGCGTAACCATTTGTCATGTCTAGCAGCTATAGAATAAC
 CTCCTTGACACCACTTGCGGAAGTCTCATCAACATGCTCTTCCTTATTACTCATTCTCTTACCAAGCAG
 AGAATGTTATCTAAAAACTACGTGTATTTACCTCTTTCTCGACTTGAACACGTCCAACCTCCTTAAGTA
 CTACCACAGCCAGGAAAGAATGGATCCAGTTCTACACGATAGCAAAGCAGAAAAACAACCAGCGTA
 CCCCTGTAGAAGCTTCTTTGTTTACAGCACTTGATCCATGTAGCCATACTCGAAATTTCAACTCATCTG
 AAACCTTTTCTGAAAGTTGAAAAAGAATGCCATAAGGGTCACCCGAAGCTTATTCACGCCCCGGGAGTT
 ATGACAATTACAACAACAGAATTCTTTCTATATATGCACGAACCTGTAATATGGAAGAAATTATGACG
 TACAAACTATAAAGTAAATATTTTACGTAACACATGGTGCTGTTGTGCTTCTTTTTCAAGAGAATACCA
 ATGACGTATGACTAAGTTTAGGATTTAATGCAGGTGACGGACCCATCTTTCAAACGATTTATATCAGTG
 GCGTCCAAATTGTTAGGTTTTGTTGGTTTCAGCAGGTTTCCTGTTGTGGGTCATATGACTTTGAACCAA
 TGGCCGGCTGCTAGGGCAGCACATAAGGATAATTCACCTGCCAAGACGGCACAGGCAACTATTCTTGC
 TAATTGACGTGCGTTGGTACCAGGAGCGGTAGCATGTGGGCCTTTACACCTAATAAGTCCAACATGG
 CACCTTGTGGTTCTAGAACAGTACCACCACCGATGGTACCTACTTCGATGGATGGCATGGATACGGAA
 ATTCTCAAATCACCGTCCACTTCTTTCATCAATGTTATACAGTTGGAACCTTCGACATTTTGTGCAGGAT

CTTGTCCTAATGCCAAGAAAACAGCTGTCACTAAATTAGCTGCATGTGCGTTAAATCCACCAACAGAC
 CCAGCCATTGCAGATCCAACCAAAT
 TCTTAGCAATGTTCAACTCAACCAATGCGGAAACATCACTTTTTAACACTTTTCTGACAACATCACCAG
 GAATAGTAGCTTCTGCGACGACACTCTTACCACGACCTTCGATCCAGTTGATGGCAGCTGGTTTTTTGT
 CGGTACAGTAGTTACCAGAAACGGAGACAACCTCCATATCTTCCCAGCCATACTCTTCTACCATTTGCT
 TTAATGAGTATTCGACACCCCTTAGAAATCATATTCATACCCATTGCGTCACCAGTAGTTGTTCTAAATC
 TCATGAAGAGTAAATCTCCTGCTAGACAAGTTTGAATATGTTGCAGACGTGCAAATCTTGATGTAGAG
 TAAAAAGCTTTTTTAATTGCGTTTTGTCCCTCTTCTGAGTCTAACCATATCTTACAGGCACCAGATCTTT
 TCAAAGTTGGGAAACGGACTACTGGGCCTCTTGTACATACCATCCTTAGTTAAAAACAGTTGTTGCACCAC
 CGCCAGCATTGATTGCCTTACAGCCACGCATGGCAGAAGCTACCAAACAACCCCTCTGTAGTTGCCATT
 GGTATATGATAAGATGTACCATCGATAACCAAGGGGCCTATAACACCAACGGGCAAAGGCATGTAAC
 CTATAACATTTTCACAACAAGCGCCAAATACGCGGTGCTAGTCATAATTTTTATATGGTAAACGATCA
 GATGCTAATACAGGAGCTTCTGCCAAAATTGAAAGAGCCTTCCTACGTACCGCAACCGCTCTCGTAGT
 ATCACCTAATTTTTTCTCCAAAGCGTACAAAGGTAACCTACCGTGAATAACCAAGGCAGCGACCTCTTT
 GTTCTTCAATTGTTTTGTATTTCCACTACTTAATAATGCTTCTAATTCTTCTAAAGGACGTATTTTCTTAT
 CCAAGCTTTCAATATCGCGGGAATCATCTTCCTCACTAGATGATGAAGGTCTGTAGTCTGATTGCG
 CAGATGATAAACTTTTGACTTTTCGATCCAGAAATGACTGTTTTATTGGTTAAAACTGGTGTAGAAGCCT
 TTTGTACAGGAGCAGTAAAAGACTTCTTGGTGACTTCAGTCTTCACCAATTGGTCTGCAGCCATTATAG
 TTTTTTCTCCTTGACGTAAAGTATAGAGGTATATTAACAATTTTTTGTGATACTTTTATGACATTTGA
 ATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGTTAAAGTGGTTATGCAGCTTTTGCATTTA
 TATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTAATTACCCCAGAAATAAGGCTAAAAAAC
 TAATCGCATTATTATCCTATGGTTGTTAATTTGATTTCGTTGATTGAAGGTTTGTGGGGCCAGGTTACTG
 CCAATTTTTCTCTTCATAACCATAAAAGCTAGTATTGTAGAATCTTTATTGTTTCGGAGCAGTGCGGCG
 CGAGGCACATCTGCGTTTCAGGAACGCGACCGGTGAAGACCAGGACGCACGGAGGAGAGTCTTCCGT
 CGGAGGGCTGTGCCCCGCTCGGCGGCTTCTAATCCGTACTTCAATATAGCAATGAGCAGTTAAGCGTA
 TTAAGTAAAGTTCCAAAGAGAAGGTTTTTTTAGGCTAAGATAATGGGGCTCTTACATTTCCACAACAT
 ATAAGTAAGATTAGATATGGATATGTATATGGTGGTATTGCCATGTAATATGATTATTAAACTTCTTG
 CGTCCATCCAAAAAAAAGTAAGAATTTTTGAAAATTCAATATAAATGACTGCCGACAACAATAGTAT
 GCCCCATGGTGCAGTATCTAGTTACGCCAAATTAGTGCAAAACCAAACACCTGAAGACATTTTGGAAG
 AGTTTCCTGAAATTATTCCATTACAACAAAGACCTAATACCCGATCTAGTGAGACGTCAAATGACGAA
 AGCGGAGAAACATGTTTTTCTGGTCATGATGAGGAGCAAATTAAGTTAATGAATGAAAATTGTATTGT
 TTTGGATTGGGACGATAATGCTATTGGTGCCGGTACCAAGAAAGTTTGTCAATTAATGGAAAATATTG
 AAAAGGGTTTACTACATCGTGCATTCTCCGTCTTTATTTTCAATGAACAAGGTGAATTACTTTTACAAC
 AAAGAGCCACTGAAAAAATAACTTTCCCTGATCTTTGGACTAACACATGCTGCTCTCATCCACTATGTA
 TTGATGACGAATTAGGTTTGAAGGGTAAGCTAGACGATAAGATTAAGGGCGCTATTACTGCGGCGGTG
 AGAAAACTAGATCATGAATTAGGTATTCCAGAAGATGAACTAAGACAAGGGGTAAGTTTCACTTTTT
 AAACAGAATCCATTACATGGCACCAAGCAATGAACCATGGGGTGAACATGAAATTGATTACATCCTAT
 TTTATAAGATCAACGCTAAAGAAAACCTTGACTGTCAACCCAAACGTCAATGAAGTTAGAGACTTCAAA
 TGGGTTTTACCAAATGATTTGAAAACCTATGTTTGCTGACCCAAGTTACAAGTTTACGCCTTGGTTAAG
 ATTATTTGCGAGAATTACTTATTCAACTGGTGGGAGCAATTAGATGACCTTTCTGAAGTGGAAAATGA
 CAGGCAAATTCATAGAATGCTATAACAACGCGTCAATAATATAGGCTACATAAAAAATCATAATAACTT
 TGTTATCATAGCAAAATGTGATATAAAACGTTTCATTTACCTGAAAAATAGTAAAAATAGGCGACAA

AAATCCTTAGTAATATGTAAACCTTATTTTCTTTATTTACCCGGGAGTCAGTCTGACTCTTGCGAGAGA
 TGAGGATGTAATAATACTAATCTCGAAGATGCCATCTAATACATATAGACATACATATATATATATAT
 ACATTCTATATATTCTTACCCAGATTCTTTGAGGTAAGACGGTTGGGTTTTATCTTTTGCAGTTGGTACT
 ATTAAGAACAATCGAATCATAAGCATTGCTTACAAAGAATACACATACGAAATATTAACGATAATGTC
 AATTACGAAGACTGAACTGGACGGTATATTGCCATTGGTGGCCAGAGGTAAAGTTAGAGACATATATG
 AGGTAGACGCTGGTACGTTGCTGTTTGTGCTACGGATCGTATCTCTGCATATGACGTTATTATGGAAA
 ACAGCATTCTGAAAAGGGGATCCTATTGACCAAACCTGTCAGAGTTCTGGTTCAAGTTCCTGTCCAAC
 GATGTTTCGTAATCATTTGGTCGACATCGCCCCAGGTAAGACTATTTTCGATTATCTACCTGCAAAATTG
 AGCGAACCAGTAACAAAACGCAACTAGAAAGACCGCTCTCTATTGGTTCACAAACATAAACTAATTCC
 ATTGGAAGTAATTGTCAGAGGCTACATCACCGGATCTGCTTGGAAAGAGTACGTAAAAACAGGTACTG
 TGCATGGTTTGAAACAACCTCAAGGACTTAAAGAATCTCAAGAGTTCCCAGAACCAATCTTCACCCCA
 TCGACCAAGGCTGAACAAGGTGAACATGACGAAAACATCTCTCTGCCAGGCCGCTGAGCTGGTGG
 GTGAAGATTTGTACGTAGAGTGGCAGAACTGGCTGTAAACTGTACTCCAAGTGCAAAGATTATGCT
 AAGGAGAAGGGCATCATCATCGCAGACACTAAATTGTTTAAAC

SEQ ID NO: 53

Trình tự pAM426

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGT
 CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGG
 GCTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACTACGTCGTAAGG
 CCGTTTCTGACAGAGTAAAATTCTTGAGGGAACCTTCACCATTATGGGAAATGCTTCAAGAAGGTATT
 GACTTAAACTCCATCAAATGGTCAGGTCATTGAGTGTTTTTTATTTGTTGTATTTTTTTTTTTTAGAGA
 AAATCCTCCAATATCAAATTAGGAATCGTAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTG
 CCCTCCTCCTTGTCATATTAATGTTAAAGTGCAATTCTTTTTCTTATCACGTTGAGCCATTAGTATCA
 ATTTGCTTACCTGTATTCTTTACTATCCTCCTTTTTCTCCTTCTTGATAAATGTATGTAGATTGCGTATA
 TAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATTTTCGTGTCGTTTCTATTATGAATTCATTTAT
 AAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTAAGCAAGGATTTTCTTAACTTCTTCGGC
 GACAGCATCACCGACTTCGGTGGTACTGTTGGAACCACTAAATCACCGATTCTGATACCTGCATCCA
 AAACCTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCA
 TTGCAGCAGACAAGATAGTGGCGATAGGGTCAACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGG
 CATGGTTCGTACAAACCAAATGCGGTGTTCTTGCTGGCAAAGAGGCCAAGGACGCAGATGGCAACA
 AACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATT
 ATAATACCATTTAGGTGGGTGGGTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTTG
 AACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCCTCCACAGTTTTTTCTCCATAATCTTGAAGAGGC
 CAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGG
 CCATTCTTGATTCCTTGCACCTCTGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCAT
 CGTCTTCTTTCTCTTACCAAAGTAAATACCTCCCACTAATTCTCTGACAACAACGAAGTCAGTACCTT
 TAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAG
 TTGGCGTACAATTGAAGTTCTTTACGGATTTTTAGTAAACCTTGTTTCAGGTCTAACACTACCGGTACCC
 CATTTAGGACCAGCCACAGCACCTAACAAAACGGCATCAACCTTCTTGGAGGCTTCAGCGCCTCATC
 TGGAAGTGGGACACCTGTAGCATCGATAGCAGCACCAATTAATGATTTTCGAAATCGAACTTGA
 CATTGGAACGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTCTTGACCAACG

TGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATTACAATGGTATATCCTTGAAATATATA
 TAAAAAAGGCGCCTTAGACCGCTCGGCCAAACAACCAATTACTTGTTGAGAAATAGAGTATAATTAT
 CCTATAAATATAACGTTTTTTGAACACACATGAACAAGGAAGTACAGGACAATTGATTTTGAAGAGAAT
 GTGGATTTTGTATGTAATTGTTGGGATTCCATTTTAAATAAGGCAATAATATTAGGTATGTGGATATACT
 AGAAGTTCTCCTCGACCGTCGATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGC
 ATCAGGAAATTGTAAACGTTAATATTTTGTTAAAAATTCGCGTTAAATTTTGTAAATCAGCTCATTTTT
 TAACCAATAGGCCGAAATCGGCCAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGT
 GTTGTTCAGTTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAAC
 CGTCTATCAGGGCGATGGCCCACTACGTGGAAGATCCGAGGCCTAGCTTTAACGAACGCAGAATTTTC
 GAGTTATTAACTTAAATACGCTGAACCCGAACATAGAAATATCGAATGGGAAAAAAAAAACTGCAT
 AAAGGCATTAAAAGAGGAGCGAATTTTTTTTAAATAAAAAATCTTAATAATCATTAAAAGATAAATAAT
 AGTCTATATATACGTATATAAATAAAAAATATTCAAAAAATAAATAAACTATTATTTTAGCGTAAAG
 GATGGGGAAAGAGAAAAAGAAAAAAATTGATCTATCGATTTCAATTCAATTCAATTTATTTCTTTTCGG
 ATAAGAAAGCAACACCTGGCAATTCCTTACCTTCCAATAATTCCAAGAAGCACCACCACAGTAGAG
 ACATGGGAGACCCGGGCCATGGTTAGATAGACATAGGGTAACTAGCAATGATTTGATCAAAATGCTTG
 TATTCATCTCCCATTTCTCGTAAATTTGTCTTTACCTGCATATTGGACCTCTAAAAATTGGCAAAGATAT
 ATAACAGCCATAAGTAAAGGTCTTGGGATATTCTTTGTTGTTAAATACTCTCTGTTTATGTCCTTCCAA
 ACGTCCTCCACTTCCTTATAAATCAGTGTCTGAGCATATTCTTCGTTGACATTGTATTCCTTCATGTAAG
 ATTCTAAAGAGCTTGAACATATGTTTTCTCTCCTGTTCCGCTTTATGAGTCATCAGGTCATTTAATCTCCT
 ACCCAGAATACCACTGTAAACGGAATAAAGGCGGAGCAGATACAGCCCACTCAACTGATTCCTTAGTGA
 AAATATCGCTCATTCTAGATAACAGGTAGTTGTTAGCAAGTTTGCACCACCAGTGATAATAACTACG
 GGATCGTGCTCTTCAGTTGTCTGGTATGTGTCCTTCATTAGCCCATTTTCGCTTCTACCATTAGATTCCTTA
 CGAATTCTTTAACGAACTCCTTCCCACAGTTGAATAAATCAGTTCTACCTTCTTTGGCCAGAACTCCT
 CCATTTCTGTGTAGGTATCCATGAATAATTTGTAAATAGGCTTCATGTATTCCGGCAACGTGTCTAAGC
 AGGTGATCGACCATCTTCCACGGCTTCAGTGAAAATCTTTAACTCCTCGTAAGTCCATATGCGTCAT
 ACGTGTCACTAATAAGTGTTATCACAGCAACTGCCTTAGTGAAAAAACTCTAGCTCTTGAATACTGG
 GGTTCGTAACCAGAACCTAAACCCCAAAAAATAGCATTCAACGATACGATCTCTCAGACATGGGGCATT
 TTTCTTAATATCAAATGCCTTCCACCACTTGCATACGTGACTCAACTCTTCCTTATGTAGGCTCTGCAAT
 AGATTGAACTCCAGTTTAGCTAACTTTAGCAGAGTTTTATTATGGGAGTCTTGTTGCTGATAGAAGGGT
 ATGTAAGTGGGCGGCCTCGATCCTTGGCAATCTCTTCCACAATGGTTGCTTTAAAGCTCTCTGGATTTC
 GTGAATAAAGCGGGGTTTGTACTAAACGCGTCCTTTGTCATAATCGATAGCCTTGATCTTGTGAATCCC
 AGGGCATCTTCAAGAATTATTTGCGCCGGAACCTCATGGACGTAGCCTCATATAATTCCAACAATCCT
 TCAACATCATTGCTAACGATTGTTTAAAAGCACCATTCTTGTCTTTATAGTTATTAAACACATCACAC
 GTGACATAGTATCCTTGTTTACGCATCAGCCTAAACCATAAGCTAGACCTGTGCGCCATTCCAATTATCA
 CCATAGGTCTCGTAAATACATTGCAATGCATGATCAATTTACGTTCAAAATGATACGGAATACCTAA
 ACGTTGAATCTCGTCAATCAGCTTCAACAAATTTGCATGTTTCATAGGAATATCCAATGCTTCCTTTAA
 CAACTGTCTTACTTCCTTCTTTAGATCGTTTACTATTTGCTCCACACCCTGTTCAACTTGTTTCTCATAA
 ATCAAAAATTGATCGCCCCAAATAGAAGGTGGGAAATTTGCAATTGGCCTTATAGGTTTCTCTTCAGTC
 AAGGCCATTGTTTTCTGCAGATCCGGGGTTTTTTCTCCTTGACGTTAAAGTATAGAGGTATATTAACAA
 TTTTTGTTGATACTTTTATTACATTTGAATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGT
 TAAAGTGGTTATGCAGTTTTTGCATTTATATATCTGTTAATAGATCAAAAATCATCGCTTCGCTGATTA
 ATTACCCAGAAATAAGGCTAAAAAACTAATCGCATTATCATCCTATGGTTGTTAATTTGATTCGTTCA

TTTGAAGGTTTGTGGGGCCAGGTTACTGCCAAITTTTCCCTCTTCATAACCATAAAAAGCTAGTATTGTAG
 AATCTTTATTGTTTCGGAGCAGTGC GGCGCGAGGCACATCTGCGTTTCAGGAACGCGACCGGTGAAGAC
 GAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGTCACCCGCTCGGCGGCTTCTAATCCGTA
 AGATCTGCTTTAATTTGGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGC
 GCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCC
 GCTACAGGGCGCGTTCGCGCCATTTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG
 CCTCTTCGCTATTACGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATT
 GGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTGTTTCGGCTGCGGCGAGCGGTATCA
 GCTCACTCAAAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG
 CAAAAGGCCAGCAAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG
 CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
 AGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGA
 TACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCTTTTCTCATAGCTCAGCTGTAGGTATCTCAGTT
 CGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCGACCGCTGCGCC
 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCAC
 TGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACT
 ACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA
 GTTGGTAGCTCTTGATCCGGCAAACAAACCCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAG
 ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
 GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTT
 TAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAA
 TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACTCCCC
 GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
 CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT
 GGTCTTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCG
 CCAGTTAATAGTTTTCGCGAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGT
 ATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAA
 AGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTCAGAAAGTAAGTTGGCCGAGTGTTATCACTCATGGT
 TATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA
 CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGG
 ATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTGGAAAAACGTTCTTCGGGGCGAAAA
 CTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGATCTTCA
 GCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAAATGCCGCAAAAAAGGG
 AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTATCA
 GGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGC
 GCACATTTCCCCGAAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAA
 CGCGAGAGCGCTAATTTTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGA
 AAGCGCTATTTTACCAACGAAGAATCTGTGCTTCATTTTTGTAAAAACAAAAATGCAACGCGAGAGCGC
 TAATTTTTCAAACAAAGAATCTGAGCTGCATTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTT
 TACCAACAAAGAATCTATACTTCTTTTTTGTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAAC
 AAAGCATCTTAGATTACTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCAC
 TGTAGGTCCGTTAAGGTTAGAAGAAGGCTACTTTGGTGTCTATTTCTCTCCATAAAAAAAGCCTGAC

TCCACTTCCCGCGTTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAAGGCATCCCCG
ATTATATTCTATACCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTC
ATTGGTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAATGTTTACA
TTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTTGTCTAAAGAGTAATACT
AGAGATAAACATAAAAAATGTAGAGGTGCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGG
TAGGTTATATAGGGATATAGCACAGAGATATATAGCAAAGAGATACTTTTGAGCAATGTTTGTGGAAG
CGGTATTCGCAATATTTTAGTAGCTCGTTACAGTCCGGTGCCTTTTTGGTTTTTTGAAAGTGCCTCTTCA
GAGCGCTTTTGGTTTTCAAAGCGCTCTGAAGTTCTATACTTTCTAGAGAATAGGAACTTCGGAATAG
GAACTTCAAAGCGTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCA
CTGTTACGTCGCACCTATATCTGCGTGTTGCCTGTATATATATATACATGAGAAGAACGGCATAGTGC
GTGTTTATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCT
GTGATATTATCCCATTCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCTTTAGCTGTTCTATATG
CTGCCACTCCTCAATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCCTTTGATATTGGATCATACTAA
GAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTC

SEQ ID NO: 54

Trình tự pAM322

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGT
CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTCTGGG
GCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACTACGTCGTAAGG
CCGTTTCTGACAGAGTAAAATTCTTGAGGGAACCTTCACCATTATGGGAAATGCTTCAAGAAGGTATT
GACTTAAACTCCATCAAATGGTCAGGTCATTGAGTGTTTTTTATTTGTTGTATTTTTTTTTTTTAGAGA
AAATCCTCCAATATCAAATTAGGAATCGTAGTTTCATGATTTTCTGTTACACCTAACTTTTTGTGTGGTG
CCCTCCTCCTTGTCATATTAATGTTAAAGTGCAATTCTTTTCTTATCACGTTGAGCCATTAGTATCA
ATTTGCTTACCTGTATTCTTTACTATCCTCCTTTTTCTCCTTCTTGATAAATGTATGTAGATTGCGTATA
TAGTTTCGTCTACCCTATGAACATATTCCATTTTGTAATTTCTGTGTCGTTTCTATTATGAATTCATTTAT
AAAGTTTATGTACAAATATCATAAAAAAAGAGAATCTTTTAAGCAAGGATTTTCTTAACTTCTTCGGC
GACAGCATCACCGACTTCGGTGTTACTGTTGGAACCACCTAAATCACCAGTTCTGATACCTGCATCCA
AAACCTTTTAACTGCATCTTCAATGGCCTTACCTTCTTCAGGCAAGTTCAATGACAATTTCAACATCA
TTGCAGCAGACAAGATAGTGGCGATAGGGTCAACCTTATTCTTTGGCAAATCTGGAGCAGAACCGTGG
CATGGTTCGTACAAACCAAATGCGGTGTTCTTGTCTGGCAAAGAGGCCAAGGACGCAGATGGCAACA
AACCCAAGGAACCTGGGATAACGGAGGCTTCATCGGAGATGATATCACCAAACATGTTGCTGGTGATT
ATAATACCATTTAGGTGGGTGGGTCTTAACTAGGATCATGGCGGCAGAATCAATCAATTGATGTTG
AACCTTCAATGTAGGGAATTCGTTCTTGATGGTTTCTCCACAGTTTTTCTCCATAATCTTGAAGAGGC
CAAAAGATTAGCTTTATCCAAGGACCAAATAGGCAATGGTGGCTCATGTTGTAGGGCCATGAAAGCGG
CCATTCTTGTGATTCTTTGCACTTCTGGAACGGTGTATTGTTCACTATCCCAAGCGACACCATCACCAT
CGTCTTCTTTCTCTTACCAAAGTAAATACCTCCCACTAATCTCTGACAACAACGAAGTCAGTACCTT
TAGCAAATTGTGGCTTGATTGGAGATAAGTCTAAAAGAGAGTCGGATGCAAAGTTACATGGTCTTAAG
TTGGCGTACAATTGAAGTTCTTTACGGATTTTGTAGTAAACCTTGTTCAGGTCTAACACTACCGGTACCC
CATTTAGGACCAGCCACAGCACCTAACAAAACGGCATCAACCTTCTTGGAGGCTTCCAGCGCCTCATC
TGGAAGTGGGACACCTGTAGCATCGATAGCAGCACCAATTAATGATTTTCGAAATCGAAGTTGA
CATTGGAACGAACATCAGAAATAGCTTTAAGAACCTTAATGGCTTCGGCTGTGATTCTTGACCAACG

TGGTCACCTGGCAAAACGACGATCTTCTTAGGGGCAGACATTACAATGGTATATCCCTTGAAATATATA
 TAAAAAAGGCGCCTTAGACCGCTCGGCCAAACAACCAATTACTTGTTGAGAAATAGAGTATAATTAT
 CCTATAAATATAACGTTTTTGAACACACATGAACAAGGAAGTACAGGACAATTGATTTTGAAGAGAAT
 GTGGATTTTGATGTAATTGTTGGGATTCCATTTTAAATAAGGCAATAATATTAGGTATGTGGATATACT
 AGAAGTTCTCCTCGACCGTCGATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGC
 ATCAGGAAATTGTAAACGTTAATATTTTGTAAAAATTCGCGTTAAATTTTTGTAAATCAGCTCATT
 TAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGT
 GTTGTTCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAAC
 CGTCTATCAGGGCGATGGCCCACTACGTGGAAGATCCGAGGCCTAGCTTTAACGAACGCAGAATTTTC
 GAGTTATTAACCTTAAATACGCTGAACCCGAACATAGAAATATCGAATGGGAAAAAAAAAACTGCAT
 AAAGGCATTAAAAGAGGAGCGAATTTTTTTTAAATAAAAAATCTTAATAATCATTAAAAGATAAATAAT
 AGTCTATATATACGTATATAAATAAAAAATATTCAAAAAATAAAATAAACTATTATTTTAGCGTAAAG
 GATGGGGAAAGAGAAAAAGAAAAAATTGATCTATCGATTTCAATTCAATTCAATTTATTTCTTTTCGG
 ATAAGAAAGCAACACCTGGCAATTCCTTACCTTCCAATAATTCCAAGAAGCACCACCACAGTAGAG
 ACATGGGAGACCCGGGCCATGGTTAGATAGACATAGGGTAAACTAGCAATGATTTGATCAAATGCTTG
 TATTCATCTCCCATTTCTCGTAAAATTGTCTTTACCTGCATATTGGACCTCTAAAAATTGGCAAAGATAT
 ATAACAGCCATAAGTAAAGGTCTTGGGATATTCTTTGTTGTTAAATACTCTCTGTTTATGTCTTTCCAA
 ACGTCTCCACTTCCTTATAAATCAGTGTCTGAGCATATTCTTCGTTGACATTGTATTCTTCATGTAAG
 ATTCTAAAGAGCTTGAACATATGTTTTCTCTCCTGTTCCGCTTTATGAGTCATCAGGTCATTTAATCTCCT
 ACCCAGAATACCACTGTAAACGGAATAAAGGCGGAGCAGATACAGCCCACTCAACTGATTCCTTAGTGA
 AAATATCGCTCATTCTTAGATAACAGGTAGTTGTTAGCAAGTTTGCACCACCAGTGATAATAACTACG
 GGATCGTGCTCTTCAGTTGTCTGGTATGTGTCTTCATTAGCCCATTTTCGCTTCTACCATTAGATTCCTTA
 CGAATTCTTTAACGAACCTCCTTCCCACAGTTGAATAAATCAGTTCTACCTTCTTTGGCCAGAACTCCT
 CCATTTCTGTGTAGGTATCCATGAATAATTTGTAAATAGGCTTCATGTATTCCGGCAACGTGTCTAAGC
 AGGTGATCGACCATCTTTCCACGGCTTCAGTGAAAATCTTTAACTCCTCGTAAGTCCATATGCGTCAT
 ACGTGTCACTCAATAAGTGTTATCACAGCAACTGCCTTAGTGAAAAAACTCTAGCTCTTGAATACTGG
 GGTTCTGTAACCAGAACCTAAACCCCAAAAAATAGCATTCAACGATACGATCTCTCAGACATGGGGCATT
 TTTCTTAATATCAAATGCCTTCCACCACTTGCATACGTGACTCAACTCTTCCTTATGTAGGCTCTGCAAT
 AGATTGAACTCCAGTTTAGCTAACTTTAGCAGAGTTTTATTATGGGAGTCTTGTTGCTGATAGAAGGGT
 ATGTAATGGGCGGCCTCGATCCTTGGCAATCTCTTCCACAATGGTTGCTTTAAAGCTCTCTGGATTCA
 GTGAATAAAGCGGGGTTTGTACTAAACGCGTCTTTGTCATAATCGATAGCCTTGATCTTGTGAATCCC
 AGGGCATCTTCAAGAATTATTTTCGCCCCGAACTCTCATGGACGTAGCCTCATATAATTCCAACAATCCT
 TCAACATCATTCGCTAACGATTGTTTAAAGCACCATTCTTGTCTTTATAGTTATTAACACATCACAC
 GTGACATAGTATCCTTGTTTACGCATCAGCCTAAACCATAAGCTAGACCTGTGCGCCATTCCAATTATCA
 CCATAGGTCTCGTAAATACATTGCAATGCATGATCAATTTACGTTCAAAATGATACGGAATACCTAA
 ACGTTGAATCTCGTCAATCAGCTTCAACAAATTTGCATGTTTCATAGGAATATCCAATGCTTCCTTTAA
 CAACTGTCTTACTTCCTTCTTTAGATCGTTTACTATTGCTCCACACCCTGTTCAACTTGTTTCTCATAA
 ATCAAAAATTGATCGCCCCAAATAGAAGGTGGGAAATTTGCAATTGGCCTTATAGGTTTCTCTTCAGTC
 AAGGCCATTGTTTTCTGCAGATCCGGGGTTTTTCTCCTTGACGTTAAAGTATAGAGGTATATTAACAA
 TTTTTGTTGATACTTTTATTACATTTGAATAAGAAGTAATACAAACCGAAAATGTTGAAAGTATTAGT
 TAAAGTGGTTATGCAGTTTTTGCATTTATATCTGTAAATAGATCAAAAATCATCGCTTCGCTGATTA
 ATTACCCAGAAATAAGGCTAAAAAACTAATCGCATTATCATCCTATGGTTGTTAATTTGATTGCTTCA

TTTGAAGGTTTGTGGGGCCAGGTTACTGCCAATTTTTCTCTTCATAACCATAAAAGCTAGTATTGTAG
 AATCTTTATTGTTTCGGAGCAGTGCGGCGCGAGGCACATCTGCGTTTCAGGAACGCGACCGGTGAAGAC
 GAGGACGCACGGAGGAGAGTCTTCCTTCGGAGGGCTGTCACCCGCTCGGCGGCTTCTAATCCGTA
 AGATCTGCTTTAATTTGGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGC
 GCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCGCGCTTAATGCGCC
 GCTACAGGGCGCGTCGCGCCATTGCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG
 CCTCTTCGCTATTACGCCAGCTGAATTGGAGCGACCTCATGCTATACCTGAGAAAGCAACCTGACCTA
 CAGGAAAGAGTTACTCAAGAATAAGAATTTTCGTTTTAAACCTAAGAGTCACTTTAAAATTTGTATA
 CACTTATTTTTTTTATAACTTATTTAATAATAAAAAATCATAAATCATAAGAAATTCGCTTATTTAGAAG
 TGTCAACAACGTATCTACCAACGATTTGACCCTTTTCCATCTTTTCGTAAATTTCTGGCAAGGTAGACA
 AGCCGACAACCTTGATTGGAGACTTGACCAACCTCTGGCGAAGAATTGTTAATTAAGAGTCAGTCGA
 CTTAAAACTAGGGACCAATAGCAATTCTGTTTTACGTTGCATTGTTGCACCTGAACTTTCCGTCATGT
 CAATTTGATCATATGAAACTCCATTGGGCAACTTCCAGTTGAAATGATAAAGAATGTTGGCTAGTGGC
 AGTTGAACATTGGCCAAACCTAACGCAGCGCCAGGACACATACGACGTCCAGCCCCAAATGGTAAAT
 ATTCATATTCGGCGCCCATCACTGTTGCCGAAGAGTTTTCAAATCTTTCAGGTATAAACGCTTCTGCAT
 CCTTCCAGTATTCAGGATCTCTATTGATCGCAAACACATTAACGATTAATTTTCGTTTTGTTAGGGATAT
 TATAACCAGCCAAGTTTACTGGCTGACGACATTCTCTAGGTAGCACTAACGGCAAGGGTGGGTGTAGT
 CTAAGAGTCTCTTTGATGACCATATTCAAGTAGGACAATTCTTGTATATCTTCTTCATGTATTTTTCTT
 TCCCATTC AAGGCCTTACGTAATTCAGCCTGAACCTTTTCCATTGCTTTCGGACATTTTATTAGCTCGCT
 TATAGCCCATTCATGGTAGAACTTGAAGTGTGGTCCCTGCACCGAACATGTCCAAAATTATTGCTTT
 GATATTATCCGAAGTCAGAGGAACTCAGCAGAATCCTTTAATCTAAGTAATACATCTAATAGGGTTT
 CGTTGGTTTTTGATGACGTATTTACGGTATGTTTCAGCTACCAAATTGTCAATTAAGTTATCAATCTTTTT
 ACGTAGGCTAGTTAATCTTGCTCTCTTACCGCTCAAGTGATGCAAGAACCTTTTATAGTGGGAAAATATC
 GGCAACATCGAAACCGCCTGTTTGTCTCAGTATTTCTTTAACAATTTTCAGTAAGTTCCTTTTGATCTTTA
 ATTCCTTACCAAACGCAGCACGGGATAGTATAGTGGCAATTAGTTTAAAAACGTTTTCACTTAAATTT
 ACTGGTCTACCACTACCTGAAGCCTTTATTTCCCTGGACTAAATTCCAACATTCTTCTTCCCTCAACGATT
 GAAATGACTTAACCTTTTTTACAGACAACAATTCAAGAGTACAAATCTTCCTTAATTGTCTCCAGTATT
 CCCCATATGGAGCAAGGACAACATCAGTGTTATGATATAAACTATTTCCCCAGTTAAAGTTTCGGGT
 CTATTAGCGAAAGTAATATCGTAGGTTGTAAGAATTTCCCTAGCCCACTTAGGACTCGACACGACTATT
 GTGGGTACCTCTCCAATTGAAGGTGCATTAGCGAACCATATTTTCTCGCTAAATCCCTTACACCCCTG
 TGTGGTGTGGTTCGGATCAAATGGTGCATGTGACCAATGATGGGTAGCCTCCAAGGTTCGGGCAAGGA
 CTTTTTAGTTGACTTACTTCTAGTGGCAAATTTGTACACGAACAACAAAATAGTTGCTAAAGCAATTGA
 TG TAGTTAAAGATAGTGCCATAGCCTTTAAAATTGACTTCATTGTTTTCTAGGCCTTTAGTGAGGGTT
 GAATTCGAATTTTCAAAAATTCTTACTTTTTTTTGGATGGACGCAAAGAAGTTTAATAATCATATTAC
 ATGGCATTACCACCATATACATATCCATATACATATCCATATCTAATCTTACTTATATGTTGTGGAAAT
 GTAAAGAGCCCCATTATCTTAGCCTAAAAAACCTTCTCTTTGGAACCTTTCAGTAATACGCTTAACTGC
 TCATTGCTATATTGAAGTACGGATTAGAAGCCGCCGAGCGGGTGACAGCCCTCCGAAGGAAGACTCTC
 CTCCGTGCGTCTCGTCTTACCGGTGCGGTTCCCTGAAACGCAGATGTGCCTCGCGCCGCACTGCTCCG
 AACAAATAAAGATTCTACAATACTAGCTTTTATGGTTATGAAGAGGAAAAATTGGCAGTAACCTGGCCC
 CACAAACCTTCAATGAACGAATCAAATTAACAACCATAGGATGATAATGCGATTAGTTTTTTAGCCT
 TATTTCTGGGGTAATTAATCAGCGAAGCGATGATTTTTGATCTATTAACAGATATATAAATGCAAAA
 CTGCATAACCACTTTAACTAATACTTTCAACATTTTCGGTTTGTATTACTTCTTATTCAAATGTAATAAA

AGTATCAACAAAAAATTGTTAATATACCTCTATACCTTAAACGTC AAGGAGAAAAAACCCCAAGCTTCC
 CGGGAAAACAATGCAATCGACAACCTCCGTTAACTATCACCTTTCGATCTTATGACTGCCTTGTTAA
 TGGTAAAGTTAGTTTCGACACGTCCAATACTTCCGATACAAATATACCACTGGCGGTTTTTCATGGAAA
 ACAGGGAATTGCTTATGATATTAACAACCAGTGTGGCCGTTTTAATTGGTTGTGTGGTTGTATTGGTAT
 GGAGAAGATCATCAAGTGCCGCTAAGAAGGCCGCCGAATCACCAGTCATTGTCTGCCAAAGAAAGT
 CACTGAAGATGAGGTTGATGACGGCAGAAAGAAAGTTACTGTATTTTTTCGGGACACAAACGGGGACT
 GCGGAAGGTTTTGCGAAAGCTCTAGTTGAAGAAGCCAAGGCAAGGTACGAAAAAGCAGTATTCAAAG
 TTATTGATTTAGATGACTACGCCGAGAAGATGATGAATACGAAGAAAAGCTAAAGAAAGAATCTTTG
 GCATTCTTCTTTTAGCTACCTATGGTGACGGAGAACCAACAGATAACGCCGCTAGATTCTATAAATGG
 TTTACTGAAGGAGAAGAAAAAGGTGAGTGGTTAGATAAGTTACAATACGCTGTCTTTGGATTGGGAAA
 TCGTCAATATGAACACTTCAATAAGATTGCAAAAGTGGTCGATGAAAAATTAGTTGAGCAGGGGGCTA
 AAAGGTTAGTGCCTGTCTGGTATGGGTGATGACGATCAATGTATCGAAGATGATTTTACTGCTTGGAAG
 GAATTGGTTTGGCCAGAATTAGATCAGCTATTGAGGGACGAAGATGACACAAGTGTCTACTCCGTA
 CACCGCCGCTGTTGGCGAATATCGTGTGTTTTTCACGATAAACCTGAACTTACGATCAAGATCAATT
 GACCAACGGACACGCAGTTCACGACGCCCAACACCCATGCAGATCGAACGTTGCGGTCAAGAAAGAA
 TTACACAGTCCCTTATCCGATAGGAGTTGTAATCATTAGAAATTTGATATTTCCAATACTGGACTATCG
 TATGAAACTGGCGACCATGTCTGGTGTATATGTGGAACCTGTCTGAAGTTGTAGATGAAGCCGAAAA
 ATTGATTGGGCTTCCTCCACATACATACTTTCTGTGCATACAGATAATGAAGATGGTACTCCACTTGG
 CGGAGCCTCGTTACCACCTCCCTTTCCACCATGTACACTTAGAAAAGCTCTTGCATCTTATGCAGATGT
 ACTTTCTTACCAAAGAAAAGTGCATTACTAGCTCTAGCCGCCCATGCTACCGACTCTACTGAAGCTGA
 CCGTTTGAAATCTTTGCTTCACCTGCTGGCAAAGACGAGTACGCACAGTGGATTGTGGCATCTCACAG
 ATCATTGCTGGAAGTGATGGAAGCCTTCCCATCGGCAAAGGCCACCATTAGGCGTGTTTTTCGCATCTGT
 TGCCCCACGTTTACAGCCTAGATACTATTCCATATCTTCTAGCCCCAAAATTGCCCCCAATCGTATTCA
 TGTGACGTGTGCGCTGGTGTATGAACAACTCCATCAGGAAGGGTACATAAAGGTGTCTGTAGTACAT
 GGATGAAAAACGCGGTGCCAATGACTGAATCTCAAGATTGTTCTGTTGGGCACCAATTTATGTTCTGACT
 TCTAATTTTAGACTACCTAGTGACCCTAAAGTACCAGTGATTATGATCGGGCCTGGGACAGGACTAGC
 GCCATTACAGAGGTTTCTTACAAGAAAGATTGGCCCCAAAGGAAGCAGGTACGGAATTAGGAACCGCA
 ATTCTATTCTTTGGTTGTCTGTAATAGAAAAGTTGACTTTATATACGAAGATGAGTTAAACAACTTCGTT
 GAAACTGGAGCGTTATCAGAATTAGTGACAGCATTCTCTAGGGAAGGTGCAACAAAAGAATACGTCC
 AACATAAAATGACCCAAAAGGCCAGCGATATATGGAATTTGCTGTCCGAGGGTGCCTATTTGTACGTT
 TGTGGTGATGCAAAGGGAATGGCTAAAGATGTTTACAGGACATTGCATACAATTGTTTACAGGAACAAG
 GTTCCTTGGATTCCCTCTAAGGCAGAACTTTATGTTAAAAACCTTACAGATGGCTGGTAGATATTTGCGTG
 ATGTTTGGTGAGCTAGCTAAGATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGG
 TCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTT
 CTGTACAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTC
 GAAGATCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTT
 CCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCTGCTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA
 AGGCGGTAATACGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
 GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACG
 AGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
 GTTTCCTTGGAAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC
 CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGT

CGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAA
 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA
 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
 AGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTC
 TTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGAAGCAGCAGATTACGCGCA
 GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAAC
 TCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAA
 TGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGT
 GAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATA
 ACTACGATACGGGAGGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC
 GGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTGCAACT
 TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGTTAATAGT
 TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTTCGTTTGGTATGGCTTCATTC
 AGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTC
 CTTCGGTCTCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACT
 GCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTC
 ATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGC
 CACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
 TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT
 TTCACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA
 CACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCT
 CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCC
 GAAAAGTGCCACCTGAACGAAGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCT
 AATTTTTCAAACAAAGAATCTGAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTT
 ACCAACGAAGAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATTTTTTCAA
 CAAAGAATCTGAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAGAGCGCTATTTTACCAACAAAGA
 ATCTATACTTCTTTTTTGTCTACAAAAATGCATCCCGAGAGCGCTATTTTTCTAACAAAGCATCTTAG
 ATTACTTTTTTCTCCTTTGTGCGCTCTATAATGCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTA
 AGGTTAGAAGAAGGCTACTTTGGTGTCTATTTTCTCTCCATAAAAAAAGCCTGACTCCACTTCCCGCG
 TTTACTGATTACTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCGATTATATTCTATA
 CCGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTCATTGGTCAGAAA
 ATTATGAACGTTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAATGTTTACATTTTCGTATTGTT
 TTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTGTCTAAAGAGTAATACTAGAGATAAACAT
 AAAAAATGTAGAGGTCGAGTTTAGATGCAAGTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAG
 GGATATAGCACAGAGATATATAGCAAAGAGATACTTTGAGCAATGTTTGTGGAAGCGGTATTTCGAA
 TATTTTAGTAGCTCGTTACAGTCCGGTGCCTTTTTGGTTTTTTGAAAGTGCGTCTTCAGAGCGCTTTTGG
 TTTTCAAAGCGCTCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAAG
 CGTTTTCCGAAAACGAGCGCTTCCGAAAATGCAACGCGAGCTGCGCACATACAGCTCACTGTTACGTC
 GCACCTATATCTGCGTGTGCTGTATATATATATACATGAGAAGAAGCGCATAGTGCGTGTATGCT
 TAAATGCGTACTTATATGCGTCTATTTATGTAGGATGAAAGGTAGTCTAGTACCTCCTGTGATATTATC
 CCATTCCATGCGGGGTATCGTATGCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCT

CAATTGGATTAGTCTCATCCTTCAATGCTATCATTTCCTTTGATATTGGATCATACTAAGAAACCATTAT
TATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC