



(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) 
1-0019793

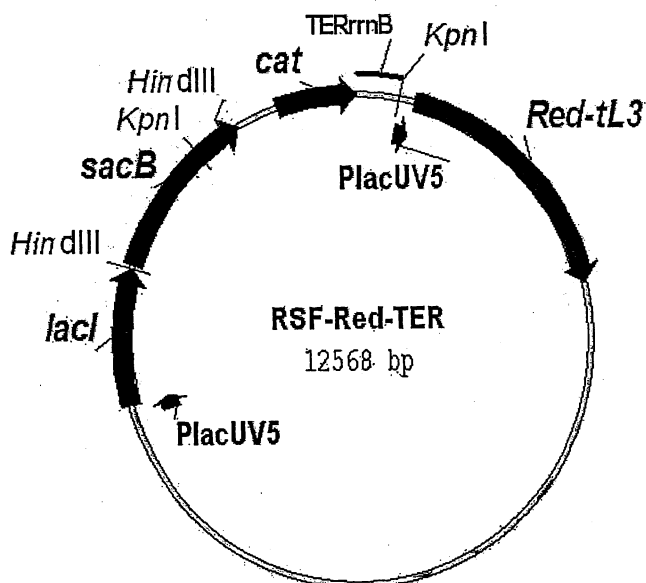
(51)⁷ **C12P 13/24**

(13) **B**

- | | | | | | |
|------|--|---------------|------------|-----------------|------------|
| (21) | 1-2011-00909 | (22) | 04.09.2009 | | |
| (86) | PCT/JP2009/065475 | 04.09.2009 | (87) | WO2010/027045A1 | 11.03.2010 |
| (30) | 2008-229736 | 08.09.2008 JP | | | |
| | 2009-032839 | 16.02.2009 JP | | | |
| (45) | 25.09.2018 | 366 | (43) | 25.08.2011 | 281 |
| (73) | AJINOMOTO CO., INC. (JP)
15-1, Kyobashi 1-chome, Chuo-ku, Tokyo 104-8315, Japan | | | | |
| (72) | Rie TAKIKAWA (JP), Yoshihiko HARA (JP), Gen NONAKA (JP), Kazuhiro TAKUMI (JP) | | | | |
| (74) | Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD) | | | | |

(54) **PHƯƠNG PHÁP SẢN XUẤT AXIT L-AMIN**

(57) Axit L-amin được sản xuất bằng cách nuôi cấy vi khuẩn thuộc họ Enterobacteriaceae có khả năng sản sinh axit L-amin và sẵn có hoạt tính glucoza dehydrogenaza mà sử dụng pyroloquinolin quinon làm coenzym, nhưng đã được cải biến sao cho hoạt tính của glucoza dehydrogenaza bị giảm, trong môi trường để sản sinh và tích lũy axit L-amin trong môi trường nuôi cấy và thu gom axit L-amin từ môi trường nuôi cấy này.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp sản xuất axit L-amin bằng cách sử dụng vi sinh vật, cụ thể là phương pháp sản xuất axit L-amin như axit L-glutamic, L-lysin, L-threonin, L-tryptophan hoặc các axit amin tương tự. Đây là các axit L-amin hữu ích trong công nghiệp, ví dụ, axit L-glutamic hữu ích làm gia vị, và L-lysin, L-threonin và L-tryptophan hữu ích làm chất phụ gia cho thức ăn của động vật, các thành phần thực phẩm cho sức khỏe, dung dịch truyền axit amin, và nhiều thành phần khác.

Tình trạng kỹ thuật của sáng chế

Các axit L-amin được sản xuất ở quy mô công nghiệp bằng cách lên men nhờ sử dụng các vi sinh vật khác nhau. Ví dụ, axit L-glutamic được sản xuất chủ yếu bằng cách lên men nhờ sử dụng vi khuẩn sản sinh axit L-glutamic của vi khuẩn được gọi là coryneform thuộc giống *Brevibacterium*, *Corynebacterium* hoặc *Microbacterium*, hoặc các chủng đột biến của chúng (xem, ví dụ, Kunihiro Akashi et al., "Amino acid fermentation", pp.195-215, 1986, Japan Scientific Societies Press). Về các phương pháp sản xuất axit L-glutamic bằng cách lên men nhờ sử dụng các vi sinh vật khác, các phương pháp sử dụng vi sinh vật thuộc giống *Bacillus*, *Streptomyces*, *Penicillium* hoặc tương tự (xem, ví dụ, đơn yêu cầu cấp patent Nhật (KOKAI) số 5-244970), các phương pháp sử dụng vi sinh vật thuộc giống *Pseudomonas*, *Arthrobacter*, *Serratia*, *Candida* hoặc tương tự (xem, ví dụ, patent Mỹ số 3,563,857), các phương pháp sử dụng vi sinh vật thuộc giống *Bacillus*, *Pseudomonas*, *Serratia*, *Aerobacter aerogenes* (hiện nay được gọi là *Enterobacter aerogenes*) hoặc tương tự (xem, ví dụ, công bố đơn yêu cầu cấp patent Nhật (KOKOKU) số 32-9393), các phương pháp sử dụng

chủng đột biến của *Escherichia coli* (xem, ví dụ, đơn yêu cầu cấp patent Nhật (KOKAI) số 5-244970), và các phương pháp khác đã biết. Ngoài ra, các phương pháp sản xuất axit L-glutamic nhờ sử dụng vi sinh vật thuộc giống *Klebsiella*, *Erwinia*, *Pantoea* hoặc *Enterobacter* cũng đã được đề cập (xem, ví dụ, patent Mỹ số 3,563,857, công bố đơn yêu cầu cấp patent Nhật (KOKOKU) số 32-9393, đơn yêu cầu cấp patent Nhật 2000-189175).

Các phương pháp sản xuất các chất đích như các axit L-amin bằng cách lên men nhờ sử dụng vi sinh vật như nêu trên bao gồm việc sử dụng vi sinh vật kiểu đại (chủng kiểu đại), sử dụng chủng dinh dưỡng thụ động thu được từ chủng kiểu đại, sử dụng chủng đột biến điều hòa sự trao đổi chất thu được từ chủng kiểu đại như chủng kháng một hoặc nhiều thuốc khác nhau, các phương pháp sử dụng chủng là chủng dinh dưỡng thụ động và chủng đột biến điều hòa sự trao đổi chất, và v.v..

Trong những năm gần đây, kỹ thuật tái tổ hợp ADN đã được sử dụng trong quy trình sản xuất các chất đích bằng cách lên men. Ví dụ, khả năng sản sinh axit L-amin của vi sinh vật có thể được cải thiện bằng cách gia tăng mức biểu hiện của gen mã hóa enzym sinh tổng hợp axit L-amin (các patent Mỹ số 5,168,056, và 5,776,736), hoặc bằng cách gia tăng dòng vào của nguồn cacbon đi vào trong hệ sinh tổng hợp axit L-amin (patent Mỹ số 5,906,925).

Glucoza dehydrogenaza phụ thuộc NAD(P) và glucoza dehydrogenaza phụ thuộc PQQ (pyroloquinolin quinon) đều đã biết. Hơn nữa, đã biết rằng kiểu phụ thuộc PQQ của glucoza dehydrogenaza (EC1.1.5.2) bao gồm các enzym của kiểu hòa tan và kiểu liên kết màng, và các enzym của kiểu liên kết màng có mặt trong không gian chất bào (không gian giữa màng bên ngoài và màng bên trong) và có nhiều trong vi khuẩn đường ruột. Sau đây, glucoza dehydrogenaza có mặt trong không gian chất bào và sử dụng PQQ làm coenzym còn được gọi là "GCD".

Đã biết rằng một số vi khuẩn như *Escherichia coli* không có khả năng tổng hợp PQQ, là coenzym của GCD, và do vậy, có thể biểu hiện hoạt tính GCD chỉ khi PQQ được bổ sung (FEMS Microbiol. Lett., 24, 329-333, 1984). Mặt khác, vi khuẩn như vi khuẩn *Pantoea* có khả năng tổng hợp PQQ và có holoenzym GCD.

Về các kỹ thuật liên quan đến GCD, phương pháp sản xuất xenluloza từ glucoza bằng cách sử dụng *Gluconobacter xylinus* trong đó gen mã hóa GCD được loại bỏ (J. Biosci. Bioeng., 99(4), 415-422, 2005), và phương pháp sản xuất axit [5S,6S]-5,6-dihydroxycyclohexa-1,3-dien-1-carboxylic bằng cách sử dụng vi khuẩn *Escherichia* có hoạt tính glucoza dehydrogenaza được giảm, hoặc vi khuẩn *Escherichia* không có hoạt tính glucoza dehydrogenaza (công bố đơn quốc tế WO2006/133898) là đã biết.

Tuy nhiên, tác dụng của sự giảm hoạt tính GCD của vi khuẩn đối với khả năng sản sinh axit L-amin của vi khuẩn chưa được báo cáo trước đó.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất vi sinh vật thuộc họ *Enterobacteriaceae* mà có thể sản sinh một cách hiệu quả axit L-amin, và phương pháp sản xuất hiệu quả axit L-amin bằng cách sử dụng vi sinh vật này.

Mục đích của sáng chế đạt được nhờ sự phát hiện rằng khả năng sản sinh axit L-amin của vi khuẩn đường ruột có sẵn hoạt tính GCD hoặc có hoạt tính GCD được cải thiện bằng cách cải biến vi khuẩn đường ruột sao cho hoạt tính GCD bị giảm.

Theo một khía cạnh, sáng chế đề xuất phương pháp sản xuất axit L-amin bao gồm bước nuôi cấy vi khuẩn thuộc họ *Enterobacteriaceae* và có khả năng sản sinh axit L-amin trong môi trường, và thu gom axit L-amin từ môi trường hoặc vi khuẩn này, trong đó vi khuẩn đã được cải biến sao cho hoạt tính glucoza

dehydrogenaza vốn có mà sử dụng pyroloquinolin quinon làm coenzym bị giảm so với vi khuẩn không bị cải biến.

Theo một khía cạnh khác, sáng chế đề xuất phương pháp như nêu trên, trong đó hoạt tính glucoza dehydrogenaza bị giảm bằng cách bất hoạt gen gcd mã hóa glucoza dehydrogenaza.

Theo một khía cạnh khác nữa, sáng chế đề xuất phương pháp như nêu trên, trong đó gen gcd bao gồm ADN mã hóa trình tự axit amin của SEQ ID NO: 2, hoặc thể biến dị của nó.

Theo một khía cạnh tiếp theo, sáng chế đề xuất phương pháp như nêu trên, trong đó axit L-amin được chọn từ nhóm bao gồm axit L-glutamic, L-lysin, L-threonin, L-arginin, L-histidin, L-isoloxin, L-valin, L-loxin, L-phenylalanin, L-tyrosin, L-tryptophan, L-xystein và hỗn hợp của chúng.

Theo khía cạnh khác, sáng chế đề xuất phương pháp như nêu trên, trong đó axit L-amin là axit L-glutamic hoặc L-xystein.

Theo một khía cạnh nữa, sáng chế đề xuất phương pháp như nêu trên, trong đó axit L-amin là axit L-glutamic, và hoạt tính hoặc các hoạt tính của enzym được chọn từ nhóm bao gồm xitrat syntaza, metyl xitrat syntaza, phosphoenolpyruvat carboxylaza, glutamat dehydrogenaza và các dạng kết hợp của chúng được gia tăng trong vi khuẩn.

Theo một khía cạnh tiếp theo, sáng chế đề xuất phương pháp như nêu trên, trong đó axit L-amin là L-xystein, và hoạt tính hoặc các hoạt tính của enzym được chọn từ nhóm bao gồm 3-phosphoglyxerat dehydrogenaza, serin axetyltransferaza, hệ vận chuyển sulfat/thiosulfat và các dạng kết hợp của chúng được gia tăng, và/hoặc mức biểu hiện của gen *yeaS* được gia tăng.

Theo một khía cạnh khác, sáng chế đề xuất phương pháp như nêu trên,

trong đó vi khuẩn thuộc giống được chọn từ nhóm bao gồm *Pantoea*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Providencia*, *Salmonella*, *Serratia*, *Morganella*, và *Yersinia*.

Mô tả vắn tắt các hình vẽ

Fig. 1 thể hiện cấu trúc của plasmit hỗ trợ RSF-Red-TER.

Fig. 2 thể hiện quy trình tạo ra plasmit hỗ trợ RSF-Red-TER.

Fig. 3 thể hiện trình tự của đoạn khởi đầu Pnlp (SEQ ID No: 67).

Mô tả chi tiết sáng chế

Sau đây, sáng chế sẽ được giải thích chi tiết hơn.

<1> Vi khuẩn

Vi khuẩn theo sáng chế có thể thuộc họ *Enterobacteriaceae*. Vi khuẩn này có sẵn hoạt tính GCD hoặc được kế thừa hoạt tính GCD, và có khả năng sản sinh axit L-amin. Hơn nữa, vi khuẩn này có thể được cải biến sao cho hoạt tính GCD bị giảm. Vi khuẩn theo sáng chế có thể thu được bằng cách cải biến vi khuẩn sản sinh axit L-amin thuộc họ *Enterobacteriaceae* mà sẵn có hoạt tính GCD sao cho hoạt tính GCD bị giảm. Theo cách khác, vi khuẩn theo sáng chế cũng có thể thu được bằng cách truyền khả năng sản sinh axit L-amin cho vi khuẩn thuộc họ *Enterobacteriaceae*, và sẵn có hoạt tính GCD, nhưng đã được cải biến sao cho hoạt tính GCD bị giảm, hoặc tăng cường khả năng sản sinh axit L-amin của vi khuẩn.

Loại axit L-amin cần được sản xuất không bị giới hạn cụ thể, và các ví dụ bao gồm các axit amin bazơ như L-lysin, L-ornithin, L-arginin, L-histidin và L-xitruilin, các axit amin béo như L-isoleuxin, L-alanin, L-valin, L-leuxin và L-glyxin, các axit amin là axit hydroxymonoaminocarboxylic như L-threonin và L-serin, các axit amin vòng như L-prolin, các axit amin thơm như L-phenylalanin, L-tyrosin và L-tryptophan, các axit amin chứa lưu huỳnh như L-cystein, L-

xystin và L-methionin, và các axit amin axit như axit L-glutamic, axit L-aspartic, L-glutamin và L-asparagin. Axit L-glutamic, L-lysin, L-threonin và L-tryptophan là các ví dụ khác. Vi sinh vật theo sáng chế có thể có khả năng sản sinh hai hoặc nhiều loại axit amin.

Thuật ngữ “axit L-amin” có thể bao gồm các axit L-amin ở dạng tự do và các muối của nó như các sulfat, các hydroclorua, các carbonat, các muối amoni, các muối natri, và các muối kali của chúng.

Vi khuẩn có khả năng sản sinh axit L-amin hoặc có thể sản sinh axit L-amin có thể chỉ đến vi khuẩn có khả năng sản sinh axit L-amin và tiết nó vào trong môi trường, nếu vi khuẩn được nuôi cấy trong môi trường. Vi khuẩn có thể sản sinh axit L-amin đến mức axit L-amin tích tụ trong môi trường với lượng bằng hoặc lớn hơn 0,5g/L, hoặc tốt hơn nữa là bằng hoặc lớn hơn 1,0g/L.

Vi khuẩn mà có thể được sử dụng làm chủng cha mẹ để thu vi khuẩn mà được cải biến sao cho hoạt tính GCD bị giảm, và các phương pháp truyền hoặc gia tăng khả năng sản sinh axit L-amin được lấy làm ví dụ dưới đây.

<2-1> Vi khuẩn

Vi khuẩn theo sáng chế có thể thuộc họ *Enterobacteriaceae*.

Họ *Enterobacteriaceae* bao gồm vi khuẩn thuộc các giống *Escherichia*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Pantoea*, *Photorhabdus*, *Providencia*, *Salmonella*, *Serratia*, *Shigella*, *Morganella*, *Yersinia*, và v.v.. Cụ thể, vi khuẩn được phân loại thành họ *Enterobacteriaceae* theo nguyên tắc phân loại được sử dụng bởi cơ sở dữ liệu NCBI (National Center for Biotechnology Information) (<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=91347>) có thể được sử dụng.

“Vi khuẩn thuộc giống *Escherichia*” có nghĩa là vi khuẩn được phân loại

thành giống *Escherichia* theo sự phân loại đã biết đối với người có hiểu biết trung bình trong lĩnh vực vi sinh vật, mặc dù vi khuẩn không bị giới hạn cụ thể ở các vi khuẩn này. Các ví dụ về vi khuẩn thuộc giống *Escherichia* bao gồm, nhưng không bị giới hạn ở, *Escherichia coli* (*E. coli*).

Các ví dụ về vi khuẩn thuộc giống *Escherichia* bao gồm, ví dụ, vi khuẩn được đề cập trong tài liệu của Neidhardt et al. (Neidhardt F.C. Ed., 1996, *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology/Second Edition, pp.2477-2483, Table 1, American Society for Microbiology Press, Washington, D.C.). Các ví dụ cụ thể bao gồm *Escherichia coli* W3110 (ATCC 27325), *Escherichia coli* MG1655 (ATCC 47076) thu được từ chủng kiểu gốc chủng kiểu đại K12, và v.v..

Các chủng này thu được từ, ví dụ, American Type Culture Collection (Address: P.O. Box 1549, Manassas, VA 20108, United States of America). Tức là, số lưu giữ được phân cho từng chủng, và các chủng có thể được sắp xếp bằng cách sử dụng các số này. Số lưu giữ của các chủng được liệt kê trong catalo của American Type Culture Collection.

Các ví dụ về vi khuẩn *Enterobacter* bao gồm, *Enterobacter agglomerans*, *Enterobacter aerogenes*, và v.v.. Cụ thể, các chủng được lấy làm ví dụ trong đơn yêu cầu cấp patent châu Âu số 952221 có thể được sử dụng. Trong những năm gần đây, vài vi khuẩn *Enterobacter agglomerans* được phân loại lại là *Pantoea agglomerans*, *Pantoea ananatis*, hoặc *Pantoea stewartii*, trên cơ sở phân tích trình tự nucleotit của 16S ARN ribosom v.v.. Vi sinh vật có thể thuộc giống *Enterobacter* hoặc *Pantoea* miễn là vi sinh vật này được phân loại thành họ Enterobacteriaceae.

Các chủng được lấy làm ví dụ về giống *Enterobacter* bao gồm chủng

Enterobacter agglomerans ATCC 12287.

Các chủng được lấy làm ví dụ về vi khuẩn *Pantoea* bao gồm *Pantoea ananatis*, *Pantoea stewartii*, *Pantoea agglomerans*, và *Pantoea citrea*. Các ví dụ cụ thể bao gồm các chủng sau đây:

Pantoea ananatis AJ13355 (FERM BP-6614, đơn yêu cầu cấp patent châu Âu số 0952221)

Pantoea ananatis AJ13356 (FERM BP-6615, đơn yêu cầu cấp patent châu Âu số 0952221)

Pantoea ananatis AJ13601 (FERM BP-7207, đơn yêu cầu cấp patent châu Âu số 0952221)

Mặc dù các chủng này được xác nhận và lưu giữ dưới dạng *Enterobacter agglomerans* khi chúng được phân lập, chúng hiện được phân loại là *Pantoea ananatis* trên cơ sở phân tích trình tự nucleotit của 16S ARN ribosom v.v., như nêu trên.

Các ví dụ về vi khuẩn *Erwinia* bao gồm *Erwinia amylovora* và *Erwinia carotovora*, và các ví dụ về vi khuẩn *Klebsiella* bao gồm *Klebsiella planticola*. Các ví dụ cụ thể bao gồm các chủng sau đây:

Erwinia amylovora ATCC 15580

Erwinia carotovora ATCC 15713

Klebsiella planticola AJ13399 (FERM BP-6600, đơn yêu cầu cấp patent châu Âu số 955368)

Klebsiella planticola AJ13410 (FERM BP-6617, đơn yêu cầu cấp patent châu Âu số 955368).

Vi khuẩn theo sáng chế có thể là vi khuẩn đường ruột như nêu trên, và có

thể thuộc họ *Enterobacteriaceae* sao cho nó có sẵn hoạt tính GCD hoặc kế thừa hoạt tính GCD, và có khả năng sản sinh axit L-amin. Vi khuẩn đường ruột mà sẵn có hoạt tính GCD hoặc kế thừa hoạt tính GCD có nghĩa là vi khuẩn kiểu đại hoặc vi khuẩn không được cải biến trong đó gen *gcd* có mặt và biểu hiện protein có hoạt tính GCD. Các ví dụ về vi khuẩn thuộc họ *Enterobacteriaceae* bao gồm vi khuẩn thuộc giống như *Pantoea*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Providencia*, *Salmonella*, *Serratia*, *Morganella*, *Yersinia*, *Citrobacter*, và *Proteus*. Cụ thể hơn, các ví dụ bao gồm vi khuẩn được đề cập trong Int. J. Syst. Bacteriol., 39(1), 61-67, 1989.

Mặc dù *Escherichia coli* có gen *gcd* và sản sinh apoenzym GCD, nó không có khả năng sản sinh PQQ. Do vậy, nó không có hoạt tính GCD trừ khi PQQ được bổ sung. Tuy nhiên, đã biết rằng nếu gen lạ được biểu hiện, chất mà thay thế cho PQQ có thể được tạo ra, và hoạt tính GCD được biểu hiện (WO2006/133898). Vi khuẩn mà thường không có hoạt tính GCD, nhưng có thể biểu hiện hoạt tính GCD như nêu trên, như vi khuẩn *Escherichia*, là ví dụ về "vi khuẩn đường ruột sẵn có hoạt tính GCD" nêu trong sáng chế. Hoạt tính GCD sẽ được giải thích dưới đây.

Sau đây, các phương pháp truyền khả năng sản sinh axit L-amin cho vi khuẩn như được nêu ở trên, hoặc các phương pháp gia tăng khả năng sản sinh axit L-amin của vi khuẩn được mô tả.

Để truyền khả năng sản sinh axit L-amin, các phương pháp được sử dụng thông thường trong quá trình nhân giống vi khuẩn coryneform hoặc vi khuẩn thuộc giống *Escherichia* (xem "Amino Acid Fermentation", Gakkai Shuppan Center (Ltd.), 1st Edition, published May 30, 1986, pp. 77-100) có thể được sử dụng. Các phương pháp này bao gồm việc thu thể đột biến dinh dưỡng thụ động, chủng kháng thể đồng đẳng axit L-amin, hoặc thể đột biến điều hòa sự trao đổi

chất, tạo ra chủng tái tổ hợp trong đó sự biểu hiện quá trình sinh tổng hợp axit L-amin được gia tăng, và v.v.. Ở đây, trong khi nhân giống vi khuẩn sản sinh axit L-amin, một hoặc nhiều đặc tính như đột biến dinh dưỡng sinh trưởng, tính kháng thể đồng đẳng, hoặc đột biến điều hòa sự trao đổi chất có thể được truyền. Sự biểu hiện của các enzym sinh tổng hợp axit L-amin có thể được gia tăng riêng biệt hoặc kết hợp của hai hoặc nhiều enzym. Hơn nữa, các phương pháp truyền các đặc tính như sự đột biến dinh dưỡng sinh trưởng, tính kháng thể đồng đẳng, hoặc sự đột biến điều hòa sự trao đổi chất có thể được kết hợp với các phương pháp gia tăng các enzym sinh tổng hợp.

Chủng đột biến dinh dưỡng thụ động, chủng kháng thể đồng đẳng axit L-amin, hoặc chủng đột biến điều hòa sự trao đổi chất có khả năng sản sinh axit L-amin có thể thu được bằng cách gây phát sinh đột biến thông thường cho chủng cha mẹ hoặc chủng kiểu đại, như chiếu tia X hoặc chiếu xạ UV, hoặc xử lý bằng tác nhân gây đột biến như N-metyl-N'-nitro-N-nitrosoguanidin, v.v., và sau đó chọn các chủng có tính tự dưỡng, tính kháng thể đồng đẳng, hoặc sự đột biến điều hòa sự trao đổi chất và cũng có khả năng sản sinh axit L-amin.

Hơn nữa, sự truyền hoặc gia tăng khả năng sản sinh axit L-amin cũng có thể đạt được bằng cách gia tăng hoạt tính enzym bởi sự tái tổ hợp gen. Ví dụ về phương pháp gia tăng hoạt tính enzym bao gồm, ví dụ, phương pháp cải biến vi khuẩn sao cho mức biểu hiện gen mã hóa enzym liên quan đến quá trình sinh tổng hợp axit L-amin được gia tăng. Về các phương pháp gia tăng mức biểu hiện của gen, sự gia tăng cũng có thể đạt được bằng cách đưa vào plasmit khuếch đại được tạo ra bằng cách đưa đoạn ADN chứa gen này vào trong plasmit thích hợp, ví dụ, vật truyền plasmit chứa ít nhất là gen ảnh hưởng đến sự sao chép và phát triển của plasmit trong vi sinh vật, gia tăng số lượng bản sao của gen trên nhiễm sắc thể bởi sự liên hợp, hiệu ứng chuyển hoặc tương tự, hoặc đưa sự đột biến

vào trong đoạn khởi đầu của gen (xem công bố quốc tế WO95/34672).

Khi gen đích được đưa vào trong plasmid khuếch đại hoặc nhiễm sắc thể nêu trên, đoạn khởi đầu bất kỳ có thể được sử dụng để biểu hiện gen miễn là nó có chức năng trong vi khuẩn đích. Đoạn khởi đầu có thể là đoạn khởi đầu của gen được sử dụng, hoặc đoạn khởi đầu được cải biến. Mức biểu hiện của gen cũng có thể được điều chỉnh bằng cách chọn thích hợp đoạn khởi đầu mà có sẵn trong vi khuẩn coryneform, hoặc bằng cách tạo ra các vùng -35 và -10 của đoạn khởi đầu gần với trình tự liên ứng. Các phương pháp gia tăng mức biểu hiện của gen của enzym như nêu trên được đề cập trong công bố quốc tế WO00/18935, đơn yêu cầu cấp patent châu Âu số 1010755, và v.v..

Sau đây, các phương pháp cụ thể truyền khả năng sản sinh axit L-amin cho vi khuẩn và vi khuẩn được truyền khả năng sản sinh axit L-amin được lấy làm ví dụ dưới đây. Mặc dù các phần mô tả sau đây chủ yếu đề cập đến vi khuẩn *Escherichia*, các phương pháp sau đây cũng có thể được áp dụng cho vi khuẩn đường ruột theo sáng chế.

Vi khuẩn sản sinh L-threonin

Các ví dụ về các vi sinh vật có khả năng sản sinh L-threonin bao gồm vi khuẩn trong đó một hoặc nhiều hoạt tính của các enzym hệ sinh tổng hợp L-threonin được gia tăng. Các ví dụ về các enzym sinh tổng hợp L-threonin bao gồm aspartokinaza III (lysC), aspartat semialdehyt dehydrogenaza (asd), aspartokinaza I (thrA), homoserin kinaza (thrB), threonin syntaza (thrC) được mã hóa bởi operon thr, và aspartat aminotransferaza (aspartat transaminaza) (aspC). Tên gọi của các gen mã hóa cho các enzym tương ứng được đề cập nằm trong dấu ngoặc đơn sau tên của các enzym (sẽ được sử dụng trong toàn bộ bản mô tả này). Trong số các enzym này, aspartat semialdehyt dehydrogenaza,

aspartokinaza I, homoserin kinaza, aspartat aminotransferaza, và threonin syntaza là các ví dụ được đặc biệt ưu tiên. Các gen mã hóa cho enzym sinh tổng hợp L-threonin có thể được đưa vào trong vi khuẩn có sự giảm khả năng phân hủy threonin. Ví dụ về vi khuẩn *Escherichia* có khả năng phân hủy threonin được giảm là chủng TDH6 mà không có hoạt tính threonin dehydrogenaza (đơn yêu cầu cấp patent Nhật số 2001-346578).

Hoạt tính enzym của các enzym sinh tổng hợp L-threonin bị ức chế bởi sản phẩm cuối, L-threonin. Do vậy, để tạo ra chủng sản sinh L-threonin, người ta mong muốn rằng các gen cho các enzym sinh tổng hợp L-threonin được cải biến sao cho các enzym này được giảm nhạy đối với sự ức chế phản hồi bởi L-threonin ở các chủng sản sinh L-threonin. Các gen thrA, thrB, và thrC nêu trên tạo thành operon threonin, mà tạo ra cấu trúc vùng suy giảm. Sự biểu hiện operon threonin bị ức chế bởi isoleuxin và threonin trong môi trường nuôi cấy và cũng bị kìm hãm bởi sự suy giảm. Do vậy, operon threonin có thể được cải biến bằng cách loại bỏ đoạn dẫn đầu ở vùng suy giảm hoặc đoạn suy giảm (xem Lynn, S.P., Burton, W.S., Donohue, T.J., Gould, R.M., Gumport, R.L, và Gardner, J.F., J. Mol. Biol. 194:59-69 (1987); WO02/26993; WO2005/049808).

Đoạn khởi đầu nguyên thể của operon threonin có mặt hướng lên phía trên của operon threonin, và có thể được thay thế bằng đoạn khởi đầu không nguyên thể (xem WO98/04715), hoặc operon threonin đã được cải biến sao cho mức biểu hiện của gen sinh tổng hợp threonin được điều chỉnh bằng gen kìm hãm và đoạn khởi đầu của thực khuẩn thể λ có thể được tạo ra (patent châu Âu số 0593792). Hơn nữa, để cải biến vi khuẩn sao cho nó bị giảm nhạy đối với sự ức chế phản hồi bởi L-threonin, chủng kháng axit α -amino- β -hydroxyisovaleric (AHV) có thể được chọn.

Số lượng bản sao của operon threonin mà được cải biến để giảm nhạy

đối với sự ức chế phản hồi bởi L-threonin có thể được gia tăng, hoặc sự biểu hiện của operon threonin có thể được gia tăng bằng cách gắn nó với đoạn khởi đầu tiềm năng trong vật chủ. Số lượng bản sao cũng có thể được gia tăng bằng cách chuyển operon threonin đến hệ gen bằng cách sử dụng gen nhảy, thể thực khuẩn Mu, hoặc tương tự, ngoài sự khuếch đại sử dụng plasmit.

Khác với sự gia tăng mức biểu hiện của các gen sinh tổng hợp L-threonin, mức biểu hiện của các gen có trong chu trình thủy phân glucoza, chu trình TCA, hoặc chuỗi hô hấp, các gen điều chỉnh sự biểu hiện của các gen này, hoặc các gen liên quan đến sự hấp thu đường cũng có thể được gia tăng. Các ví dụ về các gen đó bao gồm các gen mã hóa transhydrogenaza (pntAB, patent châu Âu số 733712), phosphoenolpyruvat carboxylaza (pepC, WO95/06114), phosphoenolpyruvat syntaza (pps, patent châu Âu số 877090), và gen mã hóa pyruvat carboxylaza từ vi khuẩn coryneform hoặc vi khuẩn *Bacillus* (WO99/18228, đơn yêu cầu cấp patent châu Âu số 1092776 A).

Tính kháng L-threonin, L-homoserin, hoặc cả hai có thể được truyền cho vật chủ bằng cách, ví dụ, gia tăng mức biểu hiện của gen mã truyền tính kháng L-threonin hoặc L-homoserin. Các ví dụ về các gen này bao gồm gen *rhtA* (Livshits, V.A. et al., 2003, Res. Microbiol., 154:123-135), gen *rhtB* (đơn yêu cầu cấp patent châu Âu số 0994190), các gen *rhtC* (đơn yêu cầu cấp patent châu Âu số 1013765), *yfiK*, và *yeaS* (đơn yêu cầu cấp patent châu Âu số 1016710). Các phương pháp truyền tính kháng L-threonin cho vật chủ được nêu trong đơn yêu cầu cấp patent châu Âu số 0994190 và WO90/04636.

Các ví dụ về vi khuẩn sản sinh L-threonin và chủng cha mẹ mà có thể được sử dụng để thu được vi khuẩn này bao gồm, nhưng không bị giới hạn ở các chủng thuộc giống *Escherichia* như *E. coli* TDH-6/pVIC40 (VKPM B-3996) (patent Mỹ số 5,175,107, patent Mỹ số 5,705,371), *E. coli* 472T23/pYN7

(ATCC 98081) (patent Mỹ số 5,631,157), *E. coli* NRRL-21593 (patent Mỹ số 5,939,307), *E. coli* FERM BP-3756 (patent Mỹ số 5,474,918), *E. coli* FERM BP-3519 và FERM BP-3520 (patent Mỹ số 5,376,538), *E. coli* MG442 (Gusyatiner et al., Genetika (in Russian), 14, 947-956 (1978)), *E. coli* VL643 và VL2055 (đơn yêu cầu cấp patent châu Âu số 1149911) và v.v..

Chủng TDH-6 không có gen *thrC*, cũng như có tính đồng hóa sucroza, và gen *ilvA* của nó có sự đột biến rò. Chủng này cũng có sự đột biến ở gen *rhtA*, mà truyền tính kháng threonin hoặc homoserin ở nồng độ cao. Chủng B-3996 chứa plasmit pVIC40, mà thu được bằng cách xen operon *thrA*BC*, bao gồm gen *thrA* đột biến, vào trong vật truyền thu được từ RSF1010. Gen *thrA* đột biến mã hóa aspartokinaza homoserin dehydrogenaza I về cơ bản bị giảm nhạy đối với sự ức chế phản hồi bởi threonin. Chủng B-3996 được lưu giữ ngày 19 tháng 11 năm 1987 ở All-Union Scientific Center of Antibiotics (Nagatinskaya Street 3-A, 117105 Moscow, Russia) dưới số truy cập RIA 1867. Chủng này cũng được lưu giữ ở Russian National Collection of Industrial Microorganisms (VKPM) (1 Dorozhny proezd., 1 Moscow 117545, Russia) ngày 07 tháng 04 năm 1987 dưới số truy cập VKPM B-3996.

E. coli VKPM B-5318 (patent châu Âu số 0593792) cũng có thể được sử dụng làm vi khuẩn sản sinh L-threonin hoặc chủng cha mẹ để thu nó. Chủng B-5318 là chất dinh dưỡng vô cơ đối với isoloxin, và gen kìm hãm thể thực khuẩn lambda nhạy nhiệt Cl và đoạn khởi đầu PR thay thế vùng điều hòa của operon threonin trong plasmit pVIC40. Chủng VKPM B-5318 được lưu giữ dưới dạng lưu giữ quốc tế ở Russian National Collection of Industrial Microorganisms (VKPM) (1 Dorozhny proezd., 1 Moscow 117545, Russia) ngày 3 tháng 5 năm 1990 dưới số truy cập VKPM B-5318.

Gen *thrA* mà mã hóa aspartokinaza homoserin dehydrogenaza I của

Escherichia coli đã được giải thích (các nucleotit ở vị trí từ 337 đến 2799, mã số ngân hàng gen NC_000913.2, gi: 49175990). Gen *thrA* nằm ở giữa các gen *thrL* và *thrB* trên nhiễm sắc thể của *E. coli* K-12. Gen *thrB* mã hóa homoserin kinaza của *Escherichia coli* đã được giải thích (các nucleotit ở vị trí từ 2801 đến 3733, mã số ngân hàng gen NC_000913.2, gi: 49175990). Gen *thrB* nằm giữa các gen *thrA* và *thrC* trên nhiễm sắc thể của *E. coli* K-12. Gen *thrC* mã hóa threonin syntaza của *Escherichia coli* đã được giải thích (các nucleotit ở vị trí từ 3734 đến 5020, mã số ngân hàng gen NC_000913.2, gi: 49175990). Gen *thrC* nằm giữa gen *thrB* và khung đọc mở *yaaX* trên nhiễm sắc thể của *E. coli* K-12. Tất cả các gen này có chức năng là operon threonin đơn. Để gia tăng mức biểu hiện của operon threonin, vùng suy giảm mà ảnh hưởng đến sự phiên mã có thể được loại bỏ khỏi operon (WO2005/049808, WO2003/097839).

Gen *thrA* đột biến mã hóa aspartokinaza homoserin dehydrogenaza I chống lại sự ức chế phản hồi bởi threonin, cũng như các gen *thrB* và *thrC* có thể thu được dưới dạng một operon từ plasmit pVIC40 đã biết mà có mặt trong chủng *E. coli* sản sinh threonin VKPM B-3996. Plasmit pVIC40 được mô tả chi tiết trong patent Mỹ số 5,705,371.

Gen *rhtA* hiện diện ở 18 min trên nhiễm sắc thể *E. coli* gần với operon *glnHPQ*, gen này mã hóa các thành phần của hệ vận chuyển glutamin. Gen *rhtA* là giống với ORF1 (gen *ybiF*, các nucleotit ở vị trí từ 764 đến 1651, mã số ngân hàng gen AAA218541, gi:440181) và nằm giữa các gen *pexB* và *ompX*. Đơn vị biểu hiện protein được mã hóa bởi ORF1 đã được chỉ định gen *rhtA* (*rht*: kháng homoserin và threonin). Ngoài ra, đã nhận thấy rằng sự đột biến *rhtA23* là sự thế A cho G ở vị trí -1 đối với codon khởi đầu ATG (ABSTRACTS of the 17th International Congress of Biochemistry and Molecular Biology in conjunction with Annual Meeting of the American Society for Biochemistry and Molecular

Biology, San Francisco, California August 24-29, 1997, abstract No. 457, đơn yêu cầu cấp patent châu Âu số 1013765).

Gen *asd* của *E. coli* đã được giải thích (các nucleotit ở vị trí từ 3572511 đến 3571408, mã số ngân hàng gen NC_000913.1, gi:16131307), và có thể thu được bằng PCR (phản ứng chuỗi polymeraza; xem White, T.J. et al., Trends Genet., 5, 185 (1989)) sử dụng các đoạn mồi được tạo ra dựa vào trình tự nucleotit của gen này. Gen *asd* của các vi sinh vật khác có thể thu được theo cách tương tự.

Ngoài ra, gen *aspC* của *E. coli* đã được giải thích (các nucleotit ở vị trí từ 983742 đến 984932, mã số ngân hàng gen NC_000913.1, gi:16128895), và có thể thu được bằng PCR. Các gen *aspC* của các vi sinh vật khác có thể thu được theo cách tương tự.

Vi khuẩn sản sinh L-lysine

Các ví dụ về vi khuẩn sản sinh L-lysine thuộc giống *Escherichia* bao gồm các thể đột biến có tính kháng chất đồng đẳng L-lysine. Chất đồng đẳng L-lysine ức chế sự phát triển của vi khuẩn thuộc giống *Escherichia*, nhưng sự ức chế này được giảm nhạy hoàn toàn hoặc một phần khi L-lysine có mặt trong môi trường. Các ví dụ về chất đồng đẳng L-lysine bao gồm, nhưng không bị giới hạn ở, oxalysin, lysine hydroxamat, S-(2-aminoethyl)-L-cysteine (AEC), γ -methyllysine, α -clocaprolactam, và v.v.. Các thể đột biến có tính kháng các chất đồng đẳng lysine này có thể thu được bằng cách xử lý gây phát sinh đột biến nhân tạo thông thường cho vi khuẩn thuộc giống *Escherichia*. Các ví dụ cụ thể về các chủng vi khuẩn hữu ích để sản xuất L-lysine bao gồm *Escherichia coli* AJ11442 (FERM BP-1543, NRRL B-12185; xem patent Mỹ số 4,346,170) và *Escherichia coli* VL611. Trong các vi sinh vật này, sự ức chế phản hồi của aspartokinase bởi L-

lysine được giảm nhảy.

Các ví dụ về vi khuẩn sản sinh L-lysine và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-lysine còn bao gồm các chủng trong đó sự biểu hiện của một hoặc nhiều gen mã hóa enzyme sinh tổng hợp L-lysine được gia tăng. Các ví dụ về các enzyme này bao gồm, nhưng không bị giới hạn ở dihydrodipicolinate syntase (dapA), aspartokinase (lysC), dihydrodipicolinate reductase (dapB), diaminopimelate decarboxylase (lysA), diaminopimelate dehydrogenase (ddh) (patent Mỹ số 6,040,160), phosphoenolpyruvate carboxylase (ppc), aspartate semialdehyde dehydrogenase (asd), diaminopimelate epimerase (dapF), tetrahydrodipicolinate succinylase (dapD), succinyl diaminopimelate deacetylase (dapE), và aspartase (aspA) (đơn yêu cầu cấp patent châu Âu số 1253195). Trong số các enzyme này, dihydrodipicolinate reductase, diaminopimelate decarboxylase, diaminopimelate dehydrogenase, phosphoenolpyruvate carboxylase, aspartate aminotransferase, diaminopimelate epimerase, aspartate semialdehyde dehydrogenase, tetrahydrodipicolinate succinylase, và succinyl diaminopimelate deacetylase là được đặc biệt ưu tiên. Ngoài ra, chủng cha mẹ có thể biểu hiện các mức gia tăng của gen liên quan đến hiệu suất năng lượng (cyo) (đơn yêu cầu cấp patent châu Âu số 1170376), gen mã hóa nicotinamide nucleotide transhydrogenase (pntAB) (patent Mỹ số 5,830,716), gen ybjE (WO2005/073390), hoặc các dạng kết hợp của chúng.

Các ví dụ về chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-lysine còn bao gồm các chủng có hoạt tính được giảm hoặc được loại bỏ của enzyme mà xúc tác phản ứng tạo ra hợp chất khác với L-lysine bằng cách phân nhánh từ chu trình sinh tổng hợp của L-lysine. Các ví dụ về các enzyme này bao gồm homoserine dehydrogenase, lysine decarboxylase (patent Mỹ số 5,827,698), và enzyme malic (WO2005/010175).

Các ví về vi khuẩn sản sinh L-lysin bao gồm *Escherichia coli* WC196 Δ cadA Δ ldc/pCABD2 (WO2006/078039). Chúng được tạo ra bằng cách đưa plasmid pCABD2 được đề cập trong patent Mỹ số 6,040,160 vào trong chủng WC196 có các gen cadA và ldcC bị đứt gãy, mà mã hóa lysin decarboxylaza. Chủng WC196 được nhân giống từ chủng W3110, thu được từ *Escherichia coli* K-12, bằng cách thay thế gen lysC kiểu dại trên nhiễm sắc thể của chủng W3110 bằng gen lysC đột biến mã hóa aspartokinaza III đột biến trong đó threonin ở vị trí 352 được thế bằng isoloxin, dẫn đến sự giảm nhạy sự ức chế phản hồi của nó bởi L-lysin (patent Mỹ số 5,661,012), và tạo ra tính kháng AEC cho chủng thu được (patent Mỹ số 5,827,698). Chủng WC196 được gọi là *Escherichia coli* AJ13069, được lưu giữ tại National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (hiện nay là National Institute of Advanced Industrial Science and Technology, International Patent Organism Depositary, Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japan) ngày 6 tháng 12 năm 1994, và được mang số lưu giữ FERM P-14690. Sau đó, nó được chuyển sang lưu giữ quốc tế dưới các điều khoản của Hiệp ước Budapest ngày 29 tháng 9 năm 1995, và nhận số lưu giữ FERM BP-5252 (patent Mỹ số 5,827,698). Chủng WC196 Δ cadA Δ ldc cũng là vi khuẩn sản sinh L-lysin được ưu tiên. Plasmid pCABD2 chứa gen dapA đột biến thu được từ *Escherichia coli* và mã hóa cho dihydrodipicolinat syntaza (DDPS) có sự đột biến để giảm nhạy đối với sự ức chế phản hồi bởi L-lysin, gen lysC đột biến thu được từ *Escherichia coli* và mã hóa cho aspartokinaza III có sự đột biến để giảm nhạy đối với sự ức chế phản hồi bởi L-lysin, gen dapB thu được từ *Escherichia coli* và mã hóa cho dihydrodipicolinat reductaza, và gen ddh thu được từ *Brevibacterium lactofermentum* và mã hóa cho diaminopimelat dehydrogenaza.

Vi khuẩn sản sinh L-xystein

Khả năng sản sinh L-xystein của vi khuẩn có thể được cải thiện bằng cách gia tăng hoạt tính của enzym của chu trình sinh tổng hợp L-xystein hoặc enzym liên quan đến sự sản sinh hợp chất có vai trò là chất nền của chu trình đó như L-serin, ví dụ, 3-phosphoglyxerat dehydrogenaza, serin axetyltransferaza, và v.v.. 3-phosphoglyxerat dehydrogenaza được đem đi ức chế phản hồi bởi serin, và do vậy hoạt tính của enzym này có thể được gia tăng bằng cách kết hợp gen *serA* đột biến mã hóa cho 3-phosphoglyxerat dehydrogenaza đột biến để cho sự ức chế phản hồi bị suy giảm hoặc được loại bỏ trong vi khuẩn.

Hơn nữa, serin axetyltransferaza được đem đi ức chế phản hồi bởi L-xystein. Do vậy, hoạt tính của enzym này có thể được gia tăng bằng cách kết hợp gen *cysE* đột biến mã hóa cho serin axetyltransferaza nhờ đó sự ức chế phản hồi được suy giảm hoặc được loại bỏ trong vi khuẩn. Vì gen mã hóa cho SAT của *Escherichia coli*, *cysE* đã được tách dòng từ chủng kiểu dại và chủng đột biến tiết L-xystein, và trình tự nucleotit của nó đã được giải thích (Denk, D. and Boeck, A., J. General Microbiol., 133, 515-525 (1987)). Trình tự nucleotit và trình tự axit amin được mã hóa bởi trình tự nucleotit được thể hiện trong các SEQ ID NO: 37 và 38.

Khả năng sản sinh L-xystein cũng có thể được cải thiện bằng cách gia tăng hoạt tính của hệ vận chuyển sulfat/thiosulfat. Nhóm protein hệ vận chuyển sulfat/thiosulfat được mã hóa bởi cụm gen *cysPTWAM* (đơn yêu cầu cấp patent Nhật số 2005-137369, patent châu Âu số 1528108).

Khả năng sản sinh L-xystein của vi khuẩn cũng có thể được cải thiện bằng cách gia tăng mức biểu hiện của gen *yeaS* (đơn yêu cầu cấp patent châu Âu số 1016710). Trình tự nucleotit của gen *yeaS* và trình tự axit amin được mã hóa

bởi gen được thể hiện trong các SEQ ID NO: 39 và 40. Đã biết rằng vi khuẩn sử dụng các codon khác nhau như GTG ngoài ATG làm codon khởi đầu (<http://depts.washington.edu/agro/genomes/students/stanstart.htm>). Mặc dù axit amin tương ứng với codon ban đầu gtg được chỉ định là Val trong các SEQ ID NO: 39 và 40, rất có khả năng là nó thực sự là Met.

Các ví dụ cụ thể về vi khuẩn *Escherichia* có khả năng sản sinh L-xystein và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn này bao gồm, nhưng không giới hạn ở, vi khuẩn *Escherichia* như *E. coli* JM15 được biến nạp với các alen của gen *cysE* khác nhau mã hóa serin axetyltransferaza chống lại sự ức chế phản hồi (patent Mỹ số 6,218,168, đơn yêu cầu cấp patent Nga số 2003121601), *E. coli* W3110 trong đó gen mã hóa protein gây tiết các chất gây độc tế bào được biểu hiện quá mức (patent Mỹ số 5,972,663), chủng *E. coli* có hoạt tính xystein desulphhydraza giảm (đơn yêu cầu cấp patent Nhật số 11-155571), và *E. coli* W3110 trong đó hoạt tính của yếu tố điều khiển sự phiên mã tích cực của regulon (đơn vị điều hòa) xystein được mã hóa bởi gen *cysB* được gia tăng (WO01/27307).

Khả năng sản sinh L-xysteincủa vi khuẩn có thể được cải thiện bằng cách cải biến vi khuẩn sao cho hoạt tính của protein được mã hóa bởi *yhaM* (còn được gọi là "YhaM") bị giảm. Gen *yhaM* là giống như các gen ECK3099, b4470 và *yhaN*, và trước đây còn được gọi là b3109 hoặc b3108.

Các ví dụ về vi khuẩn *Pantoea* có khả năng sản sinh L-xystein bao gồm chủng *Pantoea ananatis* được cải biến sao cho hoạt tính xystein desulphhydraza bị giảm, và chủng *Pantoea ananatis* có gen mã hóa serin axetyltransferaza đột biến nhờ đó sự ức chế phản hồi bởi L-xystein bị giảm. Các ví dụ về chủng cha mẹ để nhân giống vi khuẩn sản sinh L-xystein bao gồm chủng *Pantoea ananatis* AJ13355, chủng SC17, và chủng SC17(0). Chủng AJ13355 là chủng được phân

lập từ đất ở Iwata-shi, Shizuoka-ken, Nhật Bản là chủng mà có thể tăng sinh ở môi trường có độ pH thấp chứa axit L-glutamic và nguồn cacbon, và chủng SC17 là chủng được chọn làm chủng đột biến sinh đờm dài thấp từ chủng AJ13355 (patent Mỹ số 6,596,517). Chủng *Pantoea ananatis* AJ13355 được lưu giữ tại National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology, Ministry of International Trade and Industry (hiện nay là National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository, Address: Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japan) ngày 19 tháng 2 năm 1998 và nhận số lưu giữ FERM P-16644. Sau đó được chuyển sang lưu giữ quốc tế dưới các điều khoản của Hiệp ước Budapest vào ngày 11 tháng 1 năm 1999 và nhận số lưu giữ FERM BP-6614. Chủng SC17(0) là chủng được tạo ra làm chủng kháng gen λ Red để làm đứt gãy gen trong *Pantoea ananatis* (xem Ví dụ tham khảo 1).

Chủng SC17 được mang mã số AJ416, và được lưu giữ tại National Institute of Advanced Industrial Science and Technology, International Patent Organism Depository (địa chỉ: Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, 305-8566, Japan) ngày 4 tháng 2 năm 2009 dưới dạng lưu giữ quốc tế, và nhận số lưu giữ FERM BP-11091. Chủng SC17(0) được lưu giữ tại Russian National Collection of Industrial Microorganisms (VKPM), GNII Genetika (address: Russia, 117545 Moscow, 1 Dorozhny proezd. 1) ngày 21 tháng 9 năm 2005 với số lưu giữ VKPM B-9246.

Một phần L-xystein được sản sinh bởi vi khuẩn có thể thay đổi thành L-xystin trong môi trường bằng cách hình thành liên kết disulfua. Hơn nữa, như được mô tả sau đây, S-sulfoxystein có thể được tạo thành bởi phản ứng của L-xystein và axit thiosulfuric có trong môi trường (Szczepkowski T.W., Nature,

vol. 182 (1958)). Hơn nữa, L-xystein được tạo ra trong các tế bào vi khuẩn có thể được ngưng tụ với keton, aldehyt, hoặc, ví dụ, axit pyruvic, mà tồn tại trong các tế bào này, để tạo ra dẫn xuất thiazolidin nhờ hemithioketal làm chất trung gian (xem patent Nhật số 2992010). Dẫn xuất thiazolidin và hemithioketal có thể tồn tại dưới dạng hỗn hợp được làm cân bằng. Do vậy, khả năng sản sinh L-xystein không bị giới hạn ở khả năng chỉ tích tụ L-xystein trong môi trường hoặc tế bào, mà còn bao gồm khả năng tích tụ L-xystein hoặc dẫn xuất của nó như S-sulfoxystein, dẫn xuất thiazolidin, hoặc hemithioketal hoặc hỗn hợp của nó trong môi trường, ngoài L-xystein.

Vi khuẩn sản sinh L-loxin

Các ví dụ về vi khuẩn sản sinh L-loxin và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-loxin bao gồm, nhưng không bị giới hạn ở, các chủng thuộc giống *Escherichia*, như các chủng *E. coli* kháng loxin (ví dụ, chủng 57 (VKPM B-7386, patent Mỹ số 6,124,121)) hoặc các thể đồng đẳng loxin bao gồm β -2-thienylalanin, 3-hydroxyloxin, 4-azaloxin, 5,5,5-trifloloxin (công bố đơn yêu cầu cấp patent Nhật số 62-34397 và đơn yêu cầu cấp patent Nhật số 8-70879); các chủng *E. coli* thu được bằng phương pháp xử lý gen được đề cập trong WO96/06926; *E. coli* H-9068 (đơn yêu cầu cấp patent Nhật số 8-70879), và v.v..

Vi khuẩn theo sáng chế có thể được cải thiện bằng cách gia tăng mức biểu hiện của một hoặc nhiều gen liên quan đến quá trình sinh tổng hợp L-loxin. Các ví dụ về các gen này bao gồm các gen của operon *leuABCD*, ví dụ cụ thể là gen *leuA* đột biến mã hóa isopropyl malat syntaza được làm giảm nhạy đối với sự ức chế phản hồi bởi L-loxin (patent Mỹ số 6,403,342). Ngoài ra, vi khuẩn theo sáng chế có thể được cải thiện bằng cách gia tăng mức biểu hiện của một hoặc nhiều gen mã hóa protein tiết axit L-amin từ tế bào vi khuẩn. Các ví dụ về

các gen này bao gồm các gen b2682 và b2683 (các gen ygaZH) (đơn yêu cầu cấp patent châu Âu số 1239041 A2).

Vi khuẩn sản sinh L-histidin

Các ví dụ về vi khuẩn sản sinh L-histidin và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-histidin bao gồm, nhưng không bị giới hạn ở, các chủng thuộc giống *Escherichia*, như chủng *E. coli* 24 (VKPM B-5945, RU2003677); chủng *E. coli* 80 (VKPM B-7270, RU2119536); *E. coli* NRRL B-12116 – B12121 (patent Mỹ số 4,388,405); *E. coli* H-9342 (FERM BP-6675) và H-9343 (FERM BP-6676) (patent Mỹ số 6,344,347); *E. coli* H-9341 (FERM BP-6674) (patent châu Âu số 1085087); *E. coli* AI80/pFM201 (patent Mỹ số 6,258,554) và v.v..

Các ví dụ về vi khuẩn sản sinh L-histidin và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-histidin còn bao gồm các chủng trong đó sự biểu hiện của một hoặc nhiều gen mã hóa enzym sinh tổng hợp L-histidin được gia tăng. Các ví dụ về các gen này bao gồm các gen mã hóa ATP phosphoribosyltransferaza (hisG), phosphoribosyl AMP xyclohydrolaza (hisI), phosphoribosyl-ATP pyrophosphohydrolaza (hisIE), phosphoribosylformimino-5-aminoimidazol carboxamit ribotit isomeraza (hisA), amidotransferaza (hisH), histidinol phosphat aminotransferaza (hisC), histidinol phosphataza (hisB), histidinol dehydrogenaza (hisD), và v.v..

Đã biết rằng các enzym sinh tổng hợp L-histidin được mã hóa bởi hisG và hisBHAFI bị ức chế bởi L-histidin, và do vậy khả năng sản sinh L-histidin cũng có thể được gia tăng hiệu quả bằng cách đưa sự đột biến truyền tính kháng sự ức chế phản hồi vào trong gen ATP phosphoribosyltransferaza (hisG) (các patent Nga số 2003677 và 2119536).

Các ví dụ cụ thể về các chủng có khả năng sản sinh L-histidin bao gồm *E. coli* FERM P-5038 và 5048 mà đã được đưa vào bằng vật truyền mang ADN mã hóa enzym sinh tổng hợp L-histidin (đơn yêu cầu cấp patent Nhật số 56-005099), các chủng *E. coli* được đưa rht vào, gen để chuyển axit amin (đơn yêu cầu cấp patent châu Âu số 1016710), chủng *E. coli* 80 được truyền với sulfaguanidin, DL-1,2,4-triazol-3-alanin, và có tính kháng streptomycin (VKPM B-7270, patent Nga số 2119536), và v.v..

Vi khuẩn sản sinh axit L-glutamic

Các ví dụ về vi khuẩn sản sinh axit L-glutamic và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh axit L-glutamic bao gồm, nhưng không bị giới hạn ở, các chủng thuộc giống *Escherichia*, như *E. coli* VL334thrC⁺ (patent châu Âu số 1172433). *E. coli* VL334 (VKPM B-1641) là chủng dinh dưỡng thụ động L-isoloxin và L-threonin có các sự đột biến trong các gen thrC và ilvA (patent Mỹ số 4,278,765). Alen kiểu đại của gen thrC được chuyển bằng phương pháp tải nạp thông thường sử dụng thể thực khuẩn P1 sinh trưởng trên các tế bào của chủng *E. coli* K12 kiểu đại (VKPM B-7). Kết quả là, thu được chủng dinh dưỡng thụ động L-isoloxin VL334thrC⁺ (VKPM B-8961).

Các ví dụ về vi khuẩn sản sinh axit L-glutamic và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh axit L-glutamic còn bao gồm, nhưng không bị giới hạn ở, các chủng trong đó hoạt tính của một hoặc nhiều enzym sinh tổng hợp axit L-glutamic được gia tăng. Các ví dụ về các gen này bao gồm các gen mã hóa glutamat dehydrogenaza (gdhA), glutamin synthetaza (glnA), glutamat synthetaza (gltAB), isoxitrat dehydrogenaza (icdA), aconitat hydrataza (acnA, acnB), xitrat syntaza (gltA), metyl xitrat syntaza (prpC), phosphoenolpyruvat carboxylaza (ppc), pyruvat dehydrogenaza (aceEF, lpdA), pyruvat kinaza (pykA, pykF), phosphoenolpyruvat syntaza (ppsA), enolaza

(eno), phosphoglyxeromutaza (pgmA, pgmI), phosphoglyxerat kinaza (pgk), glyxeraldehyt-3-phosphat dehydrogenaza (gapA), trioza phosphat isomeraza (tpiA), fructoza bisphosphat aldolaza (fbp), phosphofructokinaza (pfkA, pfkB), glucoza phosphat isomeraza (pgi), và v.v.. Trong số các enzym này, glutamat dehydrogenaza, xitrat syntaza, phosphoenolpyruvat carboxylaza, và metyl xitrat syntaza có thể được sử dụng trong phương án khác nữa.

Các ví dụ về các chủng mà đã được cải biến sao cho sự biểu hiện của gen xitrat synthetaza, gen phosphoenolpyruvat carboxylaza, và/hoặc gen glutamat dehydrogenaza được gia tăng bao gồm các gen được đề cập trong các đơn yêu cầu cấp patent châu Âu số 1078989, 955368, và 952221.

Khả năng sản sinh axit L-glutamic của vi khuẩn có thể được cải thiện bằng cách gia tăng hoạt tính của enzym liên quan đến chuỗi hô hấp, ví dụ, enzym oxy hóa kết thúc hô hấp kháng xyan (cioA, cioB).

Các ví dụ về vi khuẩn sản sinh axit L-glutamic và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh axit L-glutamic còn bao gồm các chủng trong đó hoạt tính của enzym mà xúc tác quá trình tổng hợp hợp chất khác với axit L-glutamic bằng cách tổng hợp trực tiếp từ chu trình sinh tổng hợp axit L-glutamic, hoặc hoạt tính của enzym mà xúc tác phản ứng phân hủy hoặc tiêu thụ axit L-glutamic, bị giảm hoặc được loại bỏ. Các ví dụ về các enzym này bao gồm isoxitrat lyaza (aceA), α -ketoglutarat dehydrogenaza (sucA), phosphotransaxetylaza (pta), axetat kinaza (ack), axetohydroxy axit syntaza (ilvG), axetolactat syntaza (ilvI), format axetyltransferaza (pfl), lactat dehydrogenaza (ldh), glutamat decarboxylaza (gadAB), γ -glutamyl transferaza (ggT), γ -glutamylxystein synthetaza (gshA), γ -glutamylputresxin synthetaza (ycjK), và v.v.. Vi khuẩn *Escherichia* không có hoạt tính α -ketoglutarat hoặc có hoạt tính α -ketoglutarat dehydrogenaza được giảm và các phương pháp thu vi

khuẩn được đề cập trong các patent Mỹ số 5,378,616 và 5,573,945.

Cụ thể, các chủng này bao gồm:

E. coli W3110sucA::Km^r

E. coli AJ12624 (FERM BP-3853)

E. coli AJ12628 (FERM BP-3854)

E. coli AJ12949 (FERM BP-4881)

E. coli W3110sucA::Km^r thu được bằng cách làm đứt gãy gen α -ketoglutarat dehydrogenaza (sau đây còn được gọi là "gen sucA") của *E. coli* W3110. Chủng này hoàn toàn không có hoạt tính α -ketoglutarat dehydrogenaza.

Các ví dụ khác về vi khuẩn sản sinh axit L-glutamic bao gồm vi khuẩn *Escherichia* mà kháng chất kháng chuyển hóa axit aspartic. Các chủng này cũng có thể không có hoạt tính α -ketoglutarat dehydrogenaza và bao gồm, ví dụ, *E. coli* AJ13199 (FERM BP-5807) (patent Mỹ số 5,908,768), FFRM P-12379, mà hoạt tính được giảm tiếp để phân hủy axit L-glutamic (patent Mỹ số 5,393,671); AJ13138 (FERM BP-5565) (patent Mỹ số 6,110,714), và v.v..

Các ví dụ về vi khuẩn sản sinh axit L-glutamic *Pantoea ananatis* bao gồm chủng *Pantoea ananatis* AJ13355 nêu trên.

Hơn nữa, các ví dụ về vi khuẩn sản sinh axit L-glutamic của *Pantoea ananatis* còn bao gồm vi khuẩn *Pantoea* không có hoạt tính α -ketoglutarat dehydrogenaza (α KGDH) hoặc có hoạt tính α KGDH được giảm. Các ví dụ về chủng này bao gồm AJ13356 (patent Mỹ số 6,331,419), mà được tạo ra bằng cách loại bỏ gen cấu trúc dưới phân tử α KGDH-E1 (sucA) trong AJ13355, và chủng SC17sucA (patent Mỹ số 6,596,517) mà cũng không có gen sucA, và được chọn từ AJ13355 vì đặc tính tạo đờm thấp của nó. Chủng AJ13356 được

lưu giữ tại National Institute of Bioscience and Human-Technology, Agency of Industrial Science and Technology (hiện nay là cơ quan hành chính độc lập, National Institute of Advanced Industrial Science and Technology, International Patent Organism Depositary (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, Japan, mã bưu điện: 305-8566)) ngày 19 tháng 2 năm 1998, và nhận số lưu giữ FERM P-16645. Sau đó, được chuyển sang lưu giữ quốc tế dưới các điều khoản của Hiệp ước Budapest ngày 11 tháng 1 năm 1999, và nhận số lưu giữ FERM BP-6616. Mặc dù các chủng AJ13355 và AJ13356 được lưu giữ tại nơi lưu giữ nêu trên dưới dạng *Enterobacter agglomerans*, chúng được gọi là *Pantoea ananatis* trong bản mô tả này. Chủng SC17sucA được mang số AJ417, và được lưu giữ tại National Institute of Advanced Industrial Science and Technology, International Patent Organism Depositary vào ngày 26 tháng 2 năm 2004, dưới số lưu giữ FERM BP-08646.

Các ví dụ về vi khuẩn *Pantoea ananatis* sản sinh axit *L-glutamic* còn bao gồm các chủng SC17sucA/RSFCPG+pSTVCB, AJ13601, NP106, và NA1. Chủng SC17sucA/RSFCPG+pSTVCB thu được bằng cách đưa plasmit RSFCPG chứa gen xitrat syntaza (*gltA*), gen phosphoenolpyruvat carboxylaza (*ppsA*), và gen glutamat dehydrogenaza (*gdhA*) thu được từ *Escherichia coli*, và plasmit pSTVCB chứa gen xitrat syntaza (*gltA*) thu được từ *Brevibacterium lactofermentum*, vào trong chủng SC17sucA. Chủng AJ13601 được chọn từ chủng SC17sucA/RSFCPG+pSTVCB để chịu được nồng độ cao của axit *L-glutamic* ở độ pH thấp. Hơn nữa, chủng NP106 thu được từ chủng AJ13601 bằng cách loại bỏ plasmit RSFCPG+pSTVCB. Chủng AJ13601 được lưu giữ tại National Institute of Advanced Industrial Science and Technology, International Patent Organism Depositary (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, Japan, mã bưu điện: 305-8566) ngày 18 tháng 8 năm

1999, và mang số lưu giữ FERM P-17516. Sau đó, được chuyển sang lưu giữ quốc tế dưới các điều khoản của Hiệp ước Budapest ngày 6 tháng 7 năm 2000, và nhận số lưu giữ FERM BP-7207. Hơn nữa, chủng NP106 có RSFPPG trong đó gen *gltA* của RSFCPG nêu trên được thay bằng *prpC* (xem WO2008/020654, các ví dụ được đề cập dưới đây) có thể được sử dụng làm vi khuẩn sản sinh axit L-glutamic.

Vi khuẩn sản sinh L-phenylalanin

Các ví dụ về vi khuẩn sản sinh L-phenylalanin và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-phenylalanin bao gồm, nhưng không bị giới hạn ở, các chủng vi khuẩn *Escherichia* như *E. coli* AJ12739 (*tyrA::Tn10*, *tyrR*) (VKPM B-8197) mà không có chorismat mutaza-prephenat dehydrogenaza và gen ức chế tyrosin (WO03/044191), *E. coli* HW1089 (ATCC 55371) chứa gen *pheA34* kiểu đột biến mã hóa cho chorismat mutaza-prephenat dehydrataza đã được gây đột biến được làm giảm nhạy đối với sự ức chế phản hồi (patent Mỹ số 5,354,672), *E. coli* MWEC101-b (KR8903681), *E. coli* NRRL B-12141, NRRL B-12145, NRRL B-12146, và NRRL B-12147 (patent Mỹ số 4,407,952). Ngoài ra, các chủng sau đây có thể được sử dụng để tạo ra vi khuẩn sản sinh L-phenylalanin: *E. coli* K-12 [W3110(*tyrA*)/pPHAB (FERM BP-3566) chứa các gen mã hóa cho chorismat mutaza-prephenat dehydrataza đã được đột biến cần được làm giảm nhạy đối với sự ức chế phản hồi, *E. coli* K-12 [W3110(*tyrA*)/pPHAD] (FERM BP-12659), *E. coli* K-12 [W3110(*tyrA*)/pPHATerm] (FERM BP-12662), và *E. coli* K-12 [W3110(*tyrA*)/pBR-aroG4, pACMAB] (cũng được gọi là AJ12604 (FERM BP-3579) (patent châu Âu số 488424 B1). Hơn nữa, vi khuẩn *Escherichia* sản sinh L-phenylalanin có hoạt tính được gia tăng của protein được mã hóa bởi gen *yedA* hoặc gen *yddG* cũng có thể được sử dụng (các đơn yêu cầu cấp patent Mỹ

số 2003/0148473 và 2003/0157667, WO03/044192).

Vi khuẩn sản sinh L-tryptophan

Các ví dụ về vi khuẩn sản sinh L-tryptophan và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-tryptophan bao gồm, nhưng không bị giới hạn ở, các chủng vi khuẩn *Escherichia* như *E. coli* JP4735/pMU3028 (DSM10122) và *E. coli* JP6015/pMU91 (DSM10123) mà không có tryptophanyl-tARN synthetaza được mã hóa bởi gen *trpS* đột biến (patent Mỹ số 5,756,345), *E. coli* SV164 (pGH5) mà chứa alen *serA* mã hóa phosphoglyxerat dehydrogenaza và alen *trpE* mã hóa anthranilat syntaza, mà được làm giảm nhạy đối với sự ức chế phản hồi bởi serin và tryptophan (patent Mỹ số 6,180,373), *E. coli* AGX17 (pGX44) (NRRL B-12263), và *E. coli* AGX6(pGX50)aroP (NRRL B-12264) mà không có tryptophanaza (patent Mỹ số 4,371,614), và *E. coli* AGX17/pGX50,pACKG4-pps trong đó khả năng sản sinh phosphoenolpyruvat được gia tăng (WO97/08333, patent Mỹ số 6,319,696). Vi khuẩn sản sinh L-tryptophan thuộc giống *Escherichia* có hoạt tính được gia tăng của protein được mã hóa bởi gen *yedA* hoặc gen *yddG* cũng có thể được sử dụng (các đơn yêu cầu cấp patent Mỹ số 2003/0148473 và 2003/0157667).

Các ví dụ về vi khuẩn sản sinh L-tryptophan và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-tryptophan còn bao gồm các chủng trong đó một hoặc nhiều hoạt tính của các enzym sau đây được gia tăng: anthranilat syntaza (*trpE*), phosphoglyxerat dehydrogenaza (*serA*), 3-deoxy-D-arabinoheptulosonat-7-phosphat syntaza (*aroG*), 3-dehydroquinat syntaza (*aroB*), shikimat dehydrogenaza (*aroE*), shikimat kinaza (*aroL*), 5-enolpyruvylshikimat-3-phosphat syntaza (*aroA*), chorismat syntaza (*aroC*), prephenat dehydrataza, chorismat mutaza, và tryptophan syntaza (*trpAB*). Prephenat dehydrataza và chorismat mutaza được mã hóa bởi gen *pheA* làm enzym chức năng kép (CM-

PD). Trong số các enzym này, phosphoglyxerat dehydrogenaza, 3-deoxy-D-arabinoheptulosonat-7-phosphat syntaza, 3-dehydroquinat syntaza, shikimat dehydrataza, shikimat kinaza, 5-enolpyruvylshikimat-3-phosphat syntaza, chorismat syntaza, prephenat dehydrataza, và chorismat mutaza-prephenat dehydrataza có thể được sử dụng trong phương án khác nữa. Anthranilat syntaza và phosphoglyxerat dehydrogenaza đều trải qua sự ức chế phản hồi bởi L-tryptophan và L-serin, và do vậy, sự đột biến làm giảm nhạy sự ức chế phản hồi có thể được đưa vào trong các gen mã hóa các enzym này. Các ví dụ cụ thể về các chủng có sự đột biến này bao gồm *E. coli* SV164 có anthranilat syntaza kiểu được làm giảm nhạy và chủng biến nạp thu được bằng cách đưa plasmit pGH5 (WO94/08031) chứa gen serA đột biến mã hóa cho phosphoglyxerat dehydrogenaza được làm giảm nhạy đối với sự ức chế phản hồi vào trong *E. coli* SV164.

Các ví dụ về vi khuẩn sản sinh L-tryptophan và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-tryptophan còn bao gồm các chủng mà đã được biến nạp với operon tryptophan chứa gen mã hóa anthranilat syntaza được làm giảm nhạy sự ức chế (các đơn yêu cầu cấp patent Nhật số 57-71397, 62-244382, patent Mỹ số 4,371,614). Hơn nữa, khả năng sản sinh L-tryptophan có thể được truyền bằng cách gia tăng sự biểu hiện của gen mã hóa tryptophan syntaza ở operon tryptophan (*trpBA*). Tryptophan syntaza bao gồm cả các đơn vị có cấu trúc dưới phân tử α và β , mà lần lượt được mã hóa bởi *trpA* và *trpB*. Ngoài ra, khả năng sản sinh L-tryptophan có thể được cải thiện bằng cách gia tăng sự biểu hiện của operon isoxitrat lyaza-malat syntaza (WO2005/103275).

Vi khuẩn sản sinh L-prolin

Các ví dụ về vi khuẩn sản sinh L-prolin và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-prolin bao gồm, nhưng không bị giới hạn ở,

các chủng vi khuẩn *Escherichia* như *E. coli* 702ilvA (VKPM B-8012) mà không có gen *ilvA* và có thể sản sinh L-prolin (patent châu Âu số 1172433).

Vi khuẩn theo sáng chế có thể được cải thiện bằng cách gia tăng sự biểu hiện của một hoặc nhiều gen liên quan đến quá trình sinh tổng hợp L-prolin. Các ví dụ về các gen được ưu tiên cho vi khuẩn sản sinh L-prolin bao gồm gen *proB* mã hóa cho glutamat kinaza mà được giảm nhạy đối với sự ức chế phản hồi bởi L-prolin (patent Đức số 3127361). Ngoài ra, vi khuẩn được sử dụng cho sáng chế có thể được cải thiện bằng cách gia tăng sự biểu hiện của một hoặc nhiều gen mã hóa cho các protein gây ra sự tiết các axit L-amin từ tế bào vi khuẩn. Các ví dụ về các gen này là các gen *b2682* và *b2683* (các gen *ygaZH*) (đơn yêu cầu cấp patent châu Âu số 1239041 A2).

Vi khuẩn *Escherichia* mà sản sinh L-prolin bao gồm các chủng *E. coli* sau đây: NRRL B-12403 và NRRL B-12404 (patent Anh số 2075056), VKPM B-8012 (đơn yêu cầu cấp patent Nga 2000124295), các thể đột biến plasmit được đề cập trong patent Đức 3127361, các thể đột biến plasmit được đề cập trong Bloom F.R. et al (The 15th Miami Winter Symposium, 1983, p. 34), và v.v..

Vi khuẩn sản sinh L-arginin

Các ví dụ về vi khuẩn sản sinh L-arginin và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-arginin bao gồm, nhưng không bị giới hạn ở, các chủng vi khuẩn *Escherichia* như chủng *E. coli* 237 (VKPM B-7925) (đơn yêu cầu cấp patent Mỹ số 2002/058315 A1) và các chủng dẫn xuất của nó mang N-axetylglutamat syntaza đột biến (đơn yêu cầu cấp patent Nga số 2001112869), chủng *E. coli* 382 (VKPM B-7926) (đơn yêu cầu cấp patent châu Âu số 1170358 A1), và chủng sản sinh arginin được biến nạp với gen *argA* mã hóa N-axetylglutamat synthetaza (đơn yêu cầu cấp patent châu Âu số 1170361 A1).

Các ví dụ về vi khuẩn sản sinh L-arginin và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-arginin còn bao gồm các chủng trong đó sự biểu hiện của một hoặc nhiều gen mã hóa enzym sinh tổng hợp L-arginin được gia tăng. Các ví dụ về các gen này bao gồm gen N-axetylglutamyl phosphat reductaza (argC), gen ornithin axetyl transferaza (argJ), gen N-axetylglutamat kinaza (argB), gen axetylornithin transaminaza (argD), gen ornithin carbamoyl transferaza (argF), gen axit argininosuccinic synthetaza (argG), gen axit argininosuccinic lyaza (*argH*), và gen carbamoyl phosphat synthetaza (*carAB*).

Vi khuẩn sản sinh L-valin

Các ví dụ về vi khuẩn sản sinh L-valin và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-valin bao gồm, nhưng không bị giới hạn ở, các chủng mà đã được cải biến để biểu hiện qua mức operon *ilvGMEDA* (patent Mỹ số 5,998,178). Mong muốn loại bỏ vùng này trong operon *ilvGMEDA* mà cần để làm suy giảm sao cho sự biểu hiện của operon không bị suy giảm bởi L-valin được sinh ra. Hơn nữa, gen *ilvA* trong operon được làm đứt gãy sao cho hoạt tính threonin deaminaza bị giảm.

Các ví dụ về vi khuẩn sản sinh L-valin và chủng cha mẹ mà có thể được sử dụng để thu vi khuẩn sản sinh L-valin còn bao gồm các thể đột biến có các sự đột biến amino-axyl t-ARN synthetaza (patent Mỹ số 5,658,766). Ví dụ là *E. coli* VL1970, có sự đột biến trong gen *ileS* mã hóa isoxin tARN synthetaza. *E. coli* VL1970 được lưu giữ tại Russian National Collection of Industrial Microorganisms (VKPM) (1 Dorozhny proezd., 1 Moscow 117545, Russia) ngày 24 tháng 6 năm 1988 dưới số truy cập VKPM B-4411.

Hơn nữa, các chủng đột biến mà cần axit lipoic để sinh trưởng và/hoặc

không có H⁺-ATPaza (WO96/06926) cũng là hữu ích để thu vi khuẩn sản sinh L-valin.

Vi khuẩn sản sinh L-isoloxin

Các ví dụ về vi khuẩn sản sinh L-isoloxin và chủng cha mẹ mà có thể được sử dụng để tạo ra vi khuẩn sản sinh L-isoloxin bao gồm, nhưng không bị giới hạn ở, các thể đột biến kháng 6-dimetylaminopurin (đơn yêu cầu cấp patent Nhật số 5-304969), các thể đột biến kháng chất đồng đẳng isoloxin như thiaisoloxin và isoloxin hydroxamat, và các thể đột biến mà kháng DL-ethionin và/hoặc arginin hydroxamat (đơn yêu cầu cấp patent Nhật số 5-130882). Ngoài ra, các chủng tái tổ hợp được biến nạp với các gen mã hóa các protein liên quan đến quá trình sinh tổng hợp L-isoloxin như threonin deaminaza và axetohydroxat syntaza, cũng hữu hiệu để tạo ra vi khuẩn sản sinh L-isoloxin (đơn yêu cầu cấp patent Nhật số 2-458, FR 0356739, và patent Mỹ số 5,998,178).

Vi khuẩn sản sinh L-tyrosin

Các ví dụ về vi khuẩn sản sinh tyrosin bao gồm vi khuẩn *Escherichia* có gen prephenat dehydrataza được làm giảm nhạy (tyrA) được làm giảm nhạy đối với sự ức chế bởi tyrosin (đơn yêu cầu cấp patent châu Âu số 1616940).

Khi vi khuẩn sản sinh axit L-amin nêu trên được nhân giống bằng cách tái tổ hợp gen, các gen được sử dụng không bị giới hạn ở các gen có thông tin di truyền nêu trên hoặc các gen có các trình tự đã biết, mà còn bao gồm các gen có các sự đột biến bảo toàn, như các thể tương đồng hoặc các gen được cải biến nhân tạo, cũng có thể được sử dụng miễn là chức năng của các protein được mã hóa không bị suy thoái. Tức là, chúng có thể là các gen mã hóa trình tự axit amin đã biết chứa một hoặc nhiều sự thể đoạn, khuyết đoạn, xen đoạn, thêm đoạn hoặc tương tự của một hoặc vài gốc axit amin ở một hoặc vài vị trí.

Mặc dù số lượng “vài” gốc axit amin được đề cập ở đây có thể khác nhau tùy thuộc vào vị trí ở cấu trúc ba chiều hoặc các kiểu của các gốc axit amin của protein, cụ thể, có thể là từ 1 đến 20, tốt hơn nữa là từ 1 đến 10, vẫn tốt hơn nữa là từ 1 đến 5 gốc axit amin. Sự đột biến bảo toàn là sự đột biến trong đó sự thế đoạn thực hiện trong số Phe, Trp, và Tyr, nếu vị trí thế đoạn là axit amin thơm; trong số Leu, Ile và Val, nếu nó là axit amin kỵ nước; giữa Gln và Asn, nếu nó là axit amin phân cực; trong số Lys, Arg và His, nếu nó là axit amin kiềm; giữa Asp và Glu, nếu nó là axit amin axit; và giữa Ser và Thr, nếu nó là axit amin có nhóm hydroxyl. Sự đột biến bảo toàn là sự thế đoạn bảo toàn, và các sự thế đoạn được cho là sự thế đoạn bảo toàn bao gồm, cụ thể, sự thay thế của Ser hoặc Thr cho Ala, sự thay thế của Gln, His hoặc Lys cho Arg, sự thay thế của Glu, Gln, Lys, His hoặc Asp cho Asn, sự thay thế của Asn, Glu hoặc Gln cho Asp, sự thay thế của Ser hoặc Ala cho Cys, sự thay thế của Asn, Glu, Lys, His, Asp hoặc Arg cho Gln, sự thay thế của Gly, Asn, Gln, Lys hoặc Asp cho Glu, sự thay thế của Pro cho Gly, sự thay thế của Asn, Lys, Gln, Arg hoặc Tyr cho His, sự thay thế của Leu, Met, Val hoặc Phe cho Ile, sự thay thế của Ile, Met, Val hoặc Phe cho Leu, sự thay thế của Asn, Glu, Gln, His hoặc Arg cho Lys, sự thay thế của Ile, Leu, Val hoặc Phe cho Met, sự thay thế của Trp, Tyr, Met, Ile hoặc Leu cho Phe, sự thay thế của Thr hoặc Ala cho Ser, sự thay thế của Ser hoặc Ala cho Thr, sự thay thế của Phe hoặc Tyr cho Trp, sự thay thế của His, Phe hoặc Trp cho Tyr, và sự thay thế của Met, Ile hoặc Leu cho Val. Sự thế đoạn, khuyết đoạn, xen đoạn, thêm đoạn, đảo đoạn hoặc tương tự của các axit amin nêu trên có thể là kết quả của sự đột biến xảy ra tự nhiên hoặc sự biến dị do sự khác biệt cá thể hoặc sự khác biệt của các loài vi sinh vật từ đó các gen được tạo ra (thể đột biến hoặc thể biến dị). Các gen này có thể thu được bằng cách, ví dụ, cải biến trình tự nucleotit đã biết của gen bằng cách gây đột biến đặc hiệu vị trí sao cho các gốc axit amin

ở các vị trí đặc hiệu của protein được mã hóa bao gồm sự thể đoạn, khuyết đoạn, xen đoạn hoặc thêm đoạn của các gốc axit amin.

Hơn nữa, các gen này có (các) sự đột biến bảo toàn như nêu trên có thể mã hóa protein có độ tương đồng bằng hoặc lớn hơn 80%, tốt hơn là bằng hoặc lớn hơn 90%, tốt hơn nữa là bằng hoặc lớn hơn 95%, tốt nhất là bằng hoặc lớn hơn 97%, so với toàn bộ trình tự axit amin được mã hóa và có chức năng tương đương với chức năng của protein kiểu đại. Trong bản mô tả này, thuật ngữ "độ tương đồng" cũng có thể được sử dụng để chỉ "mức độ giống nhau".

Hơn nữa, các codon trong các chuỗi gen có thể được thay thế bằng các codon khác để sử dụng ở vật chủ trong đó các gen được đưa vào.

Các gen có các sự đột biến bảo toàn có thể thu được bằng cách các phương pháp thường được sử dụng trong các quá trình gây phát sinh đột biến như các quá trình xử lý bằng các tác nhân gây đột biến.

Hơn nữa, các gen có thể là ADN mà có thể lai với chuỗi bổ trợ của chuỗi gen đã biết hoặc đoạn dò mà có thể được tạo ra từ chuỗi bổ trợ dưới các điều kiện nghiêm ngặt và mã hóa protein có chức năng tương đương với chức năng của gen đã biết. Thuật ngữ "các điều kiện nghiêm ngặt" được đề cập ở đây là các điều kiện dưới đó thể lai đặc trưng được tạo ra, và không tạo ra thể lai không đặc trưng. Các ví dụ về các điều kiện nghiêm ngặt bao gồm các điều kiện dưới đó các ADN có độ tương đồng cao lai với nhau, ví dụ, các ADN có độ tương đồng không thấp hơn 80%, tốt hơn là có độ tương đồng không thấp hơn 90%, tốt hơn nữa là có độ tương đồng không thấp hơn 95%, tốt nhất là có độ tương đồng không thấp hơn 97%, được lai với nhau, và các ADN có độ tương đồng thấp hơn các mức trên không được lai với nhau, hoặc các điều kiện rửa một lần, tốt hơn là 2 hoặc 3 lần, ở nồng độ muối và nhiệt độ tương ứng với quá trình rửa điển hình

của phép lai phương Nam, tức là, 1 x SSC, 0,1% SDS ở 60°C, tốt hơn là 0,1 x SSC, 0,1% SDS ở 60°C, tốt hơn nữa là 0,1 x SSC, 0,1% SDS ở 68°C.

Về đoạn dò, một phần của trình tự mà bổ trợ cho gen cũng có thể được sử dụng. Đoạn dò này có thể được điều chế bằng PCR sử dụng các oligonucleotit được tạo ra trên cơ sở chuỗi gen đã biết làm đoạn mẫu và đoạn ADN chứa các trình tự nucleotit làm khuôn mẫu. Ví dụ, khi đoạn ADN có độ dài khoảng 300bp được sử dụng làm đoạn dò, các điều kiện rửa của quá trình lai có thể là 50°C, 2 x SSC và 0,1% SDS.

Các phần mô tả nêu trên liên quan đến các thể biến dị của các gen có thể được áp dụng một cách tương tự cho gen gcd được mô tả dưới đây.

<2-2> Sự giảm hoạt tính GCD

Sau đây, sự cải biến của vi khuẩn thuộc họ *Enterobacteriaceae* để làm giảm hoạt tính GCD của nó sẽ được giải thích.

Hoạt tính GCD nghĩa là hoạt tính xúc tác phản ứng sau đây.

β -D-glucoza + PQQ được oxy hóa $\rightarrow$ D- δ -gluconolacton + PQQ được khử

Hoạt tính GCD có thể được đo, ví dụ, trên cơ sở phát hiện sự hình thành DCPIP được khử nhờ các phản ứng sau đây bằng cách đo quang phổ ở bước sóng 600nm (đơn yêu cầu cấp patent Nhật số 2007-129965).

D-glucoza + PMS được oxy hóa $\rightarrow$ D-glucono-1,5-lacton + PMS được khử

PMS được khử + DCPIP được oxy hóa $\rightarrow$ PMS được oxy hóa + DCPIP được khử

PMS: phenazin metosulfat

DCPIP: 2,6-diclophenolindophenol

Cụm từ "được cải biến sao cho hoạt tính GCD bị giảm" tức là hoạt tính GCD trên tế bào của vi khuẩn thuộc họ *Enterobacteriaceae* là thấp hơn hoạt tính

của chủng không được cải biến, như chủng kiểu dại, của vi khuẩn. Điều này có nghĩa là, ví dụ, số lượng phân tử GCD trên tế bào bị giảm so với số lượng của chủng cha mẹ hoặc chủng kiểu dại, hoặc hoạt tính của GCD trên một phân tử bị giảm so với hoạt tính của chủng cha mẹ hoặc chủng kiểu dại. Hoạt tính GCD trên tế bào có thể được so sánh bằng cách so sánh hoạt tính GCD trong phần chiết tế bào của chủng kiểu dại hoặc chủng cha mẹ và chủng cải biến được nuôi cấy dưới các điều kiện giống nhau. Thuật ngữ "khử" hoạt tính bao gồm sự biến mất hoàn toàn hoạt tính. Vi khuẩn *Pantoea* kiểu dại được sử dụng làm đối chứng để so sánh có thể là, ví dụ, chủng *Pantoea ananatis* AJ13355 (FERM BP-6615), hoặc chủng tương tự.

Việc khử hoạt tính của GCD có thể đạt được bằng cách làm bất hoạt gen mã hóa GCD (gcd). "Bất hoạt" gen gcd có nghĩa là sự cải biến gen bằng cách tái tổ hợp gen hoặc gây sự đột biến trong gen mà hoạt tính của GCD được mã hóa bởi gen bị giảm hoặc được loại bỏ.

Các ví dụ về gen gcd bao gồm gen gcd của *Pantoea ananatis* có trình tự nucleotit được thể hiện trong SEQ ID NO: 1. Trình tự axit amin của GCD được mã hóa bởi gen gcd này được thể hiện trong SEQ ID NO: 2. Gen gcd có thể được tách dòng bằng cách tiến hành PCR sử dụng oligonucleotit tổng hợp được tổng hợp trên cơ sở trình tự nêu trên và nhiễm sắc thể của *Pantoea ananatis* làm khuôn mẫu. Hơn nữa, khi gen gcd được loại bỏ bằng cách tái tổ hợp tương đồng, gen có độ tương đồng cao hơn mức cụ thể, ví dụ, bằng hoặc cao hơn 80%, tốt hơn là bằng hoặc cao hơn 90%, tốt hơn nữa là bằng hoặc cao hơn 95%, so với gen gcd trên nhiễm sắc thể cũng có thể được sử dụng. Hơn nữa, gen có thể lai với gen gcd trên nhiễm sắc thể dưới các điều kiện nghiêm ngặt cũng có thể được sử dụng. Các ví dụ về các điều kiện nghiêm ngặt bao gồm, ví dụ, rửa một lần, tốt hơn nữa là rửa hai hoặc ba lần, ở các nồng độ muối tương ứng với 1 x SSC,

0,1% SDS, tốt hơn là 0,1 x SSC, 0,1% SDS, ở 60°C.

Cụ thể, sự bất hoạt gen gcd có thể đạt được bằng cách, ví dụ, loại bỏ một phần hoặc toàn bộ vùng mã hóa của gen gcd trên nhiễm sắc thể, hoặc xen trình tự khác vào trong vùng mã hóa. Các kỹ thuật này còn được gọi là kỹ thuật làm đứt gãy gen.

Gen gcd cũng có thể được bất hoạt bằng cách làm giảm mức biểu hiện của gen gcd nhờ sự cải biến trình tự điều chỉnh mức biểu hiện như đoạn khởi đầu hoặc trình tự Shine Dargarno (SD) của gen gcd, hoặc tương tự. Việc làm giảm mức biểu hiện bao gồm việc làm giảm mức phiên mã và làm giảm mức dịch mã. Sự biểu hiện của gen cũng có thể được làm giảm bằng cách cải biến vùng không dịch mã khác với các vùng điều chỉnh mức biểu hiện.

Hơn nữa, toàn bộ gen đích bao gồm các vùng hướng lên phía trên và các vùng hướng xuống phía dưới của gen đích trên nhiễm sắc thể có thể được loại bỏ. Ngoài ra, sự bất hoạt gen gcd cũng có thể đạt được bằng cách đưa sự đột biến để thay thế axit amin (đột biến nhầm nghĩa), codon kết thúc (đột biến vô nghĩa), hoặc đột biến dịch khung mà thêm hoặc loại bỏ một hoặc hai nucleotit vào trong hoặc ra khỏi vùng mã hóa của gen gcd trên nhiễm sắc thể (Journal of Biological Chemistry, 272:8611-8617 (1997); Proceedings of the National Academy of Sciences, USA, 95 5511-5515 (1998); Journal of Biological Chemistry, 266, 20833-20839 (1991)).

Sự cải biến gen có thể đạt được bằng cách tái tổ hợp gen. Các ví dụ cụ thể về phương pháp dựa vào sự tái tổ hợp gen bao gồm sự loại bỏ một phần hoặc toàn bộ trình tự của trình tự điều chỉnh mức biểu hiện của gen đích trên nhiễm sắc thể, ví dụ, vùng khởi đầu, hoặc vùng mã hóa hoặc vùng không mã hóa, và xen trình tự khác vào trong các vùng này.

Sự cải biến trình tự điều chỉnh mức biểu hiện có thể được thực hiện đối với một hoặc nhiều nucleotit, tốt hơn nữa là cho hai hoặc nhiều nucleotit, tốt nhất là ba hoặc nhiều nucleotit. Khi vùng mã hóa được loại bỏ, vùng được loại bỏ có thể là vùng có đầu tận cùng N, vùng bên trong hoặc vùng có đầu tận cùng C, hoặc thậm chí toàn bộ vùng mã hóa, miễn là chức năng của protein được hình thành bởi gen bị giảm hoặc được loại bỏ. Sự loại bỏ các vùng dài hơn có thể thường bất hoạt gen đích. Hơn nữa, các khung đọc hướng lên phía trên và hướng xuống phía dưới của vùng cần được loại bỏ không cần phải giống nhau hoàn toàn.

Khi trình tự khác được xen vào trong vùng mã hóa, trình tự này có thể được xen vào trong vùng bất kỳ của gen đích, và sự xen vùng dài hơn có thể chắc chắn làm bất hoạt gen đích. Các khung đọc hướng lên phía trên và hướng xuống phía dưới của vị trí đích không cần phải giống nhau. Trình tự khác không bị giới hạn cụ thể miễn là trình tự mà làm giảm hoặc loại bỏ chức năng của protein được mã hóa bởi gen đích được chọn, và các ví dụ bao gồm, ví dụ, gen nhảy mang gen kháng kháng sinh, gen hữu ích để sản sinh axit L-amin, và v.v..

Sự cải biến gen đích trên nhiễm sắc thể như nêu trên có thể đạt được bằng cách, ví dụ, tạo ra gen kiểu khuyết đoạn trong đó một phần trình tự của gen đích được loại bỏ sao cho nó không thể tạo ra protein mà có thể có chức năng bình thường, và biến nạp vi khuẩn với ADN chứa gen kiểu khuyết đoạn để gây ra sự tái tổ hợp tương đồng giữa gen kiểu khuyết đoạn và gen đích trên nhiễm sắc thể và nhờ đó thay thế gen kiểu khuyết đoạn cho gen đích trên nhiễm sắc thể. Protein được mã hóa bởi gen kiểu khuyết đoạn có cấu hình khác với cấu hình của protein kiểu đại, ngay cả khi nó được tạo ra, và như vậy, chức năng của nó bị giảm hoặc được loại bỏ. Sự làm đứt gãy gen trên cơ sở thay thế gen bằng cách tái tổ hợp tương đồng đã được thực hiện, và phương pháp này được gọi là

phương pháp kết hợp bức xạ đỏ (Datsenko, K.A, và Wanner, B.L., Proc. Natl. Acad. Sci. USA, 97:6640-6645 (2000)), phương pháp sử dụng AND mạch thẳng như phương pháp sử dụng sự kết hợp bức xạ đỏ kết hợp với hệ dễ bị kích thích thu được từ thể thực khuẩn λ (Cho, E.H., Gumport, R.I., Gardner, J.F., J. Bacteriol., 184:5200-5203 (2002)), phương pháp sử dụng plasmit chứa điểm khởi đầu sao chép nhạy nhiệt hoặc plasmit có khả năng vận chuyển liên hợp, phương pháp sử dụng vật truyền tự chết không có điểm khởi đầu sao chép ở vật chủ (patent Mỹ số 6,303,383, đơn yêu cầu cấp patent Nhật số 05-007491), và v.v..

Sự giảm mức phiên mã của gen đích có thể được xác nhận bằng cách so sánh lượng ARN thông tin được phiên mã từ gen đích với lượng ở chủng kiểu dại hoặc chủng không được cải biến. Các ví dụ về phương pháp đo lượng ARN thông tin bao gồm phép lai phương Bắc, RT-PCR, và v.v. (Molecular Cloning, Cold spring Harbor Laboratory Press, Cold Spring Harbor (USA), 2001). Mặc dù mức phiên mã có thể được giảm đến mức bất kỳ miễn là nó giảm so với mức quan sát được ở chủng kiểu dại hoặc chủng không được cải biến, tốt hơn là được giảm đến ít nhất bằng hoặc thấp hơn 75%, bằng hoặc thấp hơn 50%, bằng hoặc thấp hơn 25%, hoặc bằng hoặc thấp hơn 10%, của mức mà quan sát được ở, ví dụ, chủng kiểu dại hoặc chủng không được cải biến, và tốt nhất là gen không được biểu hiện hoàn toàn.

Sự giảm mức protein được mã hóa bởi gen đích có thể được xác nhận bởi pháp thâm tách phương Tây sử dụng các kháng thể mà liên kết với protein (Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (USA) 2001). Mặc dù lượng protein có thể được giảm đến mức bất kỳ miễn là nó giảm so với lượng protein được quan sát ở chủng kiểu dại hoặc chủng không được cải biến, tốt hơn là được giảm đến ít nhất bằng hoặc thấp hơn 75%, bằng hoặc thấp hơn 50%, bằng hoặc thấp hơn 25%, hoặc bằng hoặc thấp hơn 10%,

của lượng protein quan sát được ở, ví dụ, chủng kiểu dại hoặc chủng không được cải biến, và tốt nhất là protein hoàn toàn không được tạo ra (hoạt tính đã hoàn toàn biến mất).

Các ví dụ về phương pháp làm giảm hoạt tính của GCD bao gồm, ngoài các kỹ thuật xử lý di truyền nêu trên, ví dụ, phương pháp xử lý vi khuẩn đường ruột như vi khuẩn *Pantoea* bằng cách chiếu tia cực tím hoặc tác nhân gây đột biến được sử dụng cho quá trình xử lý gây đột biến thông thường như N-metyl-N'-nitro-N-nitrosoguanidin (NTG) hoặc axit nitơ, và chọn lọc chủng có hoạt tính GCD được giảm.

Hoạt tính của GCD cũng có thể được giảm bằng cách làm giảm khả năng tổng hợp PQQ. Khả năng tổng hợp PQQ có thể được giảm bằng cách, ví dụ, loại bỏ một phần hoặc toàn bộ pqqABCDEF, là operon cần cho quá trình sinh tổng hợp PQQ (J.S. Velterop, P.W. Postma, J. Bacteriology 177(17):5088-5098 (1995)).

Trong trường hợp vi sinh vật không có hoạt tính GCD như *Escherichia coli* và vi khuẩn coryneform, glucoza được hấp thu bằng cách sử dụng gen vận chuyển được gọi là glucoza PTS (hệ glucoza phosphotransferaza) hoặc glucoza permeaza. PTS được đồng hóa vào trong các tế bào ở dạng glucoza 6-phosphat, mà kết hợp với phản ứng chuyển hóa PEP (axit phosphoenolpyruvic) thành Pyr (axit pyruvic). Glucoza 6-phosphat được chuyển hóa thành fructoza-6-phosphat, và được chuyển hóa bởi hệ thủy phân glucoza (chu trình EMP, Embden-Myerhof) để tạo ra axit pyruvic.

Mặt khác, trong trường hợp của vi sinh vật có hoạt tính GCD, glucoza trước tiên được chuyển hóa thành axit gluconic trong không gian chất bào, sau đó nó được đồng hóa bởi gluconat permeaza, và axit 6-phosphogluconic được tạo ra nhờ phản ứng phosphoryl hóa.

Axit 6-phosphogluconic được chuyển hóa bởi chu trình pentoza phosphat hoặc chu trình Entner-Doudoroff (ED) để tạo ra glyxeraldehyt 3-phosphat, axit pyruvic, và v.v..

Đã biết rằng vi sinh vật có GCD như vi khuẩn axit axetic có đặc tính chuyển hóa sacaroza đặc trưng mà trước tiên chuyển hóa một phần của glucoza thành axit gluconic trong chất bào, và sau đó đồng hóa nó. Vì hiệu suất của chu trình EMP nội bào, chu trình ED, và chu trình pentoza phosphat là khác nhau phụ thuộc vào các vi sinh vật, mong muốn rằng nếu GCD được loại bỏ để thay đổi sự chuyển hóa sacaroza, mẫu chuyển hóa hướng xuống phía dưới từ đó có thể được thay đổi.

Trong trường hợp của *Pantoea ananatis*, người ta cho rằng, ở nhiệt độ nuôi cấy thông thường, ví dụ, 34°C, glucoza không bị đồng hóa hoàn toàn bởi GCD, và phần lớn nó được đồng hóa bởi PTS. Mặt khác, nếu được nuôi cấy ở nhiệt độ cao, ví dụ, 38°C, vì nhiệt độ tối ưu của GCD là cao, nên đã mong muốn rằng hoạt tính GCD được gia tăng, và mức tiêu thụ sacarit bởi GCD gia tăng. Vì *Pantoea ananatis* không có chu trình ED, axit 6-phosphogluconic được khử hydro bằng 6-phosphogluconat dehydrogenaza, và sau đó được chuyển hóa bằng chu trình pentoza phosphat. Vì một phân tử cacbon dioxit được giải phóng từ một phân tử axit 6-phosphogluconic khi khử hydro bằng 6-phosphogluconat dehydrogenaza, mong muốn rằng nếu mức tiêu thụ sacarit bởi GCD gia tăng, mức sản sinh axit amin giảm. Hơn nữa, người ta cho rằng nếu mức tiêu thụ sacarit bởi GCD gia tăng do nuôi cấy ở nhiệt độ cao, hiệu suất của chu trình pentoza phosphat là không đủ, và sự chảy tràn các sản phẩm chuyển hóa chảy vào trong các chu trình để tạo ra các sản phẩm phụ để làm giảm năng suất axit L-amin. Đã cho rằng sự giải phóng cacbon dioxit và dòng chảy quá mức của chu trình pentoza phosphat được loại trừ bằng cách giảm hoạt tính GCD, cụ thể khi

tiến hành nuôi cấy ở nhiệt độ cao, và do vậy khả năng sản sinh axit L-amin được cải thiện.

Hơn nữa, người ta dự đoán rằng, cũng trong *Pantoea ananatis* được đưa vào với chu trình ED (đơn yêu cầu cấp patent Nhật số 2003-274988), nếu mức tiêu thụ sacarit bởi GCD gia tăng, hiệu suất axit L-amin bị giảm do sự giải phóng cacbon dioxit tại thời điểm khử hydro của axit 6-phosphogluconic bởi 6-phosphogluconat dehydrogenaza, và chu trình ED và chu trình pentoza phosphat không đạt đầy đủ hiệu suất. Do vậy, người ta cho rằng khả năng sản sinh axit L-amin được cải thiện bằng cách giảm hoạt tính GCD trong *Pantoea ananatis* được đưa vào với chu trình ED.

Vì khuẩn theo sáng chế có thể là vi khuẩn trong đó hoạt tính GCD bị giảm và trong đó hoạt tính để đồng hóa sacarit được gia tăng tiếp. Hoạt tính đồng hóa sacarit có thể được gia tăng bằng cách, ví dụ, gia tăng hoạt tính của glucoza PTS hoặc glucoza permeaza. Hơn nữa, đã biết rằng các gen vận chuyển được coi là thành viên của họ chung của gen vận chuyển chính (major facilitator superfamily-MFS, Griffith, J.K. et al, Curr. Opin. Cel Biol., 4(4);684-95 (1992)) như galactoza permeaza (Flores et al., J Mol. Microbiol. Biotechnol., 2007;13:105-116), xyloza permeaza (đơn yêu cầu cấp patent châu Âu số 1807445 A1), và arabinoza permeaza cũng có hoạt tính đồng hóa glucoza v.v.. Do vậy, cũng trong vi khuẩn trong đó hoạt tính GCD bị giảm, bằng cách gia tăng hoạt tính bất kỳ của các gen vận chuyển này, sự đồng hóa sacarit như glucoza cũng được gia tăng, và khả năng sản sinh axit L-amin được cải thiện.

<2> Phương pháp sản xuất axit L-amin

Bằng cách nuôi cấy vi sinh vật theo sáng chế trong môi trường để sản sinh và tích tụ axit L-amin trong môi trường và thu gom axit L-amin từ môi

trường này, axit L-amin có thể được sản xuất.

Về môi trường được sử dụng cho quá trình nuôi cấy, môi trường điển hình chứa nguồn cacbon, nguồn nitơ, và các muối khoáng cũng như các chất sinh dưỡng vết hữu cơ như các axit amin và các vitamin nếu cần có thể được sử dụng. Có thể sử dụng môi trường tổng hợp hoặc môi trường tự nhiên. Bất kỳ kiểu nguồn cacbon và nguồn nitơ có thể được sử dụng miễn là chúng có thể được sử dụng bởi chủng được chọn cần được nuôi cấy.

Các đường như glucoza, glyxerol, fructoza, sucroza, maltoza, manosa, galactoza, dịch thủy phân tinh bột và rỉ đường có thể được sử dụng làm nguồn cacbon. Ngoài ra, các axit hữu cơ như axit axetic và axit xitric, và các rượu như ethanol cũng có thể được sử dụng riêng biệt hoặc kết hợp với các nguồn cacbon khác. Amoniac, các muối amoni như amoni sulfat, amoni carbonat, amoni clorua, amoni phosphat và amoni axetat, các muối của axit nitric và v.v. có thể được sử dụng làm nguồn nitơ. Các axit amin, vitamin, axit béo, axit nucleic, các thành phần chứa các chất này như pepton, axit casamino, dịch chiết nấm men, sản phẩm thủy phân protein đậu tương và v.v. có thể được sử dụng làm các chất sinh dưỡng vết hữu cơ. Khi chủng đột biến dinh dưỡng thụ động cần axit amin hoặc các chất tương tự cho sự sinh trưởng của nó được sử dụng, chất dinh dưỡng được yêu cầu có thể được bổ sung.

Cụ thể, khi môi trường lỏng được điều chế để thỏa mãn điều kiện kết tủa axit L-glutamic được sử dụng, sự bổ sung axit pantothenic vào môi trường tạo ra sự kết tủa axit L-glutamic hiệu quả hơn (WO2004/111258). Về các muối vô cơ, các muối của axit phosphoric, các muối magie, các muối canxi, các muối sắt, muối mangan và v.v. có thể được sử dụng.

Quá trình nuôi cấy có thể được thực hiện dưới dạng nuôi cấy hiếu khí,

trong khi nhiệt độ lên men được điều chỉnh nằm trong khoảng từ 20 đến 45°C, và độ pH là từ 3 đến 9. Khi độ pH giảm trong khi nuôi cấy, canxi carbonat có thể được bổ sung, hoặc dịch nuôi cấy được trung hòa bằng chất kiềm như khí amoniac. Axit L-amin đích được tích tụ trong môi trường nuôi cấy sau, ví dụ, 10 đến 120 giờ nuôi cấy dưới các điều kiện như nêu trên.

Mặc dù axit L-amin có thể được sản xuất một cách có hiệu quả bằng cách nuôi cấy ở nhiệt độ thích hợp cho sự sinh trưởng của vi khuẩn, hiệu quả trở nên cao hơn nhiều khi quá trình nuôi cấy được thực hiện ở nhiệt độ cao. Ví dụ, trong trường hợp của vi khuẩn *Pantoea* như *Pantoea ananatis*, nhiệt độ khoảng 34°C là được ưu tiên cho sự sinh trưởng của vi khuẩn. Tuy nhiên, mặc dù vi khuẩn trong đó GCD hoạt tính bị giảm có khả năng sản sinh axit L-amin cao hơn so với của chủng không được cải biến ngay cả ở nhiệt độ nuôi cấy này, nó thể hiện khả năng sản sinh axit L-amin được cải thiện hơn nữa ở, ví dụ, 36°C hoặc 38°C.

Hơn nữa, quá trình nuôi cấy có thể được thực hiện bằng cách kết tủa axit L-glutamic trong môi trường nhờ sử dụng môi trường lỏng được điều chỉnh để thỏa mãn điều kiện trong đó axit L-glutamic được kết tủa làm môi trường nuôi cấy. Các ví dụ về các điều kiện trong đó axit L-glutamic được kết tủa bao gồm, ví dụ, độ pH = 5,0 đến 4,0, tốt hơn là độ pH = 4,5 đến 4,0, tốt hơn nữa là độ pH = 4,3 đến 4,0, tốt nhất là độ pH = 4,0.

Khi axit L-glutamic được kết tủa trong môi trường, sự bổ sung sơ bộ các tinh thể của axit L-glutamic hoặc L-lysin làm mầm tinh thể có thể làm cho sự kết tinh hiệu quả hơn (patent châu Âu số 1233069, đơn yêu cầu cấp patent châu Âu số 1624069).

Sự thu gom axit L-amin từ dịch nuôi cấy sau khi nuôi cấy có thể được

thực hiện bằng phương pháp thu gom đã biết. Ví dụ, sau khi loại bỏ tế bào khỏi môi trường nuôi cấy, axit L-amin có thể được thu gom bằng cách cô môi trường để kết tinh axit L-amin, sắc kí trao đổi ion, hoặc các phương pháp tương tự khác. Khi quá trình nuôi cấy được thực hiện dưới các điều kiện sao cho axit L-glutamic được kết tủa, axit L-glutamic mà kết tủa trong môi trường có thể được thu gom bằng cách li tâm hoặc lọc. Trong trường hợp này, axit L-glutamic mà hòa tan trong môi trường có thể được kết tủa và sau đó được tách cùng với axit L-glutamic đã kết tủa.

Khi axit amin bazơ được tạo ra, một phương pháp có thể được sử dụng trong đó độ pH của môi trường được điều chỉnh nằm trong khoảng từ 6,5 đến 9,0 và độ pH của môi trường sau khi kết thúc quá trình nuôi cấy được điều chỉnh nằm trong khoảng từ 7,2 đến 9,0. Hơn nữa, áp suất trong thùng lên men có thể được điều chỉnh trong khi lên men là dương, hoặc cacbon dioxit hoặc hỗn hợp khí chứa cacbon dioxit có thể được bổ sung vào môi trường sao cho trước khi các ion bicarbonat và/hoặc các ion carbonat có mặt ở nồng độ ít nhất 2g/L trong môi trường nuôi cấy trong khi nuôi cấy, và các ion bicarbonat và/hoặc carbonat này đóng vai trò là ion trái dấu đối với các cation chủ yếu bao gồm axit amin bazơ, và axit amin bazơ đích sau đó được thu gom (xem đơn yêu cầu cấp patent Nhật số 2002-065287, đơn yêu cầu cấp patent Mỹ số 2002/025564).

Axit L-amin được thu gom có thể chứa tế bào vi khuẩn, thành phần môi trường, độ ẩm, và các chất chuyển hóa sản phẩm phụ của vi khuẩn ngoài axit L-amin đích. Độ tinh khiết của axit L-amin được thu gom là bằng hoặc cao hơn 50%, tốt hơn là bằng hoặc cao hơn 85%, tốt nhất là bằng hoặc cao hơn 95% (patent Nhật số 1214636, các patent Mỹ số 5,431,933, 4,956,471, 4,777,051, 4,946,654, 5,840,358, 6,238,714, đơn yêu cầu cấp patent Mỹ số 2005/0025878).

Khi L-xystein được tạo ra bằng phương pháp theo sáng chế, L-xystein

thu được có thể được sử dụng để sản xuất các dẫn xuất L-xystein. Các dẫn xuất xystein bao gồm metylxystein, etylxystein, carboxystein, sulfoxystein, axetyl xystein, và v.v..

Hơn nữa, khi dẫn xuất thiazolidin của L-xystein được tích tụ trong môi trường, L-xystein có thể được hình thành bằng cách thu gom dẫn xuất thiazolidin từ môi trường để phá vỡ sự cân bằng phản ứng giữa dẫn xuất thiazolidin và L-xystein sao cho L-xystein được tạo ra vượt trội. Hơn nữa, khi S-sulfoxystein được tích tụ trong môi trường, nó có thể được chuyển hóa thành L-xystein nhờ phản ứng với chất khử như dithiothreitol.

Ví dụ thực hiện sáng chế

Sau đây, sáng chế sẽ được giải thích chi tiết hơn nhờ tham khảo các ví dụ mà không gây ra sự giới hạn đối với sáng chế.

Ví dụ tham khảo 1: Tạo ra chủng *Pantoea ananatis* mà kháng gen λ Red

Để loại bỏ gen trong *Pantoea ananatis*, chủng nhận được tạo ra để thực hiện phương pháp gọi là "sự kết hợp bức xạ đỏ" hoặc "sự kết hợp được gây ra bởi ánh sáng đỏ" (Proc. Natl. Acad. Sci. USA, 97, 6640-6645 (2000)) theo cách có hiệu quả cao.

Trước tiên, plasmit hỗ trợ RSF-Red-TER mới biểu hiện các gen gam, bet và *exo* của λ (sau đây được gọi là "các gen λ Red") được tạo ra (Fig. 1). Chi tiết về plasmit này sẽ được đề cập trong Ví dụ tham khảo 2.

Plasmit này có thể được sử dụng ở nhiều loại vật chủ có các nền tảng di truyền khác nhau. Điều này là bởi vì 1) plasmit này có replicon của plasmit phổ vật chủ rộng RSF1010 (Scholz, et al., 1989; Buchanan-Wollaston et al., 1987), mà được duy trì ổn định bởi nhiều kiểu vi khuẩn gram âm và gram dương, và thậm chí các tế bào thực vật, 2) các gen λ Red, các gen gam, bet và *exo*, dưới sự

điều chỉnh của đoạn khởi đầu PlacUV5, mà được xác nhận bởi các ARN polymeraza của nhiều loại vi khuẩn (ví dụ, Brunschwig, E. và Darzins, A., *Gene*, 111, 1, 35-41 (1992); Dehio, M. et al, *Gene*, 215, 2, 223-229 (1998)), và 3) yếu tố tự điều chỉnh $P_{lacUV5-lacI}$ và các gen kết thúc phiên mã không phụ thuộc ρ ($TrnB$) của operon *rrnB* của *Escherichia coli* thấp hơn mức biểu hiện cơ sở của các gen λ Red (Skorokhodova, A. Yu et al, *Biotekhnologiya (Rus)*, 5, 3-21 (2004)). Hơn nữa, plasmit RSF-Red-TER chứa gen levansucraza (*sacB*), và bằng cách sử dụng gen này, plasmit có thể được thu gom từ các tế bào trong môi trường chứa sucroza.

Trong *Escherichia coli*, tần xuất kết hợp của đoạn ADN được tạo ra bằng PCR cùng với các vùng sườn ngắn được tạo ra bởi plasmit RSF-Red-TER là cao ngang với tần suất thu được khi sử dụng plasmit hỗ trợ pKD46 (Datsenko, K.A., Wanner, B.L., *Proc. Natl. Acad. Sci. USA*, 97, 6640-6645 (2000)). Tuy nhiên, sự biểu hiện của các gen λ Red gây độc cho *Pantoea ananatis*. Các tế bào được biến nạp với plasmit hỗ trợ RSF-Red-TER phát triển rất chậm trong môi trường LB chứa IPTG (isopropyl- β -D-thiogalactopyranosit, 1mM) và kháng thể thích hợp (25 μ g/ml cloramphenicol hoặc 40 μ g/ml kanamycin), và hiệu suất tái tổ hợp do λ Red gây ra là rất thấp (10^{-8}), nếu được quan sát toàn bộ.

Chủng biến dị của *Pantoea ananatis* kháng lại sự biểu hiện của cả ba gen λ Red được chọn. Vì mục đích này, plasmit RSF-Red-TER được đưa vào trong chủng *Pantoea ananatis* SC17 (patent Mỹ số 6,596,517) bằng phương pháp xung điện. Sau 18 giờ nuôi cấy, thu được khoảng 10^6 thể biến nạp, và trong số đó, 10 dòng tạo ra khuẩn lạc có kích cỡ lớn, và tất cả các dòng còn lại tạo ra các khuẩn lạc cực nhỏ. Sau 18 giờ nuôi cấy, các khuẩn lạc lớn là khoảng 2mm, và các khuẩn lạc nhỏ là khoảng 0,2mm. Ngược lại, các khuẩn lạc nhỏ không sinh trưởng ngay cả khi được nuôi cấy đến 24 giờ, các khuẩn lạc lớn tiếp tục phát

triển. Một trong số các chủng đột biến *Pantoea ananatis* khuẩn lạc lớn kháng lại sự biểu hiện của cả ba gen λ Red (gam, bet, và exo) được sử dụng để phân tích tiếp.

ADN plasmid RSF-Red-TER được tách từ một dòng trong số các dòng khuẩn lạc lớn, và từ vài dòng khuẩn lạc nhỏ, và được biến nạp vào trong *Escherichia coli* MG1655 để đánh giá khả năng của plasmid để tổng hợp gen Red hoạt hóa. Nhờ thử nghiệm kiểm soát cho quá trình kết hợp phụ thuộc ánh sáng đỏ trong thể biến nạp thu được, đã chứng minh được rằng chỉ plasmid được tách từ dòng khuẩn lạc lớn gây ra sự biểu hiện của các gen λ Red cần cho sự kết hợp phụ thuộc ánh sáng đỏ. Để đánh giá sự kết hợp được gây ra bởi ánh sáng đỏ xảy ra trong dòng khuẩn lạc lớn được chọn, phương pháp xung điện được thực hiện bằng cách sử dụng đoạn ADN thẳng được tạo ra bởi PCR. Đoạn này được tạo ra sao cho nó chứa gen đánh dấu Km^R và vùng bên sườn có kích cỡ 40bp tương đồng với gen hisD. Đoạn này được kết hợp vào trong gen hisD của *Pantoea ananatis* ở vị trí xác nhận SmaI. Hai dòng khuẩn lạc nhỏ được sử dụng làm đối chứng. Trình tự nucleotit của gen hisD của *Pantoea ananatis* được thể hiện trong SEQ ID NO: 3. Để tiến hành PCR, các oligonucleotit của các SEQ ID NO: 4 và 5 được sử dụng làm đoạn mồi, và plasmid pMW118-(λ attL- Km^r - λ attR) được sử dụng làm khuôn mẫu. Hai dòng khuẩn lạc nhỏ mà không kháng các gen λ Red được sử dụng làm đối chứng. Sự tạo ra plasmid pMW118-(λ attL- Km^r - λ attR) sẽ được giải thích chi tiết trong Ví dụ tham khảo 3.

Plasmid RSF-Red-TER có thể cảm ứng sự biểu hiện của các gen Red nhờ gen lacI được mang trên plasmid. Hai kiểu điều kiện cảm ứng được đánh giá. Ở nhóm thứ nhất, IPTG (1mM) được bổ sung 1 giờ trước phương pháp xung điện, và ở nhóm thứ hai, IPTG được bổ sung ở thời điểm bắt đầu nuôi cấy có thể tạo ra các tế bào trong đó phương pháp xung điện. Tốc độ sinh trưởng của các tế bào

mang RSF-Red-TER thu được từ dòng khuẩn lạc lớn không thấp hơn so với tốc độ sinh trưởng của chủng không có plasmit đó. Sự bổ sung IPTG chỉ hơi làm giảm tốc độ sinh trưởng của các quá trình nuôi cấy này. Mặt khác, thế hệ sau của dòng khuẩn lạc nhỏ phát triển cực thấp ngay cả khi không bổ sung IPTG, và sau khi cảm ứng, sự phát triển về cơ bản bị dừng lại. Sau phương pháp xung điện các tế bào thế hệ sau của dòng khuẩn lạc lớn, nhiều dòng Km^R đã phát triển (18 dòng sau thời gian phản ứng ngắn, và khoảng 100 dòng sau thời gian cảm ứng kéo dài). Tất cả 100 dòng mà được đánh giá có kiểu hình His^- , và khoảng 20 dòng được xác nhận bằng PCR để có cấu trúc mong muốn của nhiễm sắc thể trong các tế bào. Mặt khác, ngay cả khi phương pháp xung điện được thực hiện với thế hệ sau của dòng khuẩn lạc nhỏ, không thu được chủng được kết hợp.

Dòng khuẩn lạc lớn thu được được phát triển trên tấm chứa 7% sucroza để loại bỏ plasmit, và được biến nạp với RSF-Red-TER. Chủng không có plasmit được gọi là SC17(0).

Tất cả dòng mà phát triển sau quá trình tái biến nạp nêu trên thể hiện kích cỡ khuẩn lạc lớn giống như dòng chủng cha mẹ SC17(0). Thử nghiệm kết hợp được gây ra bởi ánh sáng đỏ được thực hiện ở chủng SC17(0) được biến nạp với plasmit RSF-Red-TER. Ba thể biến nạp độc lập được đánh giá bằng cách sử dụng đoạn ADN giống như ADN được sử dụng cho thử nghiệm nêu trên. Thời gian cảm ứng ngắn (1 giờ trước phương pháp xung điện) được sử dụng. Dòng Km^R vượt quá mười dòng đã phát triển trong từng thử nghiệm. Tất cả các dòng được đánh giá có kiểu hình His^- . Theo cách này, chủng đột biến được kí hiệu SC17(0) kháng lại sự biểu hiện của các gen λ Red được chọn. Chủng này có thể được sử dụng làm chủng nhận thích hợp cho sự kết hợp phụ thuộc ánh sáng đỏ vào trong nhiễm sắc thể *Pantoea ananatis*.

Ví dụ tham khảo 2: Tạo ra plasmit hỗ trợ RSF-Red-TER

Sơ đồ tạo ra plasmit hỗ trợ RSF-Red-TER được thể hiện trên Fig. 2.

Về bước thứ nhất trong quy trình tạo ra, vật truyền RSFsacBPlacMCS được tạo ra. Vì mục đích này, các đoạn ADN chứa gen cat của plasmit pACYC184 và vùng cấu trúc của gen sacB của *Bacillus subtilis* được khuếch đại bằng PCR sử dụng các oligonucleotit của các SEQ ID NO: 6 và 7, và 8 và 9. Các oligonucleotit này chứa các vị trí enzym giới hạn BglII, SacI, XbaI và BamHI, mà được yêu cầu và thuận tiện cho sự tách dòng tiếp theo, ở các vùng kết thúc 5'. Đoạn sacB thu được có kích cỡ 1,5kb được tách dòng vào trong vật truyền pMW119-P_{lac}lacI thu được ở trên ở vị trí XbaI-BamHI. Vật truyền này được tạo ra theo cách tương tự như được mô tả cho vật truyền pMW118-P_{lac}lacI (Skorokhodova, A. Yu et al, Biotekhnologiya (Rus), 5, 3-21 (2004)). Tuy nhiên, vật truyền này chứa gốc liên kết cao phân tử thu được từ pMW219 thay vì plasmit pMW218.

Sau đó, đoạn cat nêu trên có kích cỡ 1,0kb được xử lý với BglII và SacI, và được tách dòng trong plasmit RSF-P_{lac}lacIsacB thu được ở bước trên ở vị trí BamHI-SacI. Plasmit pMW-P_{lac}lacIsacBcat thu được có đoạn PlacUV5-lacI-sacB-cat. Để tách dòng phụ cho đoạn này vào trong vật truyền RSF1010, pMW-P_{lac}lacIsacBcat được cắt bằng BglII, được làm bằng đầu bằng đoạn ADN polymeraza I Klenow, và cắt liên tiếp với SacI. Đoạn BglII-SacI 3,8kb của plasmit pMWP_{lac}lacIsacBcat được rửa giải từ 1% gel agarosa, và được ghép nối với vật truyền RSF1010 mà đã được xử lý bằng PstI và SacI. *Escherichia coli* TG1 được biến nạp với hỗn hợp ghép nối, và được đặt trên môi trường L (môi trường chứa 10g Bacto trypton, 5g dịch chiết nấm men, 5g NaCl, và 15g thạch trong 1L nước tinh khiết, độ pH = 7,0) chứa cloramphenicol (50mg/L). Các plasmit được tách từ các dòng đã phát triển được phân tích với các enzym giới hạn để thu plasmit RSFsacB. Để tạo ra vật truyền RSFsacBP_{lac}MCS, đoạn ADN

chứa đoạn khởi đầu P_{lacUV5} được khuếch đại bởi PCR sử dụng các oligonucleotit của các SEQ ID NO: 10 và 11 làm đoạn môi và plasmit pMW119- P_{lac} I làm khuôn mẫu. Đoạn thu được có kích cỡ 146bp được cắt bằng SacI và NotI, và được ghép nối với đoạn lớn SacI-NotI của plasmit RSFsacB. Sau đó, bằng cách tiến hành PCR sử dụng các oligonucleotit của các SEQ ID NO: 12 và 13 làm đoạn môi, và plasmit pKD46 (Datsenko, K.A., Wanner, B.L., Proc. Natl. Acad. Sci. USA, 97, 6640-6645 (2000)) làm khuôn mẫu, đoạn ADN có kích cỡ 2,3kb chứa các gen λ Red $\alpha\beta\gamma$ và gen kết thúc phiên mã tL3 được khuếch đại. Đoạn thu được được tách dòng trong vật truyền RSFsacB P_{lac} MCS ở vị trí PvuI-NotI. Theo cách này, plasmit RSFRed được tạo ra.

Để loại bỏ khung đọc thông qua sự phiên mã của các gen Red, gen kết thúc sự phiên mã phụ thuộc ρ của operon *rrnB* của *Escherichia coli* được đính ở vị trí giữa gen cat và đoạn khởi đầu P_{lacUV5} . Vì mục đích này, đoạn ADN chứa đoạn khởi đầu P_{lacUV5} và gen kết thúc *TrnB* được khuếch đại bằng kỹ thuật PCR sử dụng các oligonucleotit của các SEQ ID NO: 14 và 11 làm đoạn môi và nhiễm sắc thể của *Escherichia coli* BW3350 làm khuôn mẫu. Các đoạn thu được này được xử lý với KpnI và được ghép nối. Sau đó, đoạn có kích cỡ 0,5kb chứa cả P_{lacUV5} và *TrnB* được khuếch đại bằng kỹ thuật PCR sử dụng các oligonucleotit của các SEQ ID NO: 11 và 15 làm đoạn môi. Đoạn ADN thu được được cắt bằng EcoRI, được làm bằng đầu bằng cách xử lý với đoạn ADN polymeraza I Klenow, được cắt bằng BamHI, và được ghép nối với đoạn lớn Ecl136II-BamHI của vật truyền RSFsacB P_{lac} MCS. Plasmit thu được được gọi là RSF-Red-TER.

Ví dụ tham khảo 3: Tạo ra plasmit pMW118-(λ attL-Km^r- λ attR)

Plasmit pMW118-(λ attL-Km^r- λ attR) được tạo ra từ plasmit pMW118-attL-Tc-attR (WO2005/010175) bằng cách thay thế gen đánh dấu kháng

tetracyclin với gen kháng kanamycin của plasmit pUC4K. Vì lý do đó, đoạn lớn EcoRI–HindIII từ plasmit pMW118-attL-Tc-attR được gắn với hai đoạn từ plasmit pUC4K: đoạn HindIII–PstI (676bp) và đoạn EcoRI–HindIII (585bp). pMW118-attL-Tc-attR cơ sở thu được bằng cách gắn bốn đoạn sau đây.

1) Đoạn BglII-EcoRI (114bp) bao gồm attL (SEQ ID NO: 18) mà thu được bằng cách khuếch đại PCR cho vùng tương ứng với attL của nhiễm sắc thể của *Escherichia coli* W3350 (chứa thể thực khuẩn λ) sử dụng các đoạn mồi P1 và P2 (các SEQ ID NO: 16 và 17) (các đoạn mồi này có các vị trí xác nhận phụ cho BglII và EcoRI).

2) Đoạn PstI-HindIII (182bp) bao gồm attR (SEQ ID NO: 21) mà thu được bằng cách khuếch đại PCR cho vùng tương ứng với attR của nhiễm sắc thể của *Escherichia coli* W3350 (chứa thể thực khuẩn λ) sử dụng các đoạn mồi P3 và P2 (các SEQ ID NO: 19 và 20) (các đoạn mồi này có các vị trí xác nhận phụ cho PstI và HindIII).

3) Đoạn lớn BglII-HindIII (3916bp) của pMW118-ter_rrnB. Plasmit pMW118-ter_rrnB thu được bằng cách gắn ba đoạn ADN sau đây:

Đoạn ADN lớn (2359bp) gồm đoạn AatII-EcoRI của pMW118 mà thu được bằng cách cắt pMW118 bằng EcoRI, xử lý bằng đoạn ADN polymeraza I Klenow, và sau đó cắt bằng AatII;

Đoạn AatII-BglII nhỏ (1194bp) của pUC19 gồm gen bla kháng ampicilin (Ap^R), mà thu được bằng cách khuếch đại PCR của vùng tương ứng của plasmit pUC19 sử dụng các đoạn mồi P5 và P6 (các SEQ ID NO: 22 và 23) (các đoạn mồi này có các vị trí xác nhận phụ cho PstI, AatII và BglII);

Đoạn BglII-PstI nhỏ (363bp) của gen kết thúc phiên mã ter_rrnB, mà thu được bằng cách khuếch đại PCR của vùng tương ứng của nhiễm sắc thể

Escherichia coli MG1655 sử dụng các đoạn mồi P7 và P8 (các SEQ ID NO: 24 và 25) (các đoạn mồi này có các vị trí xác nhận phụ cho PstI, BglII và PstI).

4) Đoạn EcoRI-PstI nhỏ (1388bp) (SEQ ID NO: 26) của pML-Tc-ter_thrL gồm gen kháng tetracyclin và gen kết thúc sự phiên mã ter_thrL; plasmit pML-Tc-ter_thrL thu được bằng hai bước sau đây:

- plasmit pML-ter_thrL thu được bằng cách cắt plasmit pML-MCS (Mashko, S.V. et al., Biotekhnologiya (in Russian), 2001, no. 5, 3-20) bằng XbaI và BamHI, sau đó gắn đoạn lớn (3342bp) với đoạn XbaI-BamHI (68bp) mang gen kết thúc ter_thrL thu được bằng cách khuếch đại PCR vùng tương ứng của nhiễm sắc thể *Escherichia coli* MG1655 sử dụng các đoạn mồi P9 và P10 (các SEQ ID NO: 27 và 28) (các đoạn mồi này có các vị trí xác nhận phụ cho PstI, XbaI và BamHI);

- plasmit pML-Tc-ter_thrL thu được bằng cách cắt plasmit pML-ter_thrL bằng KpnI và XbaI sau đó xử lý với đoạn Klenow của ADN polymeraza I và được ghép nối với đoạn EcoRI-Van91I nhỏ (1317bp) của pBR322 gồm gen kháng tetracyclin (pBR322 được cắt bằng EcoRI và Van91I và sau đó được xử lý bằng đoạn Klenow của ADN polymeraza I).

Ví dụ tham khảo 4: Tạo ra plasmit RSFPPG sản sinh axit glutamic

Plasmit RSFPPG (WO2008/020654) được tạo ra trong đó các gen hệ sinh tổng hợp axit L-glutamic, gen prpC (công bố quốc tế WO2006/051660), gen ppc và gen gdhA (đơn yêu cầu cấp patent châu Âu số 0999282) được khuếch đại.

Đoạn mồi 1 (SEQ ID NO: 29) và đoạn mồi 2 (SEQ ID NO: 30) để khuếch đại một phần của RSFCPG (đơn yêu cầu cấp patent châu Âu số 1233068) khác với ORF của gen gltA được tạo ra. Bằng cách sử dụng các đoạn

môi này và RSFCPG làm khuôn mẫu, PCR được thực hiện để thu đoạn có kích cỡ khoảng 14,9kb. Về prpC, kỹ thuật PCR được thực hiện bằng cách sử dụng đoạn môi 3 (SEQ ID NO: 31) và đoạn môi 4 (SEQ ID NO: 32) và ADN nhiễm sắc thể của chủng *E. coli* W3110 làm khuôn mẫu để thu đoạn có kích cỡ khoảng 1,2kb. Cả sản phẩm PCR được xử lý với BglII và KpnI, được gắn, và sau đó được sử dụng để biến nạp chủng *E. coli* JM109. Tất cả các khuẩn lạc mà đã sinh trưởng được thu gom, và các plasmit được tách từ các khuẩn lạc dưới dạng hỗn hợp. Chủng *E. coli* ME8330, là chủng khuyết xitrat syntaza (CS), được biến nạp với hỗn hợp plasmit, và huyền phù tế bào được phủ lên môi trường tối thiểu M9 (5g glucoza, 2mM magie sulfat, 3g monokali phosphat, 0,5g natri clorua, 1g amoni clorua và 6g dinatri phosphat trong 1L nước tinh khiết) chứa 50mg/L uraxil và 5mg/L thiamin HCl. Tất cả các khuẩn lạc mà đã phát triển được thu gom, các plasmit được chiết dưới dạng hỗn hợp, và chủng NP106, là chủng sản sinh axit L-glutamic của *P. ananatis*, được biến nạp với hỗn hợp plasmit. Dòng mà đã sinh trưởng được nuôi cấy trong các ống nghiệm dưới các điều kiện trung tính, và chủng thể hiện hiệu suất axit L-glutamic so với hiệu suất của chủng G106S được gọi là NA1. Plasmit được tách từ chủng này và được gọi là RSFPPG để gia tăng prpC, gdh, và ppc. Plasmit RSFPPG này được đưa vào trong chủng *Pantoea ananatis* NP106, là chủng sản sinh axit L-glutamic, để tạo ra chủng sản sinh axit L-glutamic, NP106/RSFPPG (chủng này được gọi là “chủng NA1”).

Chủng NP106 thu được như sau. Chủng *Pantoea ananatis* AJ13601 nêu trên được nuôi cấy qua đêm ở 34°C trong môi trường lỏng thu được bằng cách bổ sung thành phần môi trường tối thiểu (5g/L glucoza, 2mM magie sulfat, 3g/L monokali phosphat, 0,5g/L natri clorua, 1g/L amoni clorua, và 6g/L dinatri phosphat) vào môi trường L (10g/L Bacto trypton, 5g/L dịch chiết nấm men,

5g/L NaCl, độ pH = 7,0) (sau đây được gọi là "môi trường LBGM9") có khuấy, môi trường được pha loãng sao cho từ 100 đến 200 khuẩn lạc sẽ sinh trưởng trên mỗi đĩa, và được phủ lên đĩa LBGM9 chứa 12,5mg/L tetracyclin. Các khuẩn lạc mà đã xuất hiện được sinh sản trên đĩa LBGM9 chứa 12,5mg/L tetracyclin và 25mg/L chloramphenicol, và chủng nhạy chloramphenicol được chọn để thu chủng từ đó pSTVCB được loại bỏ, mà được gọi là G106S. Chủng G106S mà được nuôi cấy tiếp qua đêm ở 34°C trong môi trường lỏng LBGM9 có khuấy, môi trường được pha loãng sao cho từ 100 đến 200 khuẩn lạc xuất hiện trên mỗi đĩa, và được phủ lên đĩa LBGM9 mà không có thuốc. Các khuẩn lạc đã sinh trưởng được sinh sản trên đĩa LBGM9 chứa 12,5mg/L tetracyclin và đĩa LBGM9 mà không có thuốc, và chủng nhạy tetracyclin được chọn để thu chủng từ đó RSFCPG được loại bỏ, mà được gọi là NP106. NP106 thu được như nêu trên không chứa các plasmit RSFCPG và pSTVCB, mà được mang bởi chủng AJ13601.

Chủng G106S là chủng thu được bằng cách chỉ loại bỏ pSTVCB khỏi chủng AJ13601 theo cách tương tự.

Ví dụ 1: Sản xuất axit L-glutamic bằng cách sử dụng chủng khuyết gen *gcd*

(1) Tạo ra chủng khuyết gen *gcd*

Hai đoạn môi ADN tổng hợp được thể hiện trong các SEQ ID NO: 33 và 34 được tổng hợp bằng phương pháp thông thường.

Đoạn môi được thể hiện trong SEQ ID NO: 33 có cấu trúc mà dòng hướng lên phía trên trình tự tương đồng của gen *gcd* của *Pantoea ananatis* là trình tự tương đồng ở đầu 5' của λ attL-Km^r- λ attR. Đoạn môi của SEQ ID NO: 34 có cấu trúc mà dòng hướng xuống phía dưới trình tự bổ trợ của gen *gcd* của *Pantoea ananatis* sau đó là trình tự bổ trợ ở đầu 3' của λ attL-Km^r- λ attR. Nhờ

tiến hành PCR sử dụng các đoạn mồi này và pMW118-(λ attL-Km^r- λ attR) làm khuôn mẫu, đoạn có kích cỡ khoảng 1,5kbp trong đó dòng hướng lên phía trên trình tự tương đồng của gen *gcd* được gắn với đầu 5' của trình tự λ attL-Km^r- λ attR, và dòng hướng xuống phía dưới trình tự tương đồng của gen *gcd* được gắn với đầu 3' của trình tự của λ attL-Km^r- λ attR được khuếch đại.

Đoạn PCR được tinh lọc và được sử dụng để kết hợp phụ thuộc λ vào trong nhiễm sắc thể *Pantoea ananatis*. Plasmid hỗ trợ RSF-Red-TER được sử dụng làm chất mang của các gen Red thể thực khuẩn λ . Để thu các tế bào khả biến điện tử của *Pantoea ananatis*, chủng SC17(0) được biến nạp với plasmid RSF-Red-Ter, và được nuôi cấy qua đêm ở 34°C trong môi trường LB chứa 50 μ g/ml cloramphenicol. Sau đó, dịch nuôi cấy được pha loãng 100 lần với môi trường LB mới chứa 50 μ g/ml cloramphenicol, và các tế bào sinh trưởng ở 34°C dưới điều kiện hiếu khí cho đến khi OD₆₀₀ đạt giá trị 0,3. Sau đó, 1mM IPTG được bổ sung, và tiếp tục nuôi cấy cho đến khi OD₆₀₀ đạt giá trị 0,7. Các tế bào trong 10mL dịch nuôi cấy được rửa 3 lần với thể tích nước khử ion tương đương, và các tế bào được tạo huyền phù trong 80 μ l glycerol lạnh 10%. Sản phẩm PCR nêu trên được hòa tan trong 10 μ l nước được khử ion, và từ 100 đến 200ng của đoạn PCR được bổ sung vào huyền phù tế bào. Phương pháp xung điện được tiến hành bằng cách sử dụng thiết bị xung điện vi khuẩn (BioRad, United States, Catalog number 165-2089, Version 2-89). Các thông số xung được sử dụng là cường độ điện trường 18 kV/cm, và thời gian xung 5 mili giây.

Sau khi xung điện, 1ml môi trường LB được bổ sung với glucoza (0,5%) được bổ sung ngay vào huyền phù tế bào. Sau đó, các tế bào được sinh trưởng ở 34°C trong 2 giờ dưới điều kiện hiếu khí, và được chọn trên môi trường L (10g Bacto trypton, 5g dịch chiết nấm men, 5g NaCl, và 15g thạch trong 1L nước tinh khiết, độ pH = 7,0) chứa 40mg/L kanamycin để thu khoảng 20 khuẩn lạc làm thể

biến nạp. Sự xen đoạn gen kháng kanamycin ở vùng gen *gcd* được xác nhận bằng kỹ thuật PCR sử dụng hai đoạn mồi ADN tổng hợp được thể hiện trong các SEQ ID NO: 35 và 36, và chủng trong đó sự xen đoạn được xác nhận được kí hiệu là SC17(0):: Δ gcd. ADN hệ gen được tách từ chủng này, và được sử dụng để biến nạp chủng NA1 bằng phương pháp xung điện.

Chủng NA1 trong đó ADN hệ gen của SC17(0):: Δ gcd được đưa vào được chọn trên đĩa môi trường LBGM9 được bổ sung với 40mg/L kanamycin, 12,5mg/L tetracyclin hydroclorua, và 15g/L thạch. Kết quả là, khoảng 20 khuẩn lạc thu được dưới dạng thể biến nạp. Trong tất cả các chủng này, đoạn λ attL-Km^r- λ attR được dính trong vùng gen *gcd*, và một dòng trong số chúng được chọn và được gọi là NA1:: Δ gcd.

(2) Đánh giá khả năng sản sinh axit L-glutamic của chủng khuyết gen *gcd*

Để đánh giá tác dụng của sự khuyết gen *gcd* đối với sự sản sinh axit L-glutamic, môi trường sản sinh axit L-glutamic được thực hiện bằng cách sử dụng các chủng NA1:: Δ gcd và NA1.

Quá trình nuôi cấy được thực hiện trong hai bước: nuôi cấy mầm để cho phép hình thành tế bào, và nuôi cấy chính để sản sinh axit L-glutamic.

Quá trình nuôi cấy mầm được thực hiện với thành phần môi trường.

Thành phần môi trường nuôi cấy mầm:

Sucroza	50g/L
MgSO ₄ ·7H ₂ O	0,4g/L
GD113 (chống tạo bọt)	0,1mL/L
(NH ₄) ₂ SO ₄	4,0g/L
KH ₂ PO ₄	2,0g/L
Dịch chiết nấm men	4,0g/L
FeSO ₄ ·7H ₂ O	0,01g/L
MnSO ₄ ·5H ₂ O	0,01g/L
Axit xitric	0,02g/L

L-lysin hydroclorua	0,4g/L
DL-methionin	0,4g/L
Axit ϵ -diaminopimelic	0,4g/L
Canxi pantothenat	18mg/L
Tetraxyclin hydroclorua	12,5mg/L

Môi trường được thanh trùng bằng hơi nước ở 120°C trong 20 phút.

Các chủng NA1:: Δ gcd và NA1, mỗi chủng được nuôi cấy sơ bộ trên đĩa môi trường LBGM9 được bổ sung với 12,5mg/L tetraxyclin và 15g/L thạch, và các tế bào tương ứng với một đĩa được cấy vào trong 300mL môi trường nuôi cấy mầm có các thành phần nêu trên có trong bình nhỏ có thể tích 1L, và quá trình khuấy được kiểm soát ở 34°C và độ pH = 6,0 trong khoảng 12 giờ sao cho đạt được sự thông khí 1/1 vvm và nồng độ oxy bằng hoặc cao hơn 3%. Trong khi nuôi cấy, độ pH được kiểm soát để đạt 6,0 với sự bổ sung khí amoniac. Trong khi nuôi cấy, độ pH được kiểm soát đến giá trị 6,0 bằng cách bổ sung khí amoniac. Quá trình nuôi cấy mầm kết thúc ở thời điểm tiêu thụ hết sacarit trong môi trường được quan sát như một chỉ số.

Thành phần của môi trường nuôi cấy chính được thể hiện dưới đây.

Thành phần môi trường nuôi cấy (các nồng độ sau khi cấy 20% môi trường nuôi cấy mầm)

Glucosa	100g/L
MgSO ₄ ·7H ₂ O	0,4g/L
GD113	0,1mL/L
(NH ₄) ₂ SO ₄	5,0g/L
KH ₂ PO ₄	6,0g/L
Dịch chiết nấm men	6,0g/L
FeSO ₄ ·7H ₂ O	0,02g/L
MnSO ₄ ·5H ₂ O	0,02g/L
Axit xitric	0,02g/L
Betain*	2,0g/L
L-lysin hydroclorua	0,8g/L

DL-methionin	0,6g/L
Axit ϵ -diaminopimelic	0,6g/L
Canxi pantothenat	18mg/L
Tetracyclin hydroclorua	25mg/L

* N,N,N-trimetyl glyxin

Các tế bào thu được bằng cách nuôi cấy mầm ở thể tích 60mL được cấy vào trong 240mL môi trường có thành phần nêu trên có trong bình nhỏ có dung tích 1L, và được nuôi cấy ở nhiệt độ 34°C, 36°C, hoặc 38°C và ở độ pH = 4,9. Quá trình nuôi cấy kết thúc khi toàn bộ glucoza trong môi trường được tiêu thụ hết. Nồng độ axit L-glutamic được đo cho phần nổi lên trên bề mặt nuôi cấy được pha loãng thích hợp bằng nước bằng cách sử dụng thiết bị Biotech Analyzer (AS-210, Sakura SI).

Các kết quả được thể hiện trong Bảng 1. Rõ ràng là sự tích tụ axit L-glutamic của chủng khuyết gen *gcd*, chủng NA1:: Δ *gcd*, được gia tăng so với chủng đối chứng NA1.

Bảng 1: Axit L-glutamic được tạo ra (g/bình)

Chủng	34°C	36°C	38°C
NA1	13,0	12,4	13,3
NA1:: Δ <i>gcd</i>	15,4	16,0	16,7

Ví dụ 2: Sản xuất L-xystein bằng cách sử dụng chủng khuyết gen *gcd*

(1) Tạo ra vi khuẩn sản sinh L-xystein

Để đánh giá ảnh hưởng của sự khuyết gen *gcd* đối với sự sản sinh L-xystein của *P. ananatis*, vi khuẩn *P. ananatis* sản sinh L-xystein được tạo ra.

(1-1) Tạo ra plasmit biểu hiện gen *yeaS*

Trước tiên, plasmit để tạo ra chủng được tạo ra nêu trên. Phương pháp

này được mô tả dưới đây.

Nhờ quá trình PCR sử dụng ADN nhiễm sắc thể của *E. coli* MG1655 (ATCC No. 47076) làm khuôn mẫu cũng như P11 (agctgagtcg acccccagga aaaattggtt aataac, SEQ ID NO: 51) và P12 (agctgagcat gcttccaact gcgctaata cgc, SEQ ID NO: 52) làm đoạn mồi, thu được đoạn ADN chứa vùng khởi đầu của gen *nlpD* (sau đây gen khởi đầu *nlpD* kiểu đại được gọi là "Pnlp0") có kích cỡ khoảng 300bp. Ở các đầu 5' và 3' của các đoạn mồi nêu trên, các vị trí cho các enzym giới hạn SalI và PaeI được tạo ra. Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình 94°C trong 20 giây, 55°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng. Đoạn thu được được xử lý với SalI và PaeI, và được xen vào trong pMIV-5JS (đơn yêu cầu cấp patent Nhật số 2008-99668) ở vị trí SalI-PaeI để thu plasmit pMIV-Pnlp0. Trình tự nucleotit của đoạn PaeI-SalI của đoạn khởi đầu Pnlp0 được xen vào trong plasmit pMIV-Pnlp0 này là như được thể hiện trong SEQ ID NO: 41.

Sau đó, nhờ PCR sử dụng ADN nhiễm sắc thể của MG1655 làm khuôn mẫu, cũng như P13 (agctgatcta gaaaacagaa ttgcctggc ggc, SEQ ID NO: 53) và P14 (agctgaggat ccaggaagag ttgtagaaa cgc, SEQ ID NO: 54) làm đoạn mồi, thu được đoạn ADN chứa vùng kết thúc của gen *rrnB* có kích cỡ khoảng 300bp. Ở các đầu 5' của các đoạn mồi nêu trên, các vị trí cho các enzym giới hạn XbaI và BamHI được tạo ra. Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình gồm 94°C trong 20 giây, 59°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng. Đoạn thu được được xử lý với XbaI và BamHI, và được xen vào trong pMIV-Pnlp0 ở vị trí XbaI-BamHI để thu plasmit pMIV-Pnlp0-ter.

Sau đó, nhờ PCR sử dụng ADN nhiễm sắc thể của chủng MG1655 làm khuôn mẫu, cũng như P15 (agctgagtcg acgtgttcgc tgaatacggg gt, SEQ ID NO: 55) và P16 (agctgatcta gagaaagcat caggattgca gc, SEQ ID NO: 56) làm đoạn mồi, thu được đoạn ADN có kích cỡ khoảng 700bp chứa gen *yeaS*. Ở các đầu 5' của các đoạn mồi nêu trên, các vị trí cho các enzym giới hạn *SalI* và *XbaI* được tạo ra. Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình gồm 94°C trong 20 giây, 55°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng. Đoạn thu được được xử lý với *SalI* và *XbaI*, và được xen vào trong pMIV-Pnlp0-ter ở vị trí *SalI-XbaI* để thu plasmit pMIV-Pnlp0-YeaS3. Như nêu trên, đơn vị biểu hiện *yeaS* gồm vật chủ pMIV-5JS trên đó đoạn khởi đầu *nlpD*, gen *yeaS*, và gen kết thúc *rrnB* được ghép nối theo trình tự này được tạo ra.

Để cải biến vùng -10 của đoạn khởi đầu *nlpD* để làm cho nó trở thành đoạn khởi đầu mạnh hơn, vùng -10 được ngẫu nhiên hóa bằng phương pháp sau đây. Vùng khởi đầu *nlpD* chứa hai vùng được cho là có chức năng như đoạn khởi đầu (Fig. 3), và chúng được kí hiệu là *pnlp1* và *pnlp2*, trên hình vẽ. Nhờ kỹ thuật PCR sử dụng plasmit pMIV-Pnlp0 làm khuôn mẫu cũng như P11 và P17 (atcgtgaaga tctttccag tgttnannag ggtgccttgc acggtatna ngtcactgg ("n" có nghĩa là gốc tương ứng có thể là gốc bất kỳ trong số các gốc a, t, g và c), SEQ ID NO: 57) làm đoạn mồi, thu được đoạn ADN trong đó vùng -10 có trong trình tự có đầu 3' của đoạn khởi đầu *nlpD* (được gọi là -10(Pnlp1)) được ngẫu nhiên hóa (Fig. 3). Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng.

Hơn nữa, nhờ kỹ thuật PCR sử dụng plasmit pMIV-Pnlp0 làm khuôn mẫu cũng như P12 và P18 (tggaagagat cttcannnnn cgetgacctg cg ("n" có nghĩa là gốc tương ứng có thể làm gốc bất kỳ trong số các gốc a, t, g và c), SEQ ID NO: 58) làm đoạn môi, thu được đoạn ADN trong đó vùng -10 có trong trình tự có đầu 5' của đoạn khởi đầu nlpD (được gọi là -10(Pnlp2)) được ngẫu nhiên hóa (Fig. 3). Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng.

Các đoạn có đầu 3' và 5' có thể gắn kết bằng cách sử dụng các vị trí *Bgl*III được tạo ra trong các đoạn môi P17 và P18, và đoạn khởi đầu nlpD có độ dài đầy đủ trong đó hai vùng -10 được ngẫu nhiên hóa có thể tạo ra bằng cách gắn kết. Nhờ kỹ thuật PCR sử dụng đoạn này làm khuôn mẫu cũng như P11 và P12 làm đoạn môi, thu được đoạn ADN tương ứng có đoạn khởi đầu nlpD kiểu được cải biến có độ dài đầy đủ. Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 12 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng.

Đoạn đã khuếch đại được xử lý với các enzym giới hạn *Sal*I và *Pae*I, nhờ đó các vị trí được tạo ra ở các đầu 5' của các đoạn môi, và được xen vào trong plasmit pMIV-Pnlp0-YeaS3 được xử lý tương tự với *Sal*I và *Pae*I để thay thế Pnlp đột biến cho đoạn khởi đầu nlpD kiểu dại (Pnlp0) trên plasmit. Từ các plasmit này, một plasmit có trình tự đoạn khởi đầu (Pnlp8) được thể hiện trên Fig. 3 được chọn, và được gọi là pMIV-Pnlp8-YeaS7. Trình tự nucleotit của đoạn *Pae*I-*Sal*I của đoạn khởi đầu Pnlp8 được xen vào trong plasmit này là như được thể hiện trong SEQ ID NO: 42.

(1-2) Tạo ra plasmit biểu hiện *cysE* biến dị

Sau đó, từ pMW-Pomp-*cysE5* (WO2005/007841), phần đoạn Pomp-*cysE5* được cắt bằng *PaeI* và *SacI*, và được xen vào trong vị trí tương tự của pMIV-5JS để tạo ra pMIV-Pomp-CysE5. pMW-Pomp-*cysE5* là plasmit thu được bằng cách xen gen *cysE5* mã hóa cho SAT đột biến được ghép nối với gen khởi đầu *ompC* vào trong pMW118. Từ pACYC184 (số lưu giữ GenBank/EMBL X06403, có giá trị từ NIPPON GENE), gen kháng tetracyclin được cắt bằng *XbaI* và *Eco88I*, và đoạn gen này được xử lý bằng đoạn Klenow, và sau đó được xen vào trong pMIV-Pomp-CysE5 ở vị trí *PvuI* để tạo ra pMT-Pomp-CysE5. Sau đó, pMIV-Pnlp8-YeaS7 được cắt bằng *HindIII*, được làm bằng đầu bằng đoạn Klenow, và sau đó được cắt bằng *NcoI* để cắt đoạn chứa đoạn của gen kết thúc Pnlp8-YeaS-*rrnB* và gen đánh dấu kháng cloramphenicol. Đoạn này được ghép nối với đoạn cắt *SmaI* và *NcoI* của pMT-Pomp-CysE5 có pMIV-5JS tương tự như khung chính để tạo ra pMT-EY2. pMT-EY2 là plasmit có đoạn kết thúc Pnlp8-YeaS-*rrnB* và đoạn Pomp-CysE5 trên một plasmit.

(1-3) Đưa *cysE5* và *yeaS* vào trong chủng *P. ananatis* SC17

pMT-EY2 nêu trên có các vị trí dính của thể thực khuẩn Mu thu được từ pMIV-5JS (đơn yêu cầu cấp patent Nhật số 2008-99668). Bằng cách cho phép plasmit này cùng tồn tại với plasmit hỗ trợ pMH10 có Mu transposaza (Zimenkov D. et al., *Biotechnologiya* và (in Russian), 6, 1-22 (2004)) trong cùng một loại tế bào, đoạn gen kết thúc PompC-*cysE5*-Pnlp8-YeaS-*rrnB* gồm gen đánh dấu kháng cloramphenicol nằm giữa các vị trí dính thể thực khuẩn Mu trên plasmit pMT-EY2 này có thể được xen vào trong nhiễm sắc thể của chủng *P. ananatis* SC17 (patent Mỹ số 6,596,517). Hơn nữa, vì gen đánh dấu kháng cloramphenicol nằm trên plasmit pMT-EY2 có cấu trúc mà nó tồn tại giữa hai vị trí dính của thể thực khuẩn λ (λ attR và λ attL), gen đánh dấu kháng

cloramphenicol có thể được cắt và loại bỏ bằng phương pháp dưới đây.

Trước tiên, chủng SC17 được nhập nội với pMH10 bằng phương pháp xung điện được chọn bằng cách nuôi cấy qua đêm ở 30°C trên môi trường thạch LB chứa 20mg/L kanamycin. Thể biến nạp thu được được nuôi cấy ở 30°C, và pMT-E2 được đưa vào trong chủng này bằng phương pháp xung điện. Chủng này được biến nạp với cả pMH10 lẫn pMT-EY2 được cho sốc nhiệt ở 42°C trong 20 phút, và khuẩn lạc của các chủng kháng cloramphenicol được chọn lọc trên môi trường thạch LB chứa 20mg/L cloramphenicol. Nhiệt độ nuôi cấy cho sự chọn lọc này là 39°C. Như nêu trên, khoảng 50 dòng thu được, và pMH10 và pMT-EY2 được xử lý bằng cách nuôi cấy từng dòng ở 39°C trong 48 giờ trên môi trường thạch LB. Chủng có tính kháng cloramphenicol do sự xen đoạn trên nhiễm sắc thể và có tính nhạy kanamycin và ampicilin do đạt được sự xử lý cả hai plasmit. Hơn nữa, đã xác nhận rằng đoạn đích được dính vào nhiễm sắc thể của chủng thu được bằng PCR sử dụng ADN nhiễm sắc thể của chủng này làm khuôn mẫu cũng như P11 và P16 làm đoạn môi. Tất cả các dòng thu được được kí hiệu từ EY01 đến EY50, và môi trường sản xuất xystein được thực hiện bằng cách sử dụng các chủng từ EY01 đến EY50. Để nuôi cấy, phương pháp được nêu dưới đây được sử dụng. Chủng EY19 được chọn, là dòng sản sinh L-xystein với lượng lớn nhất nhờ quá trình nuôi cấy.

Gen đánh dấu kháng cloramphenicol được đưa vào trong chủng EY19 được loại bỏ bằng hệ cắt thu được từ thể thực khuẩn λ . Cụ thể, chủng EY19 được biến nạp với pMT-Int-Xis2 (WO2005/010175) mang gen Int-Xis của thể thực khuẩn λ , và (các) chủng EY19 có tính nhạy cloramphenicol thu được từ thể biến nạp thu được.

(1-4) Tạo ra chủng được gia tăng mức biểu hiện gen *cysPTWA* từ chủng EY19(s)

Tiếp theo, để gia tăng sự biểu hiện của gen *cysPTWA*, đoạn khởi đầu nằm ở phần hướng lên phía trên của cụm gen *cysPTWA* trên nhiễm sắc thể được thế bằng đoạn khởi đầu có sẵn *Pnlp8* nêu trên. Đoạn ADN chứa đoạn khởi đầu *nlp8* có kích cỡ khoảng 300bp thu được bằng PCR sử dụng *pMIV-Pnlp8-YeaS7* làm khuôn mẫu cũng như *P11* và *P12*. Chu trình PCR là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 20 chu trình gồm 94°C trong 20 giây, 59°C trong 20 giây, và 72°C trong 15 giây, và 72°C trong 5 phút làm chu trình sau cùng.

Đoạn ADN được khuếch đại chứa đoạn khởi đầu *nlp8* được xử lý bằng đoạn *Klenow*, được xen vào trong plasmit *pMW118-(λ attL-KmR- λ attR)* (WO2006/093322A2) được cắt bằng *XbaI* và sau đó được xử lý bằng đoạn *Klenow* để thu plasmit *pMW-Km-Pnlp8*. Bằng cách tiến hành PCR sử dụng *pMW-Km-Pnlp8* làm khuôn mẫu cũng như các đoạn mồi *P19* (tccgctcagc attttttca tcgctggtta ggtcatttat cccccaggaa aaattgggta, SEQ ID NO: 59) và *P20* (tttcacaccg ctcaaccgca gggcataacc ggcccttgaa gcctgctttt ttatactaag ttg, SEQ ID NO: 60), đoạn ADN có kích cỡ khoảng 1,6kb chứa đoạn *Km-Pnlp8* được khuếch đại. Chu trình PCR để khuếch đại là như sau: 95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 30 chu trình gồm 94°C trong 20 giây, 54°C trong 20 giây, và 72°C trong 90 giây, và 72°C trong 5 phút làm chu trình sau cùng. Ở cả hai đoạn mồi, trình tự có vai trò là đoạn đích trên nhiễm sắc thể để đính một đoạn cụ thể bằng phương pháp kết hợp phụ thuộc λ (phương pháp này còn được gọi là "sự kết hợp bức xạ ánh sáng đỏ" (Proc. Natl. Acad. Sci. USA, 2000, vol. 97, No. 12, pp.6640-6645)) (trong trường hợp này, trình tự gần đoạn khởi đầu *cysPTWA*) được tạo ra. Do vậy, nếu đoạn ADN thu được được xen vào trong chủng đích bằng phương pháp kết hợp phụ thuộc λ này, tạo ra cấu trúc mà *Km-Pnlp8* được đính ngay trước gen

cysPTWA trên nhiễm sắc thể, và gen cysPTWA được gắn với đoạn khởi đầu nlp8. Trình tự nucleotit của cụm gen cysPTWA được thể hiện trong SEQ ID NO: 43, và các trình tự axit amin được mã hóa bởi các gen cysP, cysT và cysW được thể hiện trong các SEQ ID NO: 44 đến 46. Trình tự nucleotit của gen cysA và trình tự axit amin được mã hóa bởi gen này được thể hiện trong các SEQ ID NO: 47 và 48.

Chủng *P. ananatis* SC17(0)/RSF-Red-TER là chủng vật chủ để thực hiện một cách hiệu quả phương pháp kết hợp phụ thuộc λ , và nó là chủng thu được bằng cách đưa plasmit hỗ trợ RSF-Red-TER mà biểu hiện các gen gam, bet và exo (sau đây được gọi là "các gen λ Red") vào trong chủng SC17(0), là chủng *P. ananatis* kháng gen λ Red (WO2008/075483). Chủng SC17(0) được lưu giữ tại Russian National Collection of Industrial Microorganisms (VKPM), GNII Genetika (address: Russia, 117545 Moscow, 1 Dorozhny proezd. 1) ngày 21 tháng 9 năm 2005 với số lưu giữ VKPM B-9246. Phương pháp tạo ra plasmit RSF-Red-TER được mô tả chi tiết trong WO2008/075483.

Chủng SC17(0)/RSF-Red-TER nêu trên được nuôi cấy với sự bổ sung IPTG để cảm ứng sự biểu hiện của các gen λ Red để tạo ra các tế bào cho phương pháp xung điện. Đoạn ADN đích nêu trên được đưa vào trong các tế bào này bằng phương pháp xung điện, và chủng tái tổ hợp trong đó đoạn khởi đầu nlp8 được đính dòng hướng lên phía trên của gen cysPTWA bằng phương pháp kết hợp phụ thuộc λ thu được bằng cách sử dụng tính kháng kanamycin làm gen đánh dấu. Bằng cách tiến hành PCR sử dụng ADN nhiễm sắc thể của chủng thu được làm khuôn mẫu, cũng như P21 (ctttgtccct ttagtgaagg, SEQ ID NO: 61) và P22 (agctgatcta gaagctgact cgagttaatg gcctcccaga cgac, SEQ ID NO: 62) làm đoạn mồi, đã nhận thấy rằng cấu trúc đích, Km-Pnlp8-cysPTWA, được tạo thành, và chủng này được gọi là chủng SC17(0)-Pnlp8-PTWA.

Sau đó, ADN nhiễm sắc thể của chủng SC17(0)-Pnlp8-PTWA được tinh lọc, và 10µg ADN nhiễm sắc thể này được đưa vào trong chủng EY19 bằng phương pháp xung điện để thu chủng kháng kanamycin. Quá trình khuếch đại được thực hiện bằng kỹ thuật PCR sử dụng ADN nhiễm sắc thể của chủng thu được làm khuôn mẫu cũng như P21 và P22 làm đoạn mồi để xác nhận rằng cấu trúc của Km-Pnlp8-cysPTWA đã được đưa vào trong nhiễm sắc thể của chủng EY19(s). Chủng thu được như nêu trên được gọi là chủng EYP197. Hơn nữa, gen đánh dấu kháng kanamycin được loại bỏ khỏi nhiễm sắc thể bằng cách sử dụng pMT-Int-Xis2 như nêu trên, và chủng mà trở nên nhạy kanamycin được kí hiệu là chủng EYP197(s).

(1-5) Tạo ra chủng mang gen 3-phosphoglyxerat dehydrogenaza (serA348) đột biến của chủng EYP197(s)

Về gen 3-phosphoglyxerat dehydrogenaza được đưa vào trong vi khuẩn sản sinh L-xystein, gen serA348 là gen mã hóa cho 3-phosphoglyxerat dehydrogenaza của *Pantoea ananatis* và mã hóa cho enzym đột biến gồm sự đột biến gây ra sự thay thế của gốc alanin cho gốc asparagin ở vị trí 348 (N348A) (J. Biol. Chem., 1996, 271 (38):23235-8) được tạo ra bằng phương pháp sau đây.

Trình tự của gen serA kiểu dại thu được từ *Pantoea ananatis* được thể hiện trong SEQ ID NO: 49. Trình tự axit amin được mã hóa bởi gen này được thể hiện trong SEQ ID NO: 50. Để thu đoạn ADN phía đầu 3' của gen serA trong đó sự đột biến nêu trên được đưa vào, PCR được thực hiện bằng cách sử dụng ADN nhiễm sắc thể của chủng SC17 làm khuôn mẫu cũng như P23 (agctgagtcg acatggcaaa ggtatcactg gaa, SEQ ID NO: 63) và P24 (gagaacgccc gggcgggctt cgtgaatatg cagc, SEQ ID NO: 64) làm đoạn mồi (95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 25 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 60 giây, và

72°C trong 5 phút làm chu trình sau cùng). Sau đó, để thu đoạn ADN phía đầu 5' trong đó sự đột biến được đưa vào, PCR được thực hiện theo cách tương tự bằng cách sử dụng ADN nhiễm sắc thể của chủng SC17 làm khuôn mẫu cũng như P25 (agctgatcta gacgtgggat cagtaaagca gg, SEQ ID NO: 65) và P26 (aaaaccgccc gggcggttctc ac, SEQ ID NO: 66) làm đoạn mồi (95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 20 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 20 giây, và 72°C trong 5 phút làm chu trình sau cùng). Cả hai đoạn PCR thu được được xử lý bằng enzym giới hạn SmaI, và được gắn bằng cách sử dụng ADN ligaza để thu đoạn ADN tương ứng với gen serA đột biến có độ dài đầy đủ có sự đột biến đích (N348A). Đoạn ADN này được khuếch đại bằng PCR sử dụng đoạn này làm khuôn mẫu cũng như P23 và P25 làm đoạn mồi (95°C trong 3 phút, sau đó 2 chu trình gồm 95°C trong 60 giây, 50°C trong 30 giây, và 72°C trong 40 giây, 15 chu trình gồm 94°C trong 20 giây, 60°C trong 20 giây, và 72°C trong 75 giây, và 72°C trong 5 phút làm chu trình sau cùng). Các vị trí enzym giới hạn SalI và XbaI được tạo ra trong các đoạn mồi P23 và P25 được xử lý bằng SalI và XbaI, và đoạn được dính vào trong pMIV-Pnlp8-ter được xử lý tương tự bằng SalI và XbaI để tạo ra pMIV-Pnlp8-serA348.

pMIV-Pnlp8-serA348 được tạo ra có vị trí dính Mu có trong pMIV-5JS (đơn yêu cầu cấp patent Nhật số 2008-99668). Bằng cách sử dụng plasmit này cùng với plasmit hỗ trợ pMH10 có Mu transposaza, đoạn gen kết thúc Pnlp8-serA348-rnB có gen đánh dấu kháng cloramphenicol có thể được xen vào trong nhiễm sắc thể của chủng *P. ananatis* SC17, như nêu trên. Plasmit pMIV-Pnlp8-serA348 và pMH10 được đưa vào trong chủng SC17(0) để thu chủng trong đó đoạn gen kết thúc Pnlp8-serA348-rnB được dính vào trong nhiễm sắc thể. Bằng kỹ thuật PCR sử dụng các đoạn mồi P11 và P25, đã nhận thấy rằng đoạn đích có

trong các tế bào. Hoạt tính 3-phosphoglyxerat dehydrogenaza trong các dịch chiết tế bào của 50 dòng thu được được đo, và chủng có hoạt tính cao nhất được chọn, và được kí hiệu là chủng SC17int-serA348. Sau đó, 10 μ g ADN nhiễm sắc thể của chủng SC17int-serA348 được đưa vào trong chủng EYP197(s) bằng phương pháp xung điện để thu chủng kháng cloramphenicol, và nhờ PCR sử dụng các đoạn mồi P11 và P25, đã nhận thấy rằng cấu trúc của Pnlp8-serA348 cùng với gen đánh dấu kháng cloramphenicol đã được đưa vào trong nhiễm sắc thể của chủng EYP197(s). Chủng thu được như nêu trên được kí hiệu là chủng EYPS1976.

Bằng phương pháp nêu trên để loại bỏ gen đánh dấu sử dụng pMT-Int-Xis2, gen đánh dấu kháng cloramphenicol được loại bỏ, và chủng trở nên nhạy cloramphenicol được kí hiệu là chủng EYPS1976(s).

(1-6) Tạo ra chủng khuyết gen gcd từ chủng EYPS1976(s)

ADN hệ gen được tạo ra từ chủng SC17(0):: Δ gcd được đề cập trong Ví dụ 1, và được đưa vào trong chủng EYPS1976(s) bằng phương pháp xung điện, và chủng khuyết gen gcd (chủng EYPS1976 Δ gcd) thu được từ chủng EYPS1976(s) bằng cách sử dụng gen kháng kanamycin làm gen đánh dấu.

(2) Nuôi cấy chủng vi khuẩn EYPS1976(s) và chủng EYPS1976 Δ gcd sản sinh L-xystein

Để đánh giá ảnh hưởng của sự loại bỏ gen gcd đối với sự sản xuất lên men L-xystein và O-axetylserin, là tiền chất của L-xystein, sự sản xuất nhờ lên men được thực hiện bằng cách sử dụng chủng vi khuẩn EYPS1976(s) sản sinh L-xystein và chủng khuyết gen gcd EYPS1976 Δ gcd thu được từ chủng EYPS1976(s), và các lượng L-xystein và O-axetylserin được tạo ra được so sánh. Để nuôi cấy, môi trường sản xuất L-xystein có thành phần sau đây được sử dụng.

Môi trường sản xuất L-xystein (các nồng độ của các thành phần là các nồng độ cuối)

Thành phần 1:

$(\text{NH}_4)_2\text{SO}_4$	15g/L
KH_2PO_4	1,5g/L
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	1g/L
Thiamin hydroclorua	0,1mg/L

Thành phần 2:

$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	1,7mg/L
$\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$	0,15mg/L
$\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$	0,7mg/L
$\text{MnCl} \cdot 4\text{H}_2\text{O}$	1,6mg/L
$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	0,3mg/L
$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	0,25mg/L

Thành phần 3:

Trypton	0,6g/L
Dịch chiết nấm men	0,3g/L
Natri clorua	0,6g/L

Thành phần 4:

Canxi carbonat	20g/L
----------------	-------

Thành phần 5:

L-Histidin monohydroclorua monohydrat	135mg/L
---------------------------------------	---------

Thành phần 6:

Natri thiosulfat	6g/L
------------------	------

Thành phần 7:

Pyridoxin hydroclorua 2mg/L

Thành phần 8:

Glucosa 40g/L

Đối với các thành phần này, đã chuẩn bị các dung dịch gốc có nồng độ gấp 10 lần (thành phần 1), nồng độ gấp 1000 lần (thành phần 2), nồng độ gấp 100/6 lần (thành phần 3), nồng độ gấp 100 lần (thành phần 5), 350g/L (thành phần 6), nồng độ gấp 1000 lần (thành phần 7), và nồng độ gấp 10 lần (thành phần 8), chúng được trộn ở thời điểm sử dụng, và thể tích thu được với nước đã khử trùng để đạt được các nồng độ cuối. Quá trình tiệt trùng được thực hiện bằng cách hấp ở 110°C trong 30 phút (các thành phần 1, 2, 3, 5 và 8), tiệt trùng bằng không khí nóng ở 180°C trong 5 giờ hoặc lâu hơn (thành phần 4), hoặc tiệt trùng bằng cách lọc (các thành phần 6 và 7).

Quá trình nuôi cấy sản xuất L-xystein được thực hiện như sau. Chủng EYPS1976(s) và chủng EYPS1976 Δ gcd được phủ và dàn đều lên môi trường thạch LB để tiến hành nuôi cấy sơ bộ qua đêm ở 34°C, và sau đó các tế bào tương ứng với 7cm trên đĩa cạo hai lần bằng vòng cấy kích cỡ 10- μ l (NUNC Blue Loop), và được cấy vào trong 2ml môi trường sản sinh L-xystein có trong ống nghiệm lớn (đường kính trong: 23mm, chiều dài: 20cm) ở các lượng tế bào về cơ bản là giống nhau ở thời điểm bắt đầu nuôi cấy.

Các chủng được nuôi cấy ở 34°C hoặc 38°C có khuấy, và quá trình nuôi cấy kết thúc sau 24 giờ. Ở thời điểm này, đã nhận thấy rằng glucosa là nguồn cacbon được tiêu thụ hoàn toàn. L-xystein được tạo ra trong môi trường được định lượng bằng phương pháp được nêu trong Gaitonde, M.K. (Biochem. J., 1967 Aug., 104(2):627-33). OAS (O-axetylserin) được tạo ra trong môi trường được định lượng bằng HPLC. Trong quá trình định lượng này, phương pháp

chuyển hóa OAS thành NAS (N-axetylserin) ổn định hơn bằng cách pha loãng mẫu bằng 200mM Tris-HCl (độ pH = 9,0) và phát hiện NAS được sử dụng. Các điều kiện HPLC là như sau.

Cột: Inertsil ODS-3 (cột kỵ nước, GL Science Co., Ltd.)

Tốc độ dòng dung dịch đệm: 1,0mL/phút

Nhiệt độ cột: 40°C

Sự dò: UV 210nm

Thể tích phủ của mẫu: 10mL

Dung dịch đệm: 0,1M $\text{KH}_2\text{PO}_4 \cdot \text{H}_3\text{PO}_4$ (độ pH = 2,2), 5mM Na 1-octansulfonat

Thử nghiệm được thực hiện sáu lần cho mỗi chủng, và các kết quả được thể hiện ở Bảng 2 với các giá trị trung bình và các độ lệch chuẩn. Như được thể hiện trong Bảng 2, thấy rằng sự khuyết gen gcd là hữu hiệu để gia tăng L-xystein và O-axetylserin ở cả hai nhiệt độ nuôi cấy 34°C và 38°C. Mức gia tăng đáng kể hơn là ở 38°C, và như vậy thấy rằng đặc biệt hữu hiệu cho quá trình nuôi cấy ở nhiệt độ cao. Cũng nhận thấy rằng nó cũng hữu hiệu để gia tăng lượng tế bào (OD) ở 38°C.

Bảng 2: Khả năng sản sinh L-xystein và NAS của chủng khuyết gen gcd

		Cys (g/L)	NAS (g/L)	OD600
34°C	EYPS1976(s)	1,02 ± 0,16	0,81 ± 0,10	15,1 ± 0,50
	EYPS1976Δgcd	1,29 ± 0,16	1,01 ± 0,10	15,0 ± 0,24
38°C	EYPS1976(s)	0,31 ± 0,12	0,54 ± 0,11	11,1 ± 1,68
	EYPS1976Δgcd	0,95 ± 0,21	0,87 ± 0,10	19,2 ± 0,57

Giải thích danh mục trình tự

SEQ ID NO: 1: Trình tự nucleotit của gen gcd của *Pantoea ananatis*

SEQ ID NO: 2: Trình tự axit amin của *Pantoea ananatis* GCD

SEQ ID NO: 3: Trình tự nucleotit gen hisD của *Pantoea ananatis*

SEQ ID NO: 4: Đoạn môi để khuếch đại đoạn kết hợp gen Km^r vào trong gen hisD

SEQ ID NO: 5: Đoạn môi để khuếch đại đoạn kết hợp gen Km^r vào trong gen hisD

SEQ ID NO: 6: Đoạn môi để khuếch đại gen cat

SEQ ID NO: 7: Đoạn môi để khuếch đại gen cat

SEQ ID NO: 8: Đoạn môi để khuếch đại gen sacB

SEQ ID NO: 9: Đoạn môi để khuếch đại gen sacB

SEQ ID NO: 10: Đoạn môi để khuếch đại đoạn ADN có đoạn khởi đầu PlacUV5

SEQ ID NO: 11: Đoạn môi để khuếch đại đoạn ADN có đoạn khởi đầu PlacUV5

SEQ ID NO: 12: Đoạn môi để khuếch đại đoạn ADN có các gen λ Red $\alpha\beta\gamma$ và tL3

SEQ ID NO: 13: Đoạn môi để khuếch đại đoạn ADN có các gen λ Red $\alpha\beta\gamma$ và tL3

SEQ ID NO: 14: Đoạn môi để khuếch đại đoạn ADN có đoạn khởi đầu PlacUV5 và TrnB

SEQ ID NO: 15: Đoạn môi để khuếch đại đoạn ADN có đoạn khởi đầu PlacUV5 và TrnB

SEQ ID NO: 16: Đoạn môi để khuếch đại attL

SEQ ID NO: 17: Đoạn môi để khuếch đại attL

SEQ ID NO: 18: Trình tự nucleotit của attL

SEQ ID NO: 19: Đoạn môi để khuếch đại attR

SEQ ID NO: 20: Đoạn môi để khuếch đại attR

SEQ ID NO: 21: Trình tự nucleotit của attR

SEQ ID NO: 22: Đoạn môi để khuếch đại đoạn ADN có gen bla

SEQ ID NO: 23: Đoạn môi để khuếch đại đoạn ADN có gen bla

- SEQ ID NO: 24: Đoạn mồi để khuếch đại đoạn ADN có gen *ter_rrnB*
- SEQ ID NO: 25: Đoạn mồi để khuếch đại đoạn ADN có gen *ter_rrnB*
- SEQ ID NO: 26: Trình tự nucleotit của đoạn ADN có gen kết thúc *ter_thrL*
- SEQ ID NO: 27: Đoạn mồi để khuếch đại đoạn ADN có gen kết thúc *ter_thrL*
- SEQ ID NO: 28: Đoạn mồi để khuếch đại đoạn ADN có gen kết thúc *ter_thrL*
- SEQ ID NO: 29: Đoạn mồi để khuếch đại gen *gltA* trừ ORF
- SEQ ID NO: 30: Đoạn mồi để khuếch đại gen *gltA* trừ ORF
- SEQ ID NO: 31: Đoạn mồi để khuếch đại gen *prpC*
- SEQ ID NO: 32: Đoạn mồi để khuếch đại gen *prpC*
- SEQ ID NO: 33: Đoạn mồi để loại bỏ gen *gcd*
- SEQ ID NO: 34: Đoạn mồi để loại bỏ gen *gcd*
- SEQ ID NO: 35: Đoạn mồi để xác nhận sự loại bỏ gen *gcd*
- SEQ ID NO: 36: Đoạn mồi để xác nhận sự loại bỏ gen *gcd*
- SEQ ID NO: 37: Trình tự nucleotit của gen *cysE* kiểu dại
- SEQ ID NO: 38: Trình tự axit amin của serin axetyltransferaza được mã hóa bởi *cysE* kiểu dại
- SEQ ID NO: 39: Trình tự nucleotit của gen *yeaS* kiểu dại
- SEQ ID NO: 40: Trình tự axit amin của *YeaS* kiểu dại
- SEQ ID NO: 41: Trình tự nucleotit của *Pnlp0*
- SEQ ID NO: 42: Trình tự nucleotit của *Pnlp8*
- SEQ ID NO: 43: Trình tự nucleotit của cụm gen *cysPTWA*
- SEQ ID NO: 44: Trình tự axit amin được mã hóa bởi gen *cysP*
- SEQ ID NO: 45: Trình tự axit amin được mã hóa bởi gen *cysT*
- SEQ ID NO: 46: Trình tự axit amin được mã hóa bởi gen *cysW*
- SEQ ID NO: 47: Trình tự nucleotit của gen *cysA*
- SEQ ID NO: 48: Trình tự axit amin được mã hóa bởi gen *cysA*

SEQ ID NO: 49: Trình tự nucleotit của gen serA kiểu dại của *Pantoea ananatis*

SEQ ID NO: 50: Trình tự axit amin được mã hóa bởi gen serA kiểu dại của *Pantoea ananatis*

Các SEQ ID NO: 51 đến 66: Các đoạn mồi P11 đến P26

Khả năng ứng dụng trong công nghiệp

Bằng cách sử dụng vi sinh vật theo sáng chế, axit L-amin như axit L-glutamic, L-lysin, L-threonin, L-arginin, L-histidin, L-isoleucin, L-valin, L-leucin, L-phenylalanin, L-tyrosin, L-tryptophan hoặc L-cystein có thể được sản xuất một cách hiệu quả bằng cách lên men.

Trong khi sáng chế đã được mô tả chi tiết với sự tham khảo các phương án ưu tiên, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này hoàn toàn có thể tiến hành các sự thay đổi khác, và các cách tương đương được sử dụng mà không trệt khỏi phạm vi của sáng chế. Mỗi tài liệu được nêu ở trên được kết hợp vào trong bản mô tả này để tham khảo.

YÊU CẦU BẢO HỘ

1. Phương pháp sản xuất axit L-amin bao gồm các bước:

A) nuôi cấy vi khuẩn thuộc họ *Enterobacteriaceae* và có khả năng sản sinh axit L-amin trong môi trường, và

B) thu gom axit L-amin từ môi trường hoặc vi khuẩn, trong đó vi khuẩn này đã được cải biến sao cho hoạt tính glucoza dehydrogenaza vốn có mà sử dụng pyroloquinolin quinon làm coenzym bị giảm so với vi khuẩn không được cải biến,

trong đó hoạt tính glucoza dehydrogenaza bị giảm bằng cách làm giảm mức biểu hiện của gen gcd trên nhiễm sắc thể của vi khuẩn, bằng phương pháp được chọn từ nhóm bao gồm:

a) xen chuỗi gen nhảy vào trong đoạn mã hóa của gen gcd sao cho sự biểu hiện bị ngắt quãng,

b) loại bỏ toàn bộ hoặc một phần của trình tự điều chỉnh sự biểu hiện của gen,

c) xen một hoặc nhiều nucleotit vào trong trình tự điều chỉnh sự biểu hiện của gen sao cho trình tự điều chỉnh sự biểu hiện không còn chức năng nữa,

d) loại bỏ một phần hoặc toàn bộ gen gcd trên nhiễm sắc thể, và

e) gây sự đột biến được chọn từ nhóm bao gồm sự thay đổi đoạn axit amin, codon kết thúc và đột biến dịch khung ở trong đoạn mã hóa của gen này trên nhiễm sắc thể sao cho sự biểu hiện bị ngắt quãng, và

f) tổ hợp của các phương pháp này.

2. Phương pháp theo điểm 1, trong đó gen gcd chứa ADN mã hóa trình tự axit amin của SEQ ID NO: 2 hoặc trình tự axit amin có độ tương đồng không thấp hơn 95% so với trình tự axit amin của SEQ ID NO: 2.

3. Phương pháp theo điểm 1, trong đó axit L-amin được chọn từ nhóm bao gồm

axit L-glutamic, L-lysin, L-threonin, L-arginin, L-histidin, L-isoloxin, L-valin, L-loxin, L-phenylalanin, L-tyrosin, L-tryptophan, L-xystein và hỗn hợp của chúng.

4. Phương pháp theo điểm 3, trong đó axit L-amin là axit L-glutamic hoặc L-xystein.

5. Phương pháp theo điểm 4, trong đó axit L-amin là axit L-glutamic, và hoạt tính hoặc các hoạt tính của enzym được chọn từ nhóm bao gồm xitrat syntaza, metyl xitrat syntaza, phosphoenolpyruvat carboxylaza, glutamat dehydrogenaza và các hỗn hợp của chúng, được tăng cường ở vi khuẩn bằng phương pháp được chọn từ nhóm bao gồm:

a) gia tăng số lượng bản sao của gen mã hóa enzym và

b) gây sự đột biến ở trong đoạn khởi đầu của gen này.

6. Phương pháp theo điểm 4, trong đó axit L-amin là L-xystein, và hoạt tính hoặc các hoạt tính của enzym được chọn từ nhóm bao gồm 3-phosphoglyxerat dehydrogenaza, serin axetyltransferaza, hệ vận chuyển sulfat/thiosulfat và các hỗn hợp của chúng được tăng cường, và/hoặc sự biểu hiện của gen *yeaS* được tăng cường, trong đó hoạt tính hoặc sự biểu hiện được gia tăng bởi phương pháp được chọn từ nhóm bao gồm:

a) gia tăng số lượng bản sao của gen mã hóa enzym và/hoặc gen *yeaS*, và

b) gây sự đột biến ở trong đoạn khởi đầu của gen này.

7. Phương pháp theo điểm 1, trong đó vi khuẩn thuộc giống được chọn từ nhóm bao gồm *Pantoea*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Providencia*, *Salmonella*, *Serratia*, *Morganella*, và *Yersinia*.

DANH MỤC TRÌNH TỰ

<110> Ajinomoto Co., Inc.

<120> Phương pháp sản xuất axit L-amin

<130> D095-C9293

<150> JP2008-229736

<151> 2008-09-08

<150> JP2009-032839

<151> 2009-02-16

<160> 66

<170> PatentIn version 3.5

<210> 1

<211> 2931

<212> ADN

<213> Pantoea anantis

<220>

<221> CDS

<222> (301)..(2691)

<400> 1

```

gogtctggcg atctcattcc ctttacgctt taacctttct tatctcctgg cactattgaa      60
tcatatcggc aaggcccat taacacctc taaagtattg ccttagatgc gttgttaaga      120
gctttaagcc aaaaaaatca aaactcatac caattttgct ataacattta acagatatgg      180
atcatoacgc ttcagggaag agcgaaataa gaatgccttt catcttttgt actttatgac      240
tgacaacaat ctatctgatt gttttcctga gttttctggc aacgaaatga ggtcaacatt      300
atg ggg aaa aac tcc tca tcc ttt agc gtg gtc cgt ttt tta acg gtg      348
Met Gly Lys Asn Ser Ser Phe Ser Val Val Arg Phe Leu Thr Val
1           5           10           15
ctg ttc gcc gtg cta acg ggt gcg ttc atg tta att ggt ggt atc tgg      396
Leu Phe Ala Val Leu Thr Gly Ala Phe Met Leu Ile Gly Gly Ile Trp
20           25           30
ctg gcc acg atc ggt ggt tcc tgg tac tac atc atc ggc ggt gca gcc      444
Leu Ala Thr Ile Gly Gly Ser Trp Tyr Tyr Ile Ile Gly Gly Ala Ala
35           40           45
atg ctg ctt acc gct ttc ctg ctg tgg cga cgt aac agc gct gcc ctg      492
Met Leu Leu Thr Ala Phe Leu Leu Trp Arg Arg Asn Ser Ala Ala Leu
50           55           60
gtt gtc tat gcg ctc tta ctg ctg gct acg ctg gcc tgg ggc gtt tgg      540
Val Val Tyr Ala Leu Leu Leu Leu Ala Thr Leu Ala Trp Gly Val Trp

```

65					70					75				80	
gaa gtc ggc acc gac ttc tgg gca ctg gca ccg cgt acc gac gta ctg															588
Glu Val Gly Thr Asp Phe Trp Ala Leu Ala Pro Arg Thr Asp Val Leu															
				85					90					95	
gtg atc ttt ggc gtc tgg ctg gtg ttg ccc ttt gtc tat cgc ggc tta															636
Val Ile Phe Gly Val Trp Leu Val Leu Pro Phe Val Tyr Arg Gly Leu															
			100					105					110		
tac cag ccg ggt aaa ggc gca ctg ggt gcc atg ggc gta gcg ctg gtt															684
Tyr Gln Pro Gly Lys Gly Ala Leu Gly Ala Met Gly Val Ala Leu Val															
			115					120					125		
gcc agt gca gcg gtg tta acc tat tcc gtc ttt aat gat ccg caa gtg															732
Ala Ser Ala Ala Val Leu Thr Tyr Ser Val Phe Asn Asp Pro Gln Val															
			130					135					140		
gtt aac ggt gca tta ccg gca aca gcg gat aat gcg cct cag gca cag															780
Val Asn Gly Ala Leu Pro Ala Thr Ala Asp Asn Ala Pro Gln Ala Gln															
			145					150					155		160
ccg ttg agc aat att gct gat ggt gac tgg ccg gcc tat gcg cgc gat															828
Pro Leu Ser Asn Ile Ala Asp Gly Asp Trp Pro Ala Tyr Ala Arg Asp															
			165					170					175		
cag caa ggg acg cgc ttc tcg ccg ctg aag cag atc aac cac gac aat															876
Gln Gln Gly Thr Arg Phe Ser Pro Leu Lys Gln Ile Asn His Asp Asn															
			180					185					190		
gtg aaa gaa ctg cag gtt gcc tgg caa ttc cag acc ggt gat atg aaa															924
Val Lys Glu Leu Gln Val Ala Trp Gln Phe Gln Thr Gly Asp Met Lys															
			195					200					205		
cgc cca agc gat ccg ggc gaa att acc gat gaa gtg acg cca atc aag															972
Arg Pro Ser Asp Pro Gly Glu Ile Thr Asp Glu Val Thr Pro Ile Lys															
			210					215					220		
att cgc gac acg ctg tat ctt tgc acg cca cat cag att tta ttt gct															1020
Ile Arg Asp Thr Leu Tyr Leu Cys Thr Pro His Gln Ile Leu Phe Ala															
			225					230					235		240
ctg gat gcg gcc acc ggc aag caa aag tgg aag ttt gat ccc ggc ctg															1068
Leu Asp Ala Ala Thr Gly Lys Gln Lys Trp Lys Phe Asp Pro Gly Leu															
			245					250					255		
aaa acc aac cca acc ttc cag cac gtg acc tgt cgt ggt gtg tca tac															1116
Lys Thr Asn Pro Thr Phe Gln His Val Thr Cys Arg Gly Val Ser Tyr															
			260					265					270		
cac gaa ttc cct gca gcg aag gat gcg tcc aat acc cag cct gcg ctg															1164
His Glu Phe Pro Ala Ala Lys Asp Ala Ser Asn Thr Gln Pro Ala Leu															
			275					280					285		
tgc tcg cgt cgt atc tac ctg cca gtc aat gac ggg cgt ttg ttc gcg															1212
Cys Ser Arg Arg Ile Tyr Leu Pro Val Asn Asp Gly Arg Leu Phe Ala															
			290					295					300		
ctg gat gcg gaa acc ggt gaa cgc tgc ccg gcc ttt ggt aac aac ggt															1260
Leu Asp Ala Glu Thr Gly Glu Arg Cys Pro Ala Phe Gly Asn Asn Gly															
			305					310					315		320
gag ctg gat ctg cag cac aag cag ccg gtc aca acg cca ggc atg tat															1308
Glu Leu Asp Leu Gln His Lys Gln Pro Val Thr Thr Pro Gly Met Tyr															
			325					330					335		

gag cca acc tcc cca ccg gtg att act gac acc acc att gtg atg got	1356
Glu Pro Thr Ser Pro Pro Val Ile Thr Asp Thr Thr Ile Val Met Ala	
340 345 350	
ggc gcg gta acc gat aac ttt tca acc cgt gaa cct tca ggc gcc atc	1404
Gly Ala Val Thr Asp Asn Phe Ser Thr Arg Glu Pro Ser Gly Ala Ile	
355 360 365	
cgt ggc ttt gat gtg aac acc ggt aag ctg ttg tgg gtg ttc gat ccg	1452
Arg Gly Phe Asp Val Asn Thr Gly Lys Leu Leu Trp Val Phe Asp Pro	
370 375 380	
ggc gcg aaa gat cct aac gcg att ccg gcg gat gaa cac acg ttc acc	1500
Gly Ala Lys Asp Pro Asn Ala Ile Pro Ala Asp Glu His Thr Phe Thr	
385 390 395 400	
atg aac tcc cct aac tcc tgg gca cct gcg gtt tac gat ccg aag ctg	1548
Met Asn Ser Pro Asn Ser Trp Ala Pro Ala Val Tyr Asp Pro Lys Leu	
405 410 415	
gat atc gtt tac ctg cca atg ggg gtg acc acg ccg gat atc tgg ggc	1596
Asp Ile Val Tyr Leu Pro Met Gly Val Thr Thr Pro Asp Ile Trp Gly	
420 425 430	
ggc aac cgc aca cct gag cag gaa cgt tat gcc agc agc gtg ctg gcg	1644
Gly Asn Arg Thr Pro Glu Gln Glu Arg Tyr Ala Ser Ser Val Leu Ala	
435 440 445	
ctg aac gcg acg acc ggt aag ctg gtg tgg tca tat cag act gtg cat	1692
Leu Asn Ala Thr Thr Gly Lys Leu Val Trp Ser Tyr Gln Thr Val His	
450 455 460	
cac gat ctg tgg gat atg gac ctg cct tcc cag ccg acg ctg gcg gat	1740
His Asp Leu Trp Asp Met Asp Leu Pro Ser Gln Pro Thr Leu Ala Asp	
465 470 475 480	
att acc gat aaa gac ggt aat acc gtg ccg gtt atc tat gcc cct gcc	1788
Ile Thr Asp Lys Asp Gly Asn Thr Val Pro Val Ile Tyr Ala Pro Ala	
485 490 495	
aaa acc ggg aac atc ttt gtt ctg gat cgc cgc aca ggt aaa act gtg	1836
Lys Thr Gly Asn Ile Phe Val Leu Asp Arg Arg Thr Gly Lys Thr Val	
500 505 510	
gtt ccg gcc ccg gaa acc cct gtt ccg cag ggc gca gct aag ggc gac	1884
Val Pro Ala Pro Glu Thr Pro Val Pro Gln Gly Ala Ala Lys Gly Asp	
515 520 525	
cat gtc tca get aca cag cct tac tct gaa ctg acc ttc cgt ccg aaa	1932
His Val Ser Ala Thr Gln Pro Tyr Ser Glu Leu Thr Phe Arg Pro Lys	
530 535 540	
cag aac ctg acg gat aag gac atg tgg ggc gcg acg atg tat gac cag	1980
Gln Asn Leu Thr Asp Lys Asp Met Trp Gly Ala Thr Met Tyr Asp Gln	
545 550 555 560	
ctg gtg tgc cgc gtg att ttc aaa cgt ctg cgc tac gaa ggt ccg ttc	2028
Leu Val Cys Arg Val Ile Phe Lys Arg Leu Arg Tyr Glu Gly Pro Phe	
565 570 575	
acg cca cct tct gag cag ggc acc ctg gtc ttc ccg ggc aac ctg ggc	2076
Thr Pro Pro Ser Glu Gln Gly Thr Leu Val Phe Pro Gly Asn Leu Gly	
580 585 590	
atg ttt gaa tgg ggc ggc att tcc gtt gat ccg cat cgt cag att gcg	2124

```

Met Phe Glu Trp Gly Gly Ile Ser Val Asp Pro His Arg Gln Ile Ala
    595                                600                                605
att gct aac cca atg gcg ctg ccg ttc gtg tct aag ctg atc cca cgc      2172
Ile Ala Asn Pro Met Ala Leu Pro Phe Val Ser Lys Leu Ile Pro Arg
    610                                615                                620
ggg ccg ggt aat ccg gaa gag cca cca aaa ggc gca acg ggc ggt tca      2220
Gly Pro Gly Asn Pro Glu Glu Pro Pro Lys Gly Ala Thr Gly Gly Ser
    625                                630                                635                                640
ggg act gaa acc ggt att cag ccg cag tac ggt gtg cca tat ggc gtt      2268
Gly Thr Glu Thr Gly Ile Gln Pro Gln Tyr Gly Val Pro Tyr Gly Val
    645                                650                                655
gaa ctg aat ccg ttc ctg tca cct ttt ggt ctg ccg tgt aaa caa cct      2316
Glu Leu Asn Pro Phe Leu Ser Pro Phe Gly Leu Pro Cys Lys Gln Pro
    660                                665                                670
gca tgg ggt tat gtt tct gct gtt gac ctg aaa acc aac gaa gtg gtg      2364
Ala Trp Gly Tyr Val Ser Ala Val Asp Leu Lys Thr Asn Glu Val Val
    675                                680                                685
tgg aaa caa cgt att ggt acc gtt cgt gac agc tca cct gta ccg ctg      2412
Trp Lys Gln Arg Ile Gly Thr Val Arg Asp Ser Ser Pro Val Pro Leu
    690                                695                                700
ccc ttt aaa atg ggt atg cca atg ctg ggc gga ccg gtt gcc acc gca      2460
Pro Phe Lys Met Gly Met Pro Met Leu Gly Gly Pro Val Ala Thr Ala
    705                                710                                715                                720
ggc aaa gtg ttc ttt att ggc gca acg gct gat aac tac ctg cgc gct      2508
Gly Lys Val Phe Phe Ile Gly Ala Thr Ala Asp Asn Tyr Leu Arg Ala
    725                                730                                735
ttc agc acc gac acc ggt gaa ctc ttg tgg cag gcg cgc ctg cca gcc      2556
Phe Ser Thr Asp Thr Gly Glu Leu Leu Trp Gln Ala Arg Leu Pro Ala
    740                                745                                750
ggg ggt cag gca acg cca atg acc tat gaa gtt aac ggc aag caa tac      2604
Gly Gly Gln Ala Thr Pro Met Thr Tyr Glu Val Asn Gly Lys Gln Tyr
    755                                760                                765
gtt gtg att gct gcc ggt ggc cat ggt tca ttc ggc acc aag ctg ggc      2652
Val Val Ile Ala Ala Gly Gly His Gly Ser Phe Gly Thr Lys Leu Gly
    770                                775                                780
gat tac gtg att gcc tat gcg ctg ccc gac cag aag taa ttaacacctg      2701
Asp Tyr Val Ile Ala Tyr Ala Leu Pro Asp Gln Lys
    785                                790                                795
aacagaggcg gactccggte cgcctttttt tatgectgct atctgccttg tgcttttgcg      2761
cgtgggggagc gccagcttaa ccaggcgcaac agcccatga ccatgcaggt ggccagaaat      2821
actggccgca ttcgccacgc gccaccaatc accccaccaa aaaggggacc actgacctgg      2881
ccgatatact ggcgcgacgt cgaatagccg agcatgcgcc ccacctgcac      2931

```

<210> 2
 <211> 796
 <212> PRT
 <213> Pantoea anantis

<400> 2

```

Met Gly Lys Asn Ser Ser Ser Phe Ser Val Val Arg Phe Leu Thr Val
1          5          10          15
Leu Phe Ala Val Leu Thr Gly Ala Phe Met Leu Ile Gly Gly Ile Trp
20          25          30
Leu Ala Thr Ile Gly Gly Ser Trp Tyr Tyr Ile Ile Gly Gly Ala Ala
35          40          45
Met Leu Leu Thr Ala Phe Leu Leu Trp Arg Arg Asn Ser Ala Ala Leu
50          55          60
Val Val Tyr Ala Leu Leu Leu Leu Ala Thr Leu Ala Trp Gly Val Trp
65          70          75          80
Glu Val Gly Thr Asp Phe Trp Ala Leu Ala Pro Arg Thr Asp Val Leu
85          90          95
Val Ile Phe Gly Val Trp Leu Val Leu Pro Phe Val Tyr Arg Gly Leu
100          105          110
Tyr Gln Pro Gly Lys Gly Ala Leu Gly Ala Met Gly Val Ala Leu Val
115          120          125
Ala Ser Ala Ala Val Leu Thr Tyr Ser Val Phe Asn Asp Pro Gln Val
130          135          140
Val Asn Gly Ala Leu Pro Ala Thr Ala Asp Asn Ala Pro Gln Ala Gln
145          150          155          160
Pro Leu Ser Asn Ile Ala Asp Gly Asp Trp Pro Ala Tyr Ala Arg Asp
165          170          175
Gln Gln Gly Thr Arg Phe Ser Pro Leu Lys Gln Ile Asn His Asp Asn
180          185          190
Val Lys Glu Leu Gln Val Ala Trp Gln Phe Gln Thr Gly Asp Met Lys
195          200          205
Arg Pro Ser Asp Pro Gly Glu Ile Thr Asp Glu Val Thr Pro Ile Lys
210          215          220
Ile Arg Asp Thr Leu Tyr Leu Cys Thr Pro His Gln Ile Leu Phe Ala
225          230          235          240
Leu Asp Ala Ala Thr Gly Lys Gln Lys Trp Lys Phe Asp Pro Gly Leu
245          250          255
Lys Thr Asn Pro Thr Phe Gln His Val Thr Cys Arg Gly Val Ser Tyr
260          265          270
His Glu Phe Pro Ala Ala Lys Asp Ala Ser Asn Thr Gln Pro Ala Leu
275          280          285
Cys Ser Arg Arg Ile Tyr Leu Pro Val Asn Asp Gly Arg Leu Phe Ala
290          295          300
Leu Asp Ala Glu Thr Gly Glu Arg Cys Pro Ala Phe Gly Asn Asn Gly
305          310          315          320
Glu Leu Asp Leu Gln His Lys Gln Pro Val Thr Thr Pro Gly Met Tyr
325          330          335
Glu Pro Thr Ser Pro Pro Val Ile Thr Asp Thr Thr Ile Val Met Ala
340          345          350
Gly Ala Val Thr Asp Asn Phe Ser Thr Arg Glu Pro Ser Gly Ala Ile
355          360          365
Arg Gly Phe Asp Val Asn Thr Gly Lys Leu Leu Trp Val Phe Asp Pro
370          375          380

```

Gly	Ala	Lys	Asp	Pro	Asn	Ala	Ile	Pro	Ala	Asp	Glu	His	Thr	Phe	Thr	385	390	395	400
Met	Asn	Ser	Pro	Asn	Ser	Trp	Ala	Pro	Ala	Val	Tyr	Asp	Pro	Lys	Leu	405	410	415	
Asp	Ile	Val	Tyr	Leu	Pro	Met	Gly	Val	Thr	Thr	Pro	Asp	Ile	Trp	Gly	420	425	430	
Gly	Asn	Arg	Thr	Pro	Glu	Gln	Glu	Arg	Tyr	Ala	Ser	Ser	Val	Leu	Ala	435	440	445	
Leu	Asn	Ala	Thr	Thr	Gly	Lys	Leu	Val	Trp	Ser	Tyr	Gln	Thr	Val	His	450	455	460	
His	Asp	Leu	Trp	Asp	Met	Asp	Leu	Pro	Ser	Gln	Pro	Thr	Leu	Ala	Asp	465	470	475	480
Ile	Thr	Asp	Lys	Asp	Gly	Asn	Thr	Val	Pro	Val	Ile	Tyr	Ala	Pro	Ala	485	490	495	
Lys	Thr	Gly	Asn	Ile	Phe	Val	Leu	Asp	Arg	Arg	Thr	Gly	Lys	Thr	Val	500	505	510	
Val	Pro	Ala	Pro	Glu	Thr	Pro	Val	Pro	Gln	Gly	Ala	Ala	Lys	Gly	Asp	515	520	525	
His	Val	Ser	Ala	Thr	Gln	Pro	Tyr	Ser	Glu	Leu	Thr	Phe	Arg	Pro	Lys	530	535	540	
Gln	Asn	Leu	Thr	Asp	Lys	Asp	Met	Trp	Gly	Ala	Thr	Met	Tyr	Asp	Gln	545	550	555	560
Leu	Val	Cys	Arg	Val	Ile	Phe	Lys	Arg	Leu	Arg	Tyr	Glu	Gly	Pro	Phe	565	570	575	
Thr	Pro	Pro	Ser	Glu	Gln	Gly	Thr	Leu	Val	Phe	Pro	Gly	Asn	Leu	Gly	580	585	590	
Met	Phe	Glu	Trp	Gly	Gly	Ile	Ser	Val	Asp	Pro	His	Arg	Gln	Ile	Ala	595	600	605	
Ile	Ala	Asn	Pro	Met	Ala	Leu	Pro	Phe	Val	Ser	Lys	Leu	Ile	Pro	Arg	610	615	620	
Gly	Pro	Gly	Asn	Pro	Glu	Glu	Pro	Pro	Lys	Gly	Ala	Thr	Gly	Gly	Ser	625	630	635	640
Gly	Thr	Glu	Thr	Gly	Ile	Gln	Pro	Gln	Tyr	Gly	Val	Pro	Tyr	Gly	Val	645	650	655	
Glu	Leu	Asn	Pro	Phe	Leu	Ser	Pro	Phe	Gly	Leu	Pro	Cys	Lys	Gln	Pro	660	665	670	
Ala	Trp	Gly	Tyr	Val	Ser	Ala	Val	Asp	Leu	Lys	Thr	Asn	Glu	Val	Val	675	680	685	
Trp	Lys	Gln	Arg	Ile	Gly	Thr	Val	Arg	Asp	Ser	Ser	Pro	Val	Pro	Leu	690	695	700	
Pro	Phe	Lys	Met	Gly	Met	Pro	Met	Leu	Gly	Gly	Pro	Val	Ala	Thr	Ala	705	710	715	720
Gly	Lys	Val	Phe	Phe	Ile	Gly	Ala	Thr	Ala	Asp	Asn	Tyr	Leu	Arg	Ala	725	730	735	
Phe	Ser	Thr	Asp	Thr	Gly	Glu	Leu	Leu	Trp	Gln	Ala	Arg	Leu	Pro	Ala	740	745	750	
Gly	Gly	Gln	Ala	Thr	Pro	Met	Thr	Tyr	Glu	Val	Asn	Gly	Lys	Gln	Tyr	755	760	765	
Val	Val	Ile	Ala	Ala	Gly	Gly	His	Gly	Ser	Phe	Gly	Thr	Lys	Leu	Gly				

770 775 780
 Asp Tyr Val Ile Ala Tyr Ala Leu Pro Asp Gln Lys
 785 790 795

<210> 3
 <211> 1308
 <212> ADN
 <213> *Pantoea ananatis*

<400> 3
 atgagcagaa tcatgaagcc cgtgaactgg gaagcctgca gcagcagaggc gcagcagggc 60
 ctgtttggcac gccctgcegt cgcctegtct gacagcatca gccagatcgt gcgcgatgtg 120
 ttgggtcagag tgaaagagga aggegatgcg gctttacgag aattcagcgc gcgctttgac 180
 aaggttgaaa cagacgacct gcgcgttacg ccacagcaga tgcaggcggc cagcgatcgc 240
 cttggtgacg agctgaaaca ggcgatggcc gtggccattg gcaatattga aacctttcac 300
 cgtgogcaga tcctgcggcc ggtggatgtg gaaacgcagc ccggcgtgcg ctgtcagcaa 360
 attacgcgcc cgatgaaatc ggtgggcttg tatattccgg gcggtttctgc cccgctgttt 420
 tctaccgttc tgatgctggt taccgccggc cggattgcgg gctgtggtcg cgtggtgctg 480
 tgctgcggcc cgcgcgattg tgatgaaatt ctctacgcgg ccaaactttg cgggtgtgaa 540
 gaagtgttcc aggtgggtgg atcacaggcg attgccggcc tggcttttgg caccgaaagc 600
 atccctaagg tagataaaat ttttggtecg ggcaacgcgt gggttaccga agccaaacgt 660
 caggtcagcc agcgccttga tggcgcgcg attgatatgc ccgctggccc gtcggaagtg 720
 ctggtgattg ccgatgaagg tgccacaccg gccttcgttg cctctgatct gctgtcgag 780
 gcggaacacg gccctgactc gcaggtgatt ttactgacgc cttecgctggc gctggccgag 840
 cgcgtcgccg aggcggtgga ggatcagctg gccagttgc cacgtgcggc gacagccgc 900
 caggcaactg aaagcagcgc cctgatcgtc gcccgggata tgcagcaatg cattgcgac 960
 tccaaccgct atggtcggga gcacctgatt ctgcaaaccc gcacgcacag ggatctggtg 1020
 gaacagatta ccagcgscgg ttccggttttc ctgggcgact ggtaaccgga atccgcagga 1080
 gattatgctt cgggcaecaa ccacgtgctg ccgacctacg gctatacgc gacatgctcc 1140
 agcctgggccc tggccgactt tcagaaacgc atgacggtac aggagctgac gccgcagggc 1200
 ttccctgaacc tggcgggcgc catcgaaacc ctggcgggcg ctgaacagct gcacgcccac 1260
 aaaaatgcgc tcacgttgcg cgttgccgca ctcaaggagc aagcatga 1308

<210> 4
 <211> 68
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mồi

<400> 4
 coatagcggt tggagatcgc aatgcattgc tgcatatccc tgaagcctgc tttttatac 60
 taagttgg 68

<210> 5
 <211> 68

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 5

gcccggccagg cactggaaaag cagccgcctg atcgtegcc cgtcaagtt agtataaaaa 60
agctgaac 68

<210> 6

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 6

tagcgagatc tctgatgtcc ggcggtgctt ttg 33

<210> 7

<211> 32

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 7

aaaaagagct cttacgcccc gccctgccac tc 32

<210> 8

<211> 34

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 8

caggatctag aaggagacat gaacgatgaa catc 34

<210> 9

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 9

gataaggatc cgaaataaaa gaaaatgccca atagga

36

<210> 10

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 10

cctttgagct cgcgggcagt gagcgcaacg c

31

<210> 11

<211> 48

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 11

ctagagcggc cgcgcatcgg gatcctcctg tgtgaaattg ttatccgc

48

<210> 12

<211> 42

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 12

ctctacgac gagggaggta taaaaaatgg atattaatac tg

42

<210> 13

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 13

tcaaaggggc cgcttcttcg tctgtttcta ctggta

36

<210> 14

<211> 31

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 14

cctttggtac cgcgggcagt gagcgcaacg c

31

<210> 15

<211> 34

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 15

aacaggaatt ctttgectgg cggcagtagc gggg

34

<210> 16

<211> 40

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P1

<400> 16

ctagtaagat cttgaagcct gcttttttat actaagttgg

40

<210> 17

<211> 41

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P2

<400> 17

atgatcgaat tcgaaatcaa ataatgattt tattttgact g

41

<210> 18

<211> 120

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn ADN chứa attL

<400> 18

agatcttgaa gcctgctttt ttataactaag ttggcattat aaaaaagcat tgcttatcaa
tttgttgcaa cgaacaggtc actatcagtc aaaataaaat cattatttga tticgaattc

60

120

<210> 19

<211> 41

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P3

<400> 19

atgccactgc agtctgttac aggtcactaa taccatctaa g

41

<210> 20

<211> 46

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P4

<400> 20

accgttaagc ttcttagacg ctcaagttag tataaaaaag ctgaac

46

<210> 21

<211> 184

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn ADN chứa attR

<400> 21

```
ctgcagtcctg ttacaggtca ctaataccat ctaagtagtt gattcatagt gactgcatat 60
gttggtgtttt acagtattat gtagtetggt ttttatgcaa aatctaattt aatatattga 120
tatttatatc attttaagtt tctcgttcag cttttttata ctaacttgag cgtctagaaa 180
gctt 184
```

<210> 22

<211> 38

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn môi P5

<400> 22

```
ttcttagacg tcaggtggca cttttcgggg aaatgtgc 38
```

<210> 23

<211> 37

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn môi P6

<400> 23

```
taacagagat ctgcgcgaga aaaaaaggat ctcaaga 37
```

<210> 24

<211> 46

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn môi P7

<400> 24

```
aacagagatc taagcttaga tectttgctt ggcggcagta gcgcgg 46
```

<210> 25

<211> 35

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P8

<400> 25

ataaactgca gcaaaaagag tttgtagaaa cgcaa

35

<210> 26

<211> 1388

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn ADN chứa gen Tc và ter_thrL

<400> 26

```

gaatttcat gtttgacagc ttatcatoga taagctttaa tgcggtagtt tatcacagtt      60
aaattgctaa cgcagtcagg caacgtgtat gaaatctaac aatgcgctca tcgtcatcct      120
cggcaccgtc accctggatg ctgtaggcat aggcttgggt atgccggtac tgcggggcct      180
cttgccgggat atcgtccatt ccgacagcat cgcagtcac tatggcggtgc tgctagcgt      240
atatgcgttg atgcaatttc tatgcgcacc cgttctcgga gcactgtccg accgctttgg      300
ccgcgcacca gtccctgctcg cttecgtaact tggagccact atcgactacg cgatcatggc      360
gaccacaccc gtccctgtgga tectetacgc cggacgcate gtggccggca tcaecggcgc      420
cacagggtgc gttgctggcg cctatatcgc cgacatcacc gatggggaag atcgggctcg      480
ccacttcggg ctcatgagcg cttgtttcgg cgtgggtatg gtggcaggcc ccgtggccgg      540
gggactgttg ggcccatct ccttgcatgc accattcctt gcggcggcgg tgetcaacgg      600
cctcaacctt ctactgggt gcttcctaata gcaggagtcg cataaggag agcgtcgacc      660
gatgccttg agagccttca acccagtcag ctcttcggg tgggcgcggg gcatgactat      720
cgtcgccgca cttatgactg tcttctttat catgcaactc gtaggacagg tgcgggcagc      780
gctctgggtc attttcggcg aggaaccgct tcgctggagc gcgacgatga tcggcctgtc      840
gcttgcggtt ttcggaatct tgcacgccct cgetcaagcc ttcgtcactg gtcccgccac      900
caaacgtttc ggcgagaagc aggcattat cgcggcatg gcggccgacg cgtgggcta      960
cgtcttgctg gcgttcgcga cgcgaggctg gatggccttc ccattatga ttcttctcgc     1020
ttccggcggc atcgggatgc ccgcgttga ggcoatgctg tcaggcagg tagatgaaga     1080
ccatcaggga cagcttcaag gatcgctcgc ggctcttacc agcctaactt cgatcactgg     1140
aaccgtgatc gtcacggcga tttatgcgc ctcggcgagc acatggaacg ggttggcatg     1200
gattgtaggc gccgcctat accttgtctg cctcccgcg ttgcgtcgcg gtgcatggag     1260
ccgggccacc tcgacctgaa tggaagccg cggcacctcg ctaacggatt caccactcca     1320
actagaaagc ttaacacaga aaaaagccg cacctgacag tgcgggcttt ttttttcgac     1380
cactgcag

```

<210> 27

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P9

<400> 27

agtaattcta gaaagcttaa cacagaaaaa agcccg

36

<210> 28

<211> 43

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P10

<400> 28

ctagtaggat ccctgcagtg gtcgaaaaaa aaagcccgca ctg

43

<210> 29

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi 1

<400> 29

ggaagatcta ttgaccttcg cacatcaacc tgg

33

<210> 30

<211> 30

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi 2

<400> 30

cggggtacct tgtaaatatt ttaaccggcc

30

<210> 31

<211> 56

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi 3

<400> 31
 ggaagatcta aggagacett aaatgagcga cacaacgata ctgcaaaaca gtaccc 56

<210> 32
 <211> 39
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi 4

<400> 32
 cgggggtacct cgtagagggt tactggcgct tatccagcg 39

<210> 33
 <211> 80
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi

<400> 33
 ctatctgatt gttttcctga gttttctggc aacgaaatga ggtcaacatt tgaagcctgc 60
 tttttataac taagttggca 80

<210> 34
 <211> 80
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi

<400> 34
 agcaggcata aaaaaaggcg gaccggagtc cgcctctgtt caggtgttaa cgctcaagtt 60
 agtataaaaa agctgaacga 80

<210> 35
 <211> 21
 <212> ADN
 <213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 35

gatatggatc atcacgcttc a

21

<210> 36

<211> 18

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi

<400> 36

catgctcggc tattcgac

18

<210> 37

<211> 1422

<212> ADN

<213> Escherichia coli

<220>

<221> CDS

<222> (301)..(1119)

<400> 37

tcggtcaggg	catggatgta	caaagcgcgc	aggagaagat	tggtcaggtg	gtggaagget	60
accgcaatac	gaaagaagtc	cgcgaactgg	cgcacgcctt	cggcgttgaa	atgccataa	120
cagaggaaat	ttatcaagta	ttatatggcg	gaaaaaacgc	gcgcgaggca	gcattgactt	180
tactaggtcg	tgcacgcaag	gacgagcgca	gcagccacta	accccaggga	acctttgtta	240
ccgctatgac	ccggcccgcg	cagaacgggc	cggtcattat	ctcatcgtgt	ggagtaagca	300
atg tgc tgt	gaa gaa ctg	gaa att gtc	tgg aac aat	att aaa gcc	gaa	348
Met Ser Cys	Glu Glu Leu	Glu Ile Val	Trp Asn Asn	Ile Lys Ala	Glu	
1	5	10	15			
gcc aga acg	ctg gcg gac	tgt gag cca	atg ctg gcc	agt ttt tac	cac	396
Ala Arg Thr	Leu Ala Asp	Cys Glu Pro	Met Leu Ala	Ser Phe Tyr	His	
	20	25	30			
gcg acg cta	ctc aag cac	gaa aac ctt	ggc agt gca	ctg agc tac	atg	444
Ala Thr Leu	Leu Lys His	Glu Asn Leu	Gly Ser Ala	Leu Ser Tyr	Met	
	35	40	45			
ctg gcg aac	aag ctg tca	tgc cca att	atg cct gct	att get atc	cgt	492
Leu Ala Asn	Lys Leu Ser	Ser Pro Ile	Met Pro Ala	Ile Ala Ile	Arg	
	50	55	60			
gaa gtg gtg	gaa gaa gcc	tac gcc gct	gac ccg gaa	atg atc gcc	tct	540
Glu Val Val	Glu Glu Ala	Tyr Ala Ala	Asp Pro Glu	Met Ile Ala	Ser	
65	70	75	80			

gcg gcc tgt gat att cag gcg gtg cgt acc cgc gac ccg gca gtc gat 588
 Ala Ala Cys Asp Ile Gln Ala Val Arg Thr Arg Asp Pro Ala Val Asp
 85 90 95
 aaa tac tca acc ccg ttg tta tac ctg aag ggt ttt cat gcc ttg cag 636
 Lys Tyr Ser Thr Pro Leu Leu Tyr Leu Lys Gly Phe His Ala Leu Gln
 100 105 110
 gcc tat cgc atc ggt cac tgg ttg tgg aat cag ggg cgt cgc gca ctg 684
 Ala Tyr Arg Ile Gly His Trp Leu Trp Asn Gln Gly Arg Arg Ala Leu
 115 120 125
 gca atc ttt ctg caa aac cag gtt tct gtg acg ttc cag gtc gat att 732
 Ala Ile Phe Leu Gln Asn Gln Val Ser Val Thr Phe Gln Val Asp Ile
 130 135 140
 cac ccg gca gca aaa att ggt cgc ggt atc atg ctt gac cac gcg aca 780
 His Pro Ala Ala Lys Ile Gly Arg Gly Ile Met Leu Asp His Ala Thr
 145 150 155 160
 ggc atc gtc gtt ggt gaa acg gcg gtg att gaa aac gac gta tcg att 828
 Gly Ile Val Val Gly Glu Thr Ala Val Ile Glu Asn Asp Val Ser Ile
 165 170 175
 ctg caa tct gtg acg ctt ggc ggt acg ggt aaa tct ggt ggt gac cgt 876
 Leu Gln Ser Val Thr Leu Gly Gly Thr Gly Lys Ser Gly Gly Asp Arg
 180 185 190
 cac ccg aaa att cgt gaa ggt gtg atg att ggc gcg ggc gcg aaa atc 924
 His Pro Lys Ile Arg Glu Gly Val Met Ile Gly Ala Gly Ala Lys Ile
 195 200 205
 ctc ggc aat att gaa gtt ggg cgc ggc gcg aag att ggc gca ggt tcc 972
 Leu Gly Asn Ile Glu Val Gly Arg Gly Ala Lys Ile Gly Ala Gly Ser
 210 215 220
 gtg gtg ctg caa ccg gtg ccg ccg cat acc acc gcc gct ggc gtt ccg 1020
 Val Val Leu Gln Pro Val Pro Pro His Thr Thr Ala Ala Gly Val Pro
 225 230 235 240
 gct cgt att gtc ggt aaa cca gac agc gat aag cca tca atg gat atg 1068
 Ala Arg Ile Val Gly Lys Pro Asp Ser Asp Lys Pro Ser Met Asp Met
 245 250 255
 gac cag cat ttc aac ggt att aac cat aca ttt gag tat ggg gat ggg 1116
 Asp Gln His Phe Asn Gly Ile Asn His Thr Phe Glu Tyr Gly Asp Gly
 260 265 270
 atc taatgtcctg tgatectgcc ggatgcgatg taatcatcta tccggcctac 1169
 Ile
 agtaactaat ctctcaatac cgtecccgga taacccaact gccgccaggc ttcatacacc 1229
 actacegaoa ccgcattgga cagattcatg ctgcggtgt ccggcaecat cggaatgcga 1289
 attttttggt cagegggcag ggcataaaga atgctcgtg gcaggccgcg tgtttccggg 1349
 ccgaacatca gataatgcc atectgatag ctacggcgc tgtgagcagg tgtaccttc 1409
 gtggtgaggg cga 1422

<210> 38
 <211> 273
 <212> PRT

<213> Escherichia coli

<400> 38

```

Met Ser Cys Glu Glu Leu Glu Ile Val Trp Asn Asn Ile Lys Ala Glu
1          5          10          15
Ala Arg Thr Leu Ala Asp Cys Glu Pro Met Leu Ala Ser Phe Tyr His
20          25          30
Ala Thr Leu Leu Lys His Glu Asn Leu Gly Ser Ala Leu Ser Tyr Met
35          40          45
Leu Ala Asn Lys Leu Ser Ser Pro Ile Met Pro Ala Ile Ala Ile Arg
50          55          60
Glu Val Val Glu Glu Ala Tyr Ala Ala Asp Pro Glu Met Ile Ala Ser
65          70          75          80
Ala Ala Cys Asp Ile Gln Ala Val Arg Thr Arg Asp Pro Ala Val Asp
85          90          95
Lys Tyr Ser Thr Pro Leu Leu Tyr Leu Lys Gly Phe His Ala Leu Gln
100         105         110
Ala Tyr Arg Ile Gly His Trp Leu Trp Asn Gln Gly Arg Arg Ala Leu
115         120         125
Ala Ile Phe Leu Gln Asn Gln Val Ser Val Thr Phe Gln Val Asp Ile
130         135         140
His Pro Ala Ala Lys Ile Gly Arg Gly Ile Met Leu Asp His Ala Thr
145         150         155         160
Gly Ile Val Val Gly Glu Thr Ala Val Ile Glu Asn Asp Val Ser Ile
165         170         175
Leu Gln Ser Val Thr Leu Gly Gly Thr Gly Lys Ser Gly Gly Asp Arg
180         185         190
His Pro Lys Ile Arg Glu Gly Val Met Ile Gly Ala Gly Ala Lys Ile
195         200         205
Leu Gly Asn Ile Glu Val Gly Arg Gly Ala Lys Ile Gly Ala Gly Ser
210         215         220
Val Val Leu Gln Pro Val Pro Pro His Thr Thr Ala Ala Gly Val Pro
225         230         235         240
Ala Arg Ile Val Gly Lys Pro Asp Ser Asp Lys Pro Ser Met Asp Met
245         250         255
Asp Gln His Phe Asn Gly Ile Asn His Thr Phe Glu Tyr Gly Asp Gly
260         265         270
Ile

```

<210> 39

<211> 1039

<212> ADN

<213> Escherichia coli

<220>

<221> CDS

<222> (201)..(839)

```

<400> 39
tgcggataac ggtagaattt ttacgcccagt attttgcoga gcactaaccg aatttttcac 60
tggagcatgc ctgattaatg attcaattat cgggttgata tcagggttaaa acctgatttt 120
ctcctttcta agccgctaca gattgggttag catattcacc tttaatcgcg catgatcgaa 180
agataattaa agaggttaat gtg ttc get gaa tac ggg gtt ctg aat tac tgg 233
      Val Phe Ala Glu Tyr Gly Val Leu Asn Tyr Trp
          1           5           10
acc tat ctg gtt ggg gcc att ttt att gtg ttg gtg cca ggg cca aat 281
Thr Tyr Leu Val Gly Ala Ile Phe Ile Val Leu Val Pro Gly Pro Asn
      15           20           25
acc ctg ttt gta ctc aaa aat agc gtc agt agc ggt atg aaa ggc ggt 329
Thr Leu Phe Val Leu Lys Asn Ser Val Ser Ser Gly Met Lys Gly Gly
      30           35           40
tat ctt gcg gcc tgc ggt gta ttt att ggc gat gcg gta ttg atg ttt 377
Tyr Leu Ala Ala Cys Gly Val Phe Ile Gly Asp Ala Val Leu Met Phe
      45           50           55
ctg gca tgg gct gga gtg gcg aca tta att aag acc acc ccg ata tta 425
Leu Ala Trp Ala Gly Val Ala Thr Leu Ile Lys Thr Thr Pro Ile Leu
      60           65           70           75
ttc aac att gta cgt tat ctt ggt gcg ttt tat ttg ctc tat ctg ggg 473
Phe Asn Ile Val Arg Tyr Leu Gly Ala Phe Tyr Leu Leu Tyr Leu Gly
      80           85           90
agt aaa att ctt tac gcg acc ctg aag ggt aaa aat agc gag gcc aaa 521
Ser Lys Ile Leu Tyr Ala Thr Leu Lys Gly Lys Asn Ser Glu Ala Lys
      95          100          105
tcc gat gag ccc caa tac ggt gct att ttt aaa cgc gcg tta att ttg 569
Ser Asp Glu Pro Gln Tyr Gly Ala Ile Phe Lys Arg Ala Leu Ile Leu
      110         115         120
agc ctg act aat ccg aaa gcc att ttg ttc tat gtg tcg ttt ttc gta 617
Ser Leu Thr Asn Pro Lys Ala Ile Leu Phe Tyr Val Ser Phe Phe Val
      125         130         135
cag ttt atc gat gtt aat gcc coa cat acg gga att tca ttc ttt att 665
Gln Phe Ile Asp Val Asn Ala Pro His Thr Gly Ile Ser Phe Phe Ile
      140         145         150         155
ctg gcg gcg acg ctg gaa ctg gtg agt ttc tgc tat ttg agc ttc ctg 713
Leu Ala Ala Thr Leu Glu Leu Val Ser Phe Cys Tyr Leu Ser Phe Leu
      160         165         170
att ata tct ggt gct ttt gtc acg cag tac ata cgt acc aaa aag aaa 761
Ile Ile Ser Gly Ala Phe Val Thr Gln Tyr Ile Arg Thr Lys Lys Lys
      175         180         185
ctg gct aaa gtt ggc aac tca ctg att ggt ttg atg ttc gtg ggt ttc 809
Leu Ala Lys Val Gly Asn Ser Leu Ile Gly Leu Met Phe Val Gly Phe
      190         195         200
get gcc cga ctg gcg acg ctg caa tcc tga tgctttcagc cgcggttgct 859
Ala Ala Arg Leu Ala Thr Leu Gln Ser
      205         210
gcgggcttcc catctataat cctccctgat tcttcgctga tatggtgcta aaaagtaacc 919
aataaatggt atttaaaatg caaattatca ggcgtaacct gaaacggetg gaataaaccg 979
ttttcagcgc attcaccgaa ggagggaataa ggatgcttca aatccacag aatttatattc 1039

```

<210> 40
 <211> 212
 <212> FRT
 <213> Escherichia coli

<400> 40
 Val Phe Ala Glu Tyr Gly Val Leu Asn Tyr Trp Thr Tyr Leu Val Gly
 1 5 10 15
 Ala Ile Phe Ile Val Leu Val Pro Gly Pro Asn Thr Leu Phe Val Leu
 20 25 30
 Lys Asn Ser Val Ser Ser Gly Met Lys Gly Gly Tyr Leu Ala Ala Cys
 35 40 45
 Gly Val Phe Ile Gly Asp Ala Val Leu Met Phe Leu Ala Trp Ala Gly
 50 55 60
 Val Ala Thr Leu Ile Lys Thr Thr Pro Ile Leu Phe Asn Ile Val Arg
 65 70 75 80
 Tyr Leu Gly Ala Phe Tyr Leu Leu Tyr Leu Gly Ser Lys Ile Leu Tyr
 85 90 95
 Ala Thr Leu Lys Gly Lys Asn Ser Glu Ala Lys Ser Asp Glu Pro Gln
 100 105 110
 Tyr Gly Ala Ile Phe Lys Arg Ala Leu Ile Leu Ser Leu Thr Asn Pro
 115 120 125
 Lys Ala Ile Leu Phe Tyr Val Ser Phe Phe Val Gln Phe Ile Asp Val
 130 135 140
 Asn Ala Pro His Thr Gly Ile Ser Phe Phe Ile Leu Ala Ala Thr Leu
 145 150 155 160
 Glu Leu Val Ser Phe Cys Tyr Leu Ser Phe Leu Ile Ile Ser Gly Ala
 165 170 175
 Phe Val Thr Gln Tyr Ile Arg Thr Lys Lys Lys Leu Ala Lys Val Gly
 180 185 190
 Asn Ser Leu Ile Gly Leu Met Phe Val Gly Phe Ala Ala Arg Leu Ala
 195 200 205
 Thr Leu Gln Ser
 210

<210> 41
 <211> 313
 <212> ADN
 <213> Escherichia coli

<400> 41
 gcatgcttcc aactgcgcta atgacgcagc tggacgaagg cgggattctc gtcttaccgc 60
 taggggagga gcaccagtat ttgaaacggg tgcgtcgtcg gggaggcgaa tttattatcg 120
 ataccgtgga ggccgtgcgc tttgtccctt tagtgaaggg tgagctggct taaaacgtga 180
 ggaaatacct ggatttttcc tggttatfff gccgcaggtc agcgtatcgt gaacatcttt 240
 tccagtgttc agtaggggtgc cttgcacggc aattatgtca ctggttatta accaatTTTT 300

cctgggggtc gac

313

<210> 42

<211> 312

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> đoạn khởi đầu nlpD đột biến

<400> 42

```

gc atgcttcc aactgeccta atgacgcagc tggacgaagg cgggattctc gtcttaccgc 60
taggggagga gcaccagtat ttgaaacggg tgcgtogtcg gggaggcgaa tttattatcg 120
ataccgtgga ggcgcgtgcgc tttgtccott tagtgaaggg tgagctggct taaaacgtga 180
ggaaatacct ggatttttcc tggttatttt gcgcgaggtc agcgtataat gaagatcttt 240
tccagtgttg acaagggtcc ttgcacgggt ataatgtcac tggttattaa ccaatttttc 300
ctgggggtcg ac 312

```

<210> 43

<211> 4403

<212> ADN

<213> *Pantoea ananatis*

<220>

<221> CDS

<222> (301)..(1311)

<220>

<221> CDS

<222> (1317)..(2147)

<220>

<221> CDS

<222> (2150)..(3022)

<400> 43

```

tacagcggaa cctggcacgg gccagaaggg ttgatgccgt cggatgacac actcaagagc 60
tggacgetca gcaaaattgt ctggcagcgc taagtctttt ttcacaccgc tcaaccgcag 120
ggcataaccg gccctgcgcg tecaattctg tttttcgtct gtcttttccc gcgccttat 180
gcctttttcg actttgaaat cagcaaacga tatataaaac cgttacgggt ttacgtgag 240
ttataaataa actgctgtat ctgcagatga gatctgcac aaatttcctc aggggtgaacc 300
atg acc tta cca gcg atg aaa aaa atc gtg agc gga ctc gca ctg tcg 348
Met Thr Leu Pro Ala Met Lys Lys Ile Val Ser Gly Leu Ala Leu Ser
1          5          10          15
ctg agt ctg gcc ggt gcc gca aac gcg acc gag ctg ttg aac agc tct 396
Leu Ser Leu Ala Gly Ala Ala Asn Ala Thr Glu Leu Leu Asn Ser Ser
20          25          30

```

tac gat gtc gca cgt gaa tta ttt gtc gcc ctg aat gcg cct ttt gtc	444
Tyr Asp Val Ala Arg Glu Leu Phe Val Ala Leu Asn Ala Pro Phe Val	
35 40 45	
agc cag tgg gat gcc agc cat cct gac gac aag ctg acc att aag atg	492
Ser Gln Trp Asp Ala Ser His Pro Asp Asp Lys Leu Thr Ile Lys Met	
50 55 60	
tcc cat gcc ggg tca tcc aaa cag gcg ctg gcg atc ctg caa ggc ctg	540
Ser His Ala Gly Ser Ser Lys Gln Ala Leu Ala Ile Leu Gln Gly Leu	
65 70 75 80	
cgt gcc gat gtg gtg acc tat aac cag gtc acc gat gtg cag gtg ctg	588
Arg Ala Asp Val Val Thr Tyr Asn Gln Val Thr Asp Val Gln Val Leu	
85 90 95	
cac gat aaa ggc aaa ctg atc cct gcc gac tgg caa acc cgc ctg ccg	636
His Asp Lys Gly Lys Leu Ile Pro Ala Asp Trp Gln Thr Arg Leu Pro	
100 105 110	
aat aac agt tcc ccg ttt tac tcc acc atg gcg ttc ctg gtg cgc aag	684
Asn Asn Ser Ser Pro Phe Tyr Ser Thr Met Ala Phe Leu Val Arg Lys	
115 120 125	
gga aac cca aag cag att cac gac tgg tcc gat tta acc cgt gac gat	732
Gly Asn Pro Lys Gln Ile His Asp Trp Ser Asp Leu Thr Arg Asp Asp	
130 135 140	
gtg aag ctg att ttt cct aat ccc aaa acc tcc ggc aac gga cgt tat	780
Val Lys Leu Ile Phe Pro Asn Pro Lys Thr Ser Gly Asn Gly Arg Tyr	
145 150 155 160	
acc tat ctt gct gcc tgg ggc gcc gcc agc aac act gac ggg ggc gat	828
Thr Tyr Leu Ala Ala Trp Gly Ala Ala Ser Asn Thr Asp Gly Gly Asp	
165 170 175	
cag gct aaa acc cgc gct ttt atg aca aaa ttt ctg aaa aat gtt gaa	876
Gln Ala Lys Thr Arg Ala Phe Met Thr Lys Phe Leu Lys Asn Val Glu	
180 185 190	
gtc ttc gat acc ggt ggc cga ggt gct acg acc acc ttt gct gaa cgc	924
Val Phe Asp Thr Gly Gly Arg Gly Ala Thr Thr Thr Phe Ala Glu Arg	
195 200 205	
ggt ctg ggc gat gtg ttg atc agt ttt gag tct gaa gtg aat aac atc	972
Gly Leu Gly Asp Val Leu Ile Ser Phe Glu Ser Glu Val Asn Asn Ile	
210 215 220	
cgc aac cag tac ggc aaa gac gac tac gaa gtc gtg gtg cct aaa acc	1020
Arg Asn Gln Tyr Gly Lys Asp Asp Tyr Glu Val Val Val Pro Lys Thr	
225 230 235 240	
gat att ctc gcg gag ttt ccc gtt gcc tgg gta gat aaa aac gtc gag	1068
Asp Ile Leu Ala Glu Phe Pro Val Ala Trp Val Asp Lys Asn Val Glu	
245 250 255	
cag aat aaa aca gcc gat gca gcg aaa gcc tat ctg acc tgg ctg tat	1116
Gln Asn Lys Thr Ala Asp Ala Ala Lys Ala Tyr Leu Thr Trp Leu Tyr	
260 265 270	
tct cct gcg gcg cag aaa att att acg gat ttc tat tac cgc gtg aac	1164
Ser Pro Ala Ala Gln Lys Ile Ile Thr Asp Phe Tyr Tyr Arg Val Asn	
275 280 285	
aat ccg cag tta atg gcg cag caa aaa gcc cgt ttt cct gcc acg aac	1212

Asn Pro Gln Leu Met Ala Gln Gln Lys Ala Arg Phe Pro Ala Thr Asn	
290 295 300	
ctg ttt cgt gtt gaa gac att ttt ggc ggc tgg gat aac gtg atg aaa	1260
Leu Phe Arg Val Glu Asp Ile Phe Gly Gly Trp Asp Asn Val Met Lys	
305 310 315 320	
acc cat ttc gcc agc ggt ggc gag cta gac cag tta tta gcg gcg ggg	1308
Thr His Phe Ala Ser Gly Gly Glu Leu Asp Gln Leu Leu Ala Ala Gly	
325 330 335	
egg tgate atg ttt gca gcc agc caa aaa cgc gtc ctg ccc ggt ttc ggt	1358
Arg Met Phe Ala Ala Ser Gln Lys Arg Val Leu Pro Gly Phe Gly	
340 345 350	
ctc agc ctg ggc acc agc ctg ctc ttt acc tgt ctg gtg ctg ctg ctg	1406
Leu Ser Leu Gly Thr Ser Leu Leu Phe Thr Cys Leu Val Leu Leu Leu	
355 360 365	
cca atc agc gca ctg att atg cag ctg tgc cag atg acg ttg cag caa	1454
Pro Ile Ser Ala Leu Ile Met Gln Leu Ser Gln Met Thr Leu Gln Gln	
370 375 380	
tac tgg gac gtg gtc acc aat ccg cag ctc atc gcg gcc tat aag gtc	1502
Tyr Trp Asp Val Val Thr Asn Pro Gln Leu Ile Ala Ala Tyr Lys Val	
385 390 395	
acg ctg ctg tgc gcc ggt gtg gcc tca ctg ttt aat gcc gta ttc ggc	1550
Thr Leu Leu Ser Ala Gly Val Ala Ser Leu Phe Asn Ala Val Phe Gly	
400 405 410 415	
atg tta atg gcg tgg atc tta acg cgt tac cgt ttt ccg ggc cgc acg	1598
Met Leu Met Ala Trp Ile Leu Thr Arg Tyr Arg Phe Pro Gly Arg Thr	
420 425 430	
ctg ctc gat ggt ctg atg gat ctg ccg ttt gcg ctg ccg acc gcg gtt	1646
Leu Leu Asp Gly Leu Met Asp Leu Pro Phe Ala Leu Pro Thr Ala Val	
435 440 445	
gct ggc ctg acg ctg gcc ggt ctg ttt tcc gtg aac gcc tgg tac gga	1694
Ala Gly Leu Thr Leu Ala Gly Leu Phe Ser Val Asn Gly Trp Tyr Gly	
450 455 460	
caa tgg ttc gcg cat ttt gat atc aag atc tcc tat acc tgg atc ggt	1742
Gln Trp Phe Ala His Phe Asp Ile Lys Ile Ser Tyr Thr Trp Ile Gly	
465 470 475	
atc gcg ctc gcg atg gcc ttc acc agt att ccg ttt gtg gtg cgt acc	1790
Ile Ala Leu Ala Met Ala Phe Thr Ser Ile Pro Phe Val Val Arg Thr	
480 485 490 495	
gtg cag ccg gtg ctg gaa gag ctg ggg cct gaa tat gag gaa gcg gct	1838
Val Gln Pro Val Leu Glu Glu Leu Gly Pro Glu Tyr Glu Glu Ala Ala	
500 505 510	
caa acg ctg ggc gcc acg ccc tgg cag agc ttc cgc ccg gtc gtt ctg	1886
Gln Thr Leu Gly Ala Thr Pro Trp Gln Ser Phe Arg Arg Val Val Leu	
515 520 525	
cct gaa gtg gca ccg gcc tta ctt gcg ggc acc gcg ctg tgc ttt acc	1934
Pro Glu Val Ala Pro Ala Leu Leu Ala Gly Thr Ala Leu Ser Phe Thr	
530 535 540	
cgc agc ctg ggc gag ttt ggt gcg gta atc ttt att gcc gcc aac atc	1982
Arg Ser Leu Gly Glu Phe Gly Ala Val Ile Phe Ile Ala Gly Asn Ile	

545	550	555	
gct tgg aaa acc gaa gtg acc tcg ctg atg atc ttc gtg cgc ctg cag			2030
Ala Trp Lys Thr Glu Val Thr Ser Leu Met Ile Phe Val Arg Leu Gln			
560	565	570	575
gag ttt gac tat ccg gca gcc agc gcc att gcc tcg gtc att ctg gcg			2078
Glu Phe Asp Tyr Pro Ala Ala Ser Ala Ile Ala Ser Val Ile Leu Ala			
580	585	590	
gca tca ctg ctg tta ctt ttc gct atc aat acc tta caa agc cgc ttt			2126
Ala Ser Leu Leu Leu Leu Phe Ala Ile Asn Thr Leu Gln Ser Arg Phe			
595	600	605	
ggc cgt cgt ctg gga gcc cat ta atg gca gag att tcg caa ctc aat			2173
Gly Arg Arg Leu Gly Gly His Met Ala Glu Ile Ser Gln Leu Asn			
610	615	620	
cat gcc gac cgc cag cct gtt aac tgg gcc aag tgg ctg ctt att ggt			2221
His Ala Asp Arg Gln Pro Val Asn Trp Ala Lys Trp Leu Leu Ile Gly			
625	630	635	
att ggt gcg ctg ata tcc ttg ctg ctg ctg gtc gtg ccg atg gtg tcc			2269
Ile Gly Ala Leu Ile Ser Leu Leu Leu Val Val Pro Met Val Ser			
640	645	650	
atc ttc tgg gag gcc ctg cat aaa gga ctg gcc gtc acc tta agt aat			2317
Ile Phe Trp Glu Ala Leu His Lys Gly Leu Gly Val Thr Leu Ser Asn			
655	660	665	670
ctg acc gac agc gac atg ctc cat gcc ata tgg ctc acg gtg ctg gtc			2365
Leu Thr Asp Ser Asp Met Leu His Ala Ile Trp Leu Thr Val Leu Val			
675	680	685	
gca ttg att acc gtg ccg gtg aat tta gtg ttc gcc acg ctg ctg gcc			2413
Ala Leu Ile Thr Val Pro Val Asn Leu Val Phe Gly Thr Leu Leu Ala			
690	695	700	
tgg ctg gtg aca cgc ttt acc ttt ccg gga cgt cag ctg ctt ttg acg			2461
Trp Leu Val Thr Arg Phe Thr Phe Pro Gly Arg Gln Leu Leu Leu Thr			
705	710	715	
ctg ttc gat att ccc ttt gcg gta tcg cct gtg gtc gcc ggt ctg atg			2509
Leu Phe Asp Ile Pro Phe Ala Val Ser Pro Val Val Ala Gly Leu Met			
720	725	730	
tat ctc ctg ttc tgg gcc att aac gcc ccg gcg gcc gcc tgg ctg gat			2557
Tyr Leu Leu Phe Trp Gly Ile Asn Gly Pro Ala Gly Gly Trp Leu Asp			
735	740	745	750
gcc cat aat att cag gtg atg ttc tcc tgg cct gcc atg gtg ctg gtc			2605
Ala His Asn Ile Gln Val Met Phe Ser Trp Pro Gly Met Val Leu Val			
755	760	765	
acc gtc ttc gtt acc tgt ccg ttt gtg gtg cgc gaa ctg gtg ccg gtg			2653
Thr Val Phe Val Thr Cys Pro Phe Val Val Arg Glu Leu Val Pro Val			
770	775	780	
atg ctg agc cag gcc agt cat gaa gat gaa gcc gcg gtg ctg tta ggt			2701
Met Leu Ser Gln Gly Ser His Glu Asp Glu Ala Ala Val Leu Leu Gly			
785	790	795	
gcc tcg gcc tgg cag atg ttc cgt cgc gtg acg ctg ccg aat att cgc			2749
Ala Ser Gly Trp Gln Met Phe Arg Arg Val Thr Leu Pro Asn Ile Arg			
800	805	810	

```

tgg gcc atg ctg tat ggc gtc gtg ctg acc aac gcc cgc gcg att ggt      2797
Trp Ala Met Leu Tyr Gly Val Val Leu Thr Asn Ala Arg Ala Ile Gly
815                      820                      825                      830
gag ttt gcc gcg gtt tcc gtg gtt tcg ggt tct att cgc ggt gaa acc      2845
Glu Phe Gly Ala Val Ser Val Val Ser Gly Ser Ile Arg Gly Glu Thr
                        835                      840                      845
tac act tta ccg ctt cag gtt gaa tta ctg cat cag gat tac aac acg      2893
Tyr Thr Leu Pro Leu Gln Val Glu Leu Leu His Gln Asp Tyr Asn Thr
                        850                      855                      860
gtg gcc gcc ttt act gcc gca gcc tta ctg acc gtg atg gca atc gtg      2941
Val Gly Ala Phe Thr Ala Ala Ala Leu Leu Thr Val Met Ala Ile Val
                        865                      870                      875
acg ctg ttt ctg aaa agc att gtg caa tgg cgt tta gag caa cag cae      2989
Thr Leu Phe Leu Lys Ser Ile Val Gln Trp Arg Leu Glu Gln Gln His
                        880                      885                      890
aaa cgc ctg caa ctg gag gac aat cat gag cat tgagattaac cagatcaaca      3042
Lys Arg Leu Gln Leu Glu Asp Asn His Glu His
895                      900                      905
aatcctttgg tcgcacagcg gtgctgaacg atatctcaact ggatattcct tctggccaga      3102
tggtggcett actggggccg tccggttccg gtaaaaccac gctgctgcgc atcattgctg      3162
gactggaaca tcagaacagc ggtcagattc gttttcacga ccacgatgtc agccgcctgc      3222
acgcccgga tcgccaggtc ggatttgtct tccagcacta tgcctgttcc cgtcatatga      3282
cggctcttga caatattgcc ttggccctga ccgtgctgcc gcgcgctgag cgtccgtcca      3342
gtgcggaaat taaaaaacgc gtcacgcgcc tgcctggagat ggtgcagctt tccatctgg      3402
cgaaccgttt cccggccag ctttcgggag ggcagaagca gcgcgtgcgc ctggcaagag      3462
ccctggccgt ggaaccgcaa atcctgttgc tggatgagcc ctttgggtgc ctggaogctc      3522
aggtgcgtaa agagctgcgc cgttgggttac gtcagctgca cgaagaattg aagttcacca      3582
gcgtgttcgt caccacgat caggaagagg cgatggaagt ggccgatcgc gtggtggtga      3642
tgagccaggg cagcatcgaa cagggtggga cgcgggatga agtctggcgc gatcccgcca      3702
cgcgcttcgt gctggaattc ctgggtgagg ttaaccgctt cgacggtgaa gtgcatggtt      3762
ctcagttcca tgtcggggcg caccactggc cgttaggcta tacctctgca catcagggcg      3822
cggtcgatct gtctctgcgc ccgtgggaaa tcgacgttcc gcgcagaagt agcctggaaa      3882
cgccgctgcc cgttcaggtc ttagaagtga gtctcgtgg tcacttctgg cagctggtgg      3942
tgcagccaac gggatggcag agcgagccct tctcgttgtt ctttgacggt gaacagaccg      4002
cgccgttgcc cggcgagcgc ctgttctgtg ggctgcagca ggccagactg taccagggcg      4062
cgacaccgtt acggggoggtt gcccttgcac acagcgccctg atagggttgag tgaatgttaa      4122
acgcccgga ggcgttcccg cgatccgggc tttttaatgg caagggttgt aaactgtaga      4182
cctgataaga cgcgcaagcg tcgcatacagg caacaccacg tatggataga gatcgtgagt      4242
acattagaac aaacaatagg caatacgct ctggtgaagt tgcagcgaat ggggocggat      4302
aaccgagtg aagtgtggtt aaaactggaa ggcaataacc cggcaggttc ggtgaaagat      4362
cgtgcggcac ttctgatgat cgtogaggcg gaaaagcgcg g      4403

```

<210> 44

<211> 337

<212> PRT

<213> Pantoea ananatis

<400> 44

Met Thr Leu Pro Ala Met Lys Lys Ile Val Ser Gly Leu Ala Leu Ser
 1 5 10 15
 Leu Ser Leu Ala Gly Ala Ala Asn Ala Thr Glu Leu Leu Asn Ser Ser
 20 25 30
 Tyr Asp Val Ala Arg Glu Leu Phe Val Ala Leu Asn Ala Pro Phe Val
 35 40 45
 Ser Gln Trp Asp Ala Ser His Pro Asp Asp Lys Leu Thr Ile Lys Met
 50 55 60
 Ser His Ala Gly Ser Ser Lys Gln Ala Leu Ala Ile Leu Gln Gly Leu
 65 70 75 80
 Arg Ala Asp Val Val Thr Tyr Asn Gln Val Thr Asp Val Gln Val Leu
 85 90 95
 His Asp Lys Gly Lys Leu Ile Pro Ala Asp Trp Gln Thr Arg Leu Pro
 100 105 110
 Asn Asn Ser Ser Pro Phe Tyr Ser Thr Met Ala Phe Leu Val Arg Lys
 115 120 125
 Gly Asn Pro Lys Gln Ile His Asp Trp Ser Asp Leu Thr Arg Asp Asp
 130 135 140
 Val Lys Leu Ile Phe Pro Asn Pro Lys Thr Ser Gly Asn Gly Arg Tyr
 145 150 155 160
 Thr Tyr Leu Ala Ala Trp Gly Ala Ala Ser Asn Thr Asp Gly Gly Asp
 165 170 175
 Gln Ala Lys Thr Arg Ala Phe Met Thr Lys Phe Leu Lys Asn Val Glu
 180 185 190
 Val Phe Asp Thr Gly Gly Arg Gly Ala Thr Thr Thr Phe Ala Glu Arg
 195 200 205
 Gly Leu Gly Asp Val Leu Ile Ser Phe Glu Ser Glu Val Asn Asn Ile
 210 215 220
 Arg Asn Gln Tyr Gly Lys Asp Asp Tyr Glu Val Val Val Pro Lys Thr
 225 230 235 240
 Asp Ile Leu Ala Glu Phe Pro Val Ala Trp Val Asp Lys Asn Val Glu
 245 250 255
 Gln Asn Lys Thr Ala Asp Ala Ala Lys Ala Tyr Leu Thr Trp Leu Tyr
 260 265 270
 Ser Pro Ala Ala Gln Lys Ile Ile Thr Asp Phe Tyr Tyr Arg Val Asn
 275 280 285
 Asn Pro Gln Leu Met Ala Gln Gln Lys Ala Arg Phe Pro Ala Thr Asn
 290 295 300
 Leu Phe Arg Val Glu Asp Ile Phe Gly Gly Trp Asp Asn Val Met Lys
 305 310 315 320
 Thr His Phe Ala Ser Gly Gly Glu Leu Asp Gln Leu Leu Ala Ala Gly
 325 330 335
 Arg

<210> 45
 <211> 277
 <212> PRT

<213> Pantoea ananatis

<400> 45

Met Phe Ala Ala Ser Gln Lys Arg Val Leu Pro Gly Phe Gly Leu Ser
 1 5 10 15
 Leu Gly Thr Ser Leu Leu Phe Thr Cys Leu Val Leu Leu Leu Pro Ile
 20 25 30
 Ser Ala Leu Ile Met Gln Leu Ser Gln Met Thr Leu Gln Gln Tyr Trp
 35 40 45
 Asp Val Val Thr Asn Pro Gln Leu Ile Ala Ala Tyr Lys Val Thr Leu
 50 55 60
 Leu Ser Ala Gly Val Ala Ser Leu Phe Asn Ala Val Phe Gly Met Leu
 65 70 75 80
 Met Ala Trp Ile Leu Thr Arg Tyr Arg Phe Pro Gly Arg Thr Leu Leu
 85 90 95
 Asp Gly Leu Met Asp Leu Pro Phe Ala Leu Pro Thr Ala Val Ala Gly
 100 105 110
 Leu Thr Leu Ala Gly Leu Phe Ser Val Asn Gly Trp Tyr Gly Gln Trp
 115 120 125
 Phe Ala His Phe Asp Ile Lys Ile Ser Tyr Thr Trp Ile Gly Ile Ala
 130 135 140
 Leu Ala Met Ala Phe Thr Ser Ile Pro Phe Val Val Arg Thr Val Gln
 145 150 155 160
 Pro Val Leu Glu Glu Leu Gly Pro Glu Tyr Glu Glu Ala Ala Gln Thr
 165 170 175
 Leu Gly Ala Thr Pro Trp Gln Ser Phe Arg Arg Val Val Leu Pro Glu
 180 185 190
 Val Ala Pro Ala Leu Leu Ala Gly Thr Ala Leu Ser Phe Thr Arg Ser
 195 200 205
 Leu Gly Glu Phe Gly Ala Val Ile Phe Ile Ala Gly Asn Ile Ala Trp
 210 215 220
 Lys Thr Glu Val Thr Ser Leu Met Ile Phe Val Arg Leu Gln Glu Phe
 225 230 235 240
 Asp Tyr Pro Ala Ala Ser Ala Ile Ala Ser Val Ile Leu Ala Ala Ser
 245 250 255
 Leu Leu Leu Leu Phe Ala Ile Asn Thr Leu Gln Ser Arg Phe Gly Arg
 260 265 270
 Arg Leu Gly Gly His
 275

<210> 46

<211> 291

<212> PRT

<213> Pantoea ananatis

<400> 46

Met Ala Glu Ile Ser Gln Leu Asn His Ala Asp Arg Gln Pro Val Asn
 1 5 10 15

Trp Ala Lys Trp Leu Leu Ile Gly Ile Gly Ala Leu Ile Ser Leu Leu
 20 25 30
 Leu Leu Val Val Pro Met Val Ser Ile Phe Trp Glu Ala Leu His Lys
 35 40 45
 Gly Leu Gly Val Thr Leu Ser Asn Leu Thr Asp Ser Asp Met Leu His
 50 55 60
 Ala Ile Trp Leu Thr Val Leu Val Ala Leu Ile Thr Val Pro Val Asn
 65 70 75 80
 Leu Val Phe Gly Thr Leu Leu Ala Trp Leu Val Thr Arg Phe Thr Phe
 85 90 95
 Pro Gly Arg Gln Leu Leu Leu Thr Leu Phe Asp Ile Pro Phe Ala Val
 100 105 110
 Ser Pro Val Val Ala Gly Leu Met Tyr Leu Leu Phe Trp Gly Ile Asn
 115 120 125
 Gly Pro Ala Gly Gly Trp Leu Asp Ala His Asn Ile Gln Val Met Phe
 130 135 140
 Ser Trp Pro Gly Met Val Leu Val Thr Val Phe Val Thr Cys Pro Phe
 145 150 155 160
 Val Val Arg Glu Leu Val Pro Val Met Leu Ser Gln Gly Ser His Glu
 165 170 175
 Asp Glu Ala Ala Val Leu Leu Gly Ala Ser Gly Trp Gln Met Phe Arg
 180 185 190
 Arg Val Thr Leu Pro Asn Ile Arg Trp Ala Met Leu Tyr Gly Val Val
 195 200 205
 Leu Thr Asn Ala Arg Ala Ile Gly Glu Phe Gly Ala Val Ser Val Val
 210 215 220
 Ser Gly Ser Ile Arg Gly Glu Thr Tyr Thr Leu Pro Leu Gln Val Glu
 225 230 235 240
 Leu Leu His Gln Asp Tyr Asn Thr Val Gly Ala Phe Thr Ala Ala Ala
 245 250 255
 Leu Leu Thr Val Met Ala Ile Val Thr Leu Phe Leu Lys Ser Ile Val
 260 265 270
 Gln Trp Arg Leu Glu Gln Gln His Lys Arg Leu Gln Leu Glu Asp Asn
 275 280 285
 His Glu His
 290

<210> 47

<211> 1403

<212> ADN

<213> *Pantoea ananatis*

<220>

<221> CDS

<222> (15)..(1103)

<400> 47

actggaggac aatc atg agc att gag att aac cag atc aac aaa tcc ttt

50

				Met	Ser	Ile	Glu	Ile	Asn	Gln	Ile	Asn	Lys	Ser	Phe	
				1				5					10			
ggc	cgc	aca	gcg	gtg	ctg	aac	gat	atc	tca	ctg	gat	att	cct	tct	ggc	98
Gly	Arg	Thr	Ala	Val	Leu	Asn	Asp	Ile	Ser	Leu	Asp	Ile	Pro	Ser	Gly	
			15				20					25				
cag	atg	gtg	gcc	tta	ctg	ggg	ccg	tcc	ggc	tcc	ggc	aaa	acc	acg	ctg	146
Gln	Met	Val	Ala	Leu	Leu	Gly	Pro	Ser	Gly	Ser	Gly	Lys	Thr	Thr	Leu	
			30			35					40					
ctg	cgc	atc	att	gct	gga	ctg	gaa	cat	cag	aac	agc	ggc	cag	att	cgt	194
Leu	Arg	Ile	Ile	Ala	Gly	Leu	Glu	His	Gln	Asn	Ser	Gly	Gln	Ile	Arg	
45				50				55				60				
ttt	cac	gac	cac	gat	gtc	agc	cgc	ctg	cac	gcc	cgc	gat	cgc	cag	gtc	242
Phe	His	Asp	His	Asp	Val	Ser	Arg	Leu	His	Ala	Arg	Asp	Arg	Gln	Val	
				65				70				75				
gga	ttt	gtc	ttc	cag	cac	tat	gcg	ctg	ttc	cgt	cat	atg	acg	gtc	ttc	290
Gly	Phe	Val	Phe	Gln	His	Tyr	Ala	Leu	Phe	Arg	His	Met	Thr	Val	Phe	
			80				85				90					
gac	aat	att	gcc	ttt	ggc	ctg	acc	gtg	ctg	ccg	cgc	cgt	gag	cgt	ccg	338
Asp	Asn	Ile	Ala	Phe	Gly	Leu	Thr	Val	Leu	Pro	Arg	Arg	Glu	Arg	Pro	
			95			100					105					
tcc	agt	gcg	gaa	att	aaa	aaa	cgc	gtc	acg	cgc	ctg	ctg	gag	atg	gtg	386
Ser	Ser	Ala	Glu	Ile	Lys	Lys	Arg	Val	Thr	Arg	Leu	Leu	Glu	Met	Val	
			110			115					120					
cag	ctt	tcc	cat	ctg	gcg	aac	cgt	ttc	ccg	gcc	cag	ctt	tcc	gga	ggg	434
Gln	Leu	Ser	His	Leu	Ala	Asn	Arg	Phe	Pro	Ala	Gln	Leu	Ser	Gly	Gly	
125				130				135				140				
cag	aag	cag	cgc	gtc	gcg	ctg	gca	aga	gcc	ctg	gcc	gtg	gaa	ccg	caa	482
Gln	Lys	Gln	Arg	Val	Ala	Leu	Ala	Arg	Ala	Leu	Ala	Val	Glu	Pro	Gln	
				145			150				155					
atc	ctg	ttg	ctg	gat	gag	ccc	ttt	ggc	gcg	ctg	gac	gct	cag	gtg	cgt	530
Ile	Leu	Leu	Leu	Asp	Glu	Pro	Phe	Gly	Ala	Leu	Asp	Ala	Gln	Val	Arg	
			160				165				170					
aaa	gag	ctg	cgc	cgt	tgg	tta	cgt	cag	ctg	cac	gaa	gaa	ttg	aag	ttc	578
Lys	Glu	Leu	Arg	Arg	Trp	Leu	Arg	Gln	Leu	His	Glu	Glu	Leu	Lys	Phe	
			175			180					185					
acc	agc	gtg	ttc	gtc	acc	cac	gat	cag	gaa	gag	gcg	atg	gaa	gtg	gcc	626
Thr	Ser	Val	Phe	Val	Thr	His	Asp	Gln	Glu	Glu	Ala	Met	Glu	Val	Ala	
			190			195					200					
gat	cgc	gtg	gtg	gtg	atg	agc	cag	ggc	agc	atc	gaa	cag	gtg	ggg	acg	674
Asp	Arg	Val	Val	Val	Met	Ser	Gln	Gly	Ser	Ile	Glu	Gln	Val	Gly	Thr	
205				210				215				220				
ccg	gat	gaa	gtc	tgg	cgc	gat	ccc	gcc	acg	cgc	ttc	gtg	ctg	gaa	ttc	722
Pro	Asp	Glu	Val	Trp	Arg	Asp	Pro	Ala	Thr	Arg	Phe	Val	Leu	Glu	Phe	
				225			230				235					
ctg	ggc	gag	gtt	aac	cgc	ttc	gac	ggc	gaa	gtg	cat	ggc	tct	cag	ttc	770
Leu	Gly	Glu	Val	Asn	Arg	Phe	Asp	Gly	Glu	Val	His	Gly	Ser	Gln	Phe	
			240			245					250					
cat	gtc	ggg	gcg	cac	cac	tgg	ccg	tta	ggc	tat	acc	tct	gca	cat	cag	818
His	Val	Gly	Ala	His	His	Trp	Pro	Leu	Gly	Tyr	Thr	Ser	Ala	His	Gln	

```

      255      260      265
ggc ggc gtc gat ctg ttc ctg cgc ccg tgg gaa atc gac gtt tgc cgc      866
Gly Ala Val Asp Leu Phe Leu Arg Pro Trp Glu Ile Asp Val Ser Arg
      270      275      280
aga agt agc ctg gaa acg ccg ctg ccc gtt cag gtc tta gaa gtg agt      914
Arg Ser Ser Leu Glu Thr Pro Leu Pro Val Gln Val Leu Glu Val Ser
285      290      295      300
cct cgt ggt cac ttc tgg cag ctg gtg gtg cag cca acg gga tgg cag      962
Pro Arg Gly His Phe Trp Gln Leu Val Val Gln Pro Thr Gly Trp Gln
      305      310      315
agc gag ccc ttc tgc ctg gtc ttt gac ggt gaa cag acc ggc ccg ttg      1010
Ser Glu Pro Phe Ser Leu Val Phe Asp Gly Glu Gln Thr Ala Pro Leu
      320      325      330
cgc ggc gag cgc ctg ttc gtg ggg ctg cag cag gcc aga ctg tac cag      1058
Arg Gly Glu Arg Leu Phe Val Gly Leu Gln Gln Ala Arg Leu Tyr Gln
      335      340      345
ggc ggc aca ccg tta cgg gcg gtt gcc ttt gca cac agc gcc tga      1103
Gly Ala Thr Pro Leu Arg Ala Val Ala Phe Ala His Ser Ala
      350      355      360
taggttgagt gaatgttaaa cgcccgagg cgttccgcg gatccgggct ttttaatggc      1163
aaggtttgta acctgtagac ctgataagac gcgcaagcgt cgcacagggc aacaccacgt      1223
atggatagag atcgtgagta cattagaaca aacaataggc aatacgctc tgggtgaagtt      1283
gcagcgaatg gggccggata acggcagtg agtgtgggta aaactggaag gcaataaccc      1343
ggcaggttcg gtgaaagatc gtgcggcact ttcgatgate gtcgagggcg aaaagcgcgg      1403

```

<210> 48

<211> 362

<212> PRT

<213> *Pantoea ananatis*

<400> 48

```

Met Ser Ile Glu Ile Asn Gln Ile Asn Lys Ser Phe Gly Arg Thr Ala
1      5      10      15
Val Leu Asn Asp Ile Ser Leu Asp Ile Pro Ser Gly Gln Met Val Ala
      20      25      30
Leu Leu Gly Pro Ser Gly Ser Gly Lys Thr Thr Leu Leu Arg Ile Ile
      35      40      45
Ala Gly Leu Glu His Gln Asn Ser Gly Gln Ile Arg Phe His Asp His
      50      55      60
Asp Val Ser Arg Leu His Ala Arg Asp Arg Gln Val Gly Phe Val Phe
65      70      75      80
Gln His Tyr Ala Leu Phe Arg His Met Thr Val Phe Asp Asn Ile Ala
      85      90      95
Phe Gly Leu Thr Val Leu Pro Arg Arg Glu Arg Pro Ser Ser Ala Glu
      100      105      110
Ile Lys Lys Arg Val Thr Arg Leu Leu Glu Met Val Gln Leu Ser His
      115      120      125
Leu Ala Asn Arg Phe Pro Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg

```

130	135	140
Val Ala Leu Ala Arg Ala Leu Ala Val Glu Pro Gln Ile Leu Leu Leu		
145	150	155
Asp Glu Pro Phe Gly Ala Leu Asp Ala Gln Val Arg Lys Glu Leu Arg		160
	165	170
Arg Trp Leu Arg Gln Leu His Glu Glu Leu Lys Phe Thr Ser Val Phe		175
	180	185
Val Thr His Asp Gln Glu Glu Ala Met Glu Val Ala Asp Arg Val Val		190
	195	200
Val Met Ser Gln Gly Ser Ile Glu Gln Val Gly Thr Pro Asp Glu Val		205
	210	215
Trp Arg Asp Pro Ala Thr Arg Phe Val Leu Glu Phe Leu Gly Glu Val		220
225	230	235
Asn Arg Phe Asp Gly Glu Val His Gly Ser Gln Phe His Val Gly Ala		240
	245	250
His His Trp Pro Leu Gly Tyr Thr Ser Ala His Gln Gly Ala Val Asp		255
	260	265
Leu Phe Leu Arg Pro Trp Glu Ile Asp Val Ser Arg Arg Ser Ser Leu		270
	275	280
Glu Thr Pro Leu Pro Val Gln Val Leu Glu Val Ser Pro Arg Gly His		285
	290	295
Phe Trp Gln Leu Val Val Gln Pro Thr Gly Trp Gln Ser Glu Pro Phe		300
305	310	315
Ser Leu Val Phe Asp Gly Glu Gln Thr Ala Pro Leu Arg Gly Glu Arg		320
	325	330
Leu Phe Val Gly Leu Gln Gln Ala Arg Leu Tyr Gln Gly Ala Thr Pro		335
	340	345
Leu Arg Ala Val Ala Phe Ala His Ser Ala		350
	355	360

<210> 49
 <211> 1779
 <212> ADN
 <213> Pantoea ananatis

<220>
 <221> CDS
 <222> (301)..(1536)

<400> 49
 gtcaaaaaccg tcaaaaaata aagcacccgg cgcacaacag ggcgcgccgct ttttttgatt 60
 taaaaaaaact tttctcacca ggctgaaatt tggtagactta tgtcacataa ccgtcatcgg 120
 cagcggggttc gttcttctcg atgcggccaa cccacgattt tgtctggcaa agtaoqtoct 180
 ctgagccctg ccattgetggc ggtcaggcaa tcgtttgtat tgccgcaggc gatttttttg 240
 atatttttgac agacggetga ctgcgttcag tcctcgttga attctgaata ggggttgggaa 300
 atg gca aag gta tca ctg gaa aaa gac aaa att aag ttc ctg ctg gtg 348
 Met Ala Lys Val Ser Leu Glu Lys Asp Lys Ile Lys Phe Leu Leu Val
 1 5 10 15

gaa ggt gtc cat cag agc gcg ctg gaa aat ctt cgt gct gca ggt tac	396
Glu Gly Val His Gln Ser Ala Leu Glu Asn Leu Arg Ala Ala Gly Tyr	
20 25 30	
acc aat att gaa ttc cac aaa ggc gca ctg gat gcc gag gcg tta aaa	444
Thr Asn Ile Glu Phe His Lys Gly Ala Leu Asp Ala Glu Ala Leu Lys	
35 40 45	
gct tcc gct cgc gat gcg cat ttt atc ggt atc cgt tcc cgt tcc caa	492
Ala Ser Ala Arg Asp Ala His Phe Ile Gly Ile Arg Ser Arg Ser Gln	
50 55 60	
ctg acc gaa gag att ttt gcc gct gca gaa aaa ctg gta gcg gtg ggc	540
Leu Thr Glu Glu Ile Phe Ala Ala Ala Glu Lys Leu Val Ala Val Gly	
65 70 75 80	
tgt ttc tgt atc gga acg aat cag gtt gat tta aat gcc gca gcg aaa	588
Cys Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asn Ala Ala Ala Lys	
85 90 95	
cgc ggt atc ccg gtt ttt aac gca cct ttc tca aat acg cgc tct gtg	636
Arg Gly Ile Pro Val Phe Asn Ala Pro Phe Ser Asn Thr Arg Ser Val	
100 105 110	
gcc gag ctg gtt att ggc gag atg ctg ctg atg ctg cgc ggt gtt ccg	684
Ala Glu Leu Val Ile Gly Glu Met Leu Leu Met Leu Arg Gly Val Pro	
115 120 125	
gaa gcg aat gcc aaa gcg cac cgt ggt atc tgg aat aaa atc gcc aaa	732
Glu Ala Asn Ala Lys Ala His Arg Gly Ile Trp Asn Lys Ile Ala Lys	
130 135 140	
ggc tct ttt gaa gcg cgc ggt aaa aag ctg ggt atc att ggc tat ggc	780
Gly Ser Phe Glu Ala Arg Gly Lys Lys Leu Gly Ile Ile Gly Tyr Gly	
145 150 155 160	
cat atc ggt atg caa ctg ggc gtg ctg gca gaa agt ctg ggc atg cac	828
His Ile Gly Met Gln Leu Gly Val Leu Ala Glu Ser Leu Gly Met His	
165 170 175	
gtt tac ttc tat gac atc gaa aac aag ctg ccg ttg ggc aac gca tca	876
Val Tyr Phe Tyr Asp Ile Glu Asn Lys Leu Pro Leu Gly Asn Ala Ser	
180 185 190	
cag gtt cgt agc ctg acg cag ttg cta aat atg agt gac gtt gtc agc	924
Gln Val Arg Ser Leu Thr Gln Leu Leu Asn Met Ser Asp Val Val Ser	
195 200 205	
ctg cat gtc ccg gaa aac gcc tct acg caa aat atg att tct gcc aat	972
Leu His Val Pro Glu Thr Ala Ser Thr Gln Asn Met Ile Ser Ala Asn	
210 215 220	
gag ctg gct cag atg aag cct ggc ggc ctg ctg ata aat gcc tca cgc	1020
Glu Leu Ala Gln Met Lys Pro Gly Gly Leu Leu Ile Asn Ala Ser Arg	
225 230 235 240	
ggc acc gtg gta gat att cct gct ttg tgc gaa gcg ctg gcc agc aag	1068
Gly Thr Val Val Asp Ile Pro Ala Leu Cys Glu Ala Leu Ala Ser Lys	
245 250 255	
cag gtt ggt ggc gct gcg att gat gtg ttc cct gta gag ccg gcg acc	1116
Gln Val Gly Gly Ala Ala Ile Asp Val Phe Pro Val Glu Pro Ala Thr	
260 265 270	
aac agc gat ccg ttt gtt tcc cca ctg agc gaa ttc gac aac gtt atc	1164

Asn Ser Asp Pro Phe Val Ser Pro Leu Ser Glu Phe Asp Asn Val Ile	
275	280 285
ctg acg ccg cac atc ggg gga tgg acg gaa gaa gct cag gag aat atc	1212
Leu Thr Pro His Ile Gly Gly Ser Thr Glu Glu Ala Gln Glu Asn Ile	
290	295 300
ggg att gaa gtc gcg ggc aag ctg gcg aaa tat tgg gat aac ggt tca	1260
Gly Ile Glu Val Ala Gly Lys Leu Ala Lys Tyr Ser Asp Asn Gly Ser	
305	310 315 320
acg ctg tcc gcc gtc aat ttc ccg gaa gtg tca ttg ccg atg cac ggc	1308
Thr Leu Ser Ala Val Asn Phe Pro Glu Val Ser Leu Pro Met His Gly	
325	330 335
att agc gcc agt cgt ctg ctg cat att cac gaa aac cgt ccg ggc gtt	1356
Ile Ser Ala Ser Arg Leu Leu His Ile His Glu Asn Arg Pro Gly Val	
340	345 350
ctc acc gcg atc aac gag att ttc gct gaa caa ggc atc aac att gcc	1404
Leu Thr Ala Ile Asn Gln Ile Phe Ala Glu Gln Gly Ile Asn Ile Ala	
355	360 365
gct cag tac ctg caa acc tct ccg atg atg ggt tat gtg gtc atc gac	1452
Ala Gln Tyr Leu Gln Thr Ser Pro Met Met Gly Tyr Val Val Ile Asp	
370	375 380
att gat gct gag cac gaa ctg gca gag aaa gct ctg caa ctg atg aag	1500
Ile Asp Ala Glu His Glu Leu Ala Glu Lys Ala Leu Gln Leu Met Lys	
385	390 395 400
gcg att ccg gga acg att cgc gcc cgc ctg ctt tac tgatcccacg	1546
Ala Ile Pro Gly Thr Ile Arg Ala Arg Leu Leu Tyr	
405	410
ctgtcaccta cccgggcaca caagcatgccc cgggtttatt catcccatag ccacagtttt	1606
gatggcgctca gcacggcccg caaaggaatg tcccacgccg ctgtaggcag cgcgtcaacc	1666
cgtgacagt catgagcgt gccacccggt aaaaacccat gctgtttcca gttctgtaag	1726
gtgcgatcgt agaagccgcc cccatttct aaacgctgtc cggcgcgac gaa	1779

<210> 50

<211> 412

<212> PRT

<213> Pantoea ananatis

<400> 50

Met Ala Lys Val Ser Leu Glu Lys Asp Lys Ile Lys Phe Leu Leu Val	
1	5 10 15
Glu Gly Val His Gln Ser Ala Leu Glu Asn Leu Arg Ala Ala Gly Tyr	
20	25 30
Thr Asn Ile Glu Phe His Lys Gly Ala Leu Asp Ala Glu Ala Leu Lys	
35	40 45
Ala Ser Ala Arg Asp Ala His Phe Ile Gly Ile Arg Ser Arg Ser Gln	
50	55 60
Leu Thr Glu Glu Ile Phe Ala Ala Ala Glu Lys Leu Val Ala Val Gly	
65	70 75 80
Cys Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asn Ala Ala Ala Lys	

			85					90				95			
Arg	Gly	Ile	Pro	Val	Phe	Asn	Ala	Pro	Phe	Ser	Asn	Thr	Arg	Ser	Val
			100					105					110		
Ala	Glu	Leu	Val	Ile	Gly	Glu	Met	Leu	Leu	Met	Leu	Arg	Gly	Val	Pro
			115					120					125		
Glu	Ala	Asn	Ala	Lys	Ala	His	Arg	Gly	Ile	Trp	Asn	Lys	Ile	Ala	Lys
			130					135					140		
Gly	Ser	Phe	Glu	Ala	Arg	Gly	Lys	Lys	Leu	Gly	Ile	Ile	Gly	Tyr	Gly
145					150					155					160
His	Ile	Gly	Met	Gln	Leu	Gly	Val	Leu	Ala	Glu	Ser	Leu	Gly	Met	His
				165						170					175
Val	Tyr	Phe	Tyr	Asp	Ile	Glu	Asn	Lys	Leu	Pro	Leu	Gly	Asn	Ala	Ser
			180					185						190	
Gln	Val	Arg	Ser	Leu	Thr	Gln	Leu	Leu	Asn	Met	Ser	Asp	Val	Val	Ser
			195					200					205		
Leu	His	Val	Pro	Glu	Thr	Ala	Ser	Thr	Gln	Asn	Met	Ile	Ser	Ala	Asn
			210					215					220		
Glu	Leu	Ala	Gln	Met	Lys	Pro	Gly	Gly	Leu	Leu	Ile	Asn	Ala	Ser	Arg
225					230						235				240
Gly	Thr	Val	Val	Asp	Ile	Pro	Ala	Leu	Cys	Glu	Ala	Leu	Ala	Ser	Lys
				245					250						255
Gln	Val	Gly	Gly	Ala	Ala	Ile	Asp	Val	Phe	Pro	Val	Glu	Pro	Ala	Thr
			260					265						270	
Asn	Ser	Asp	Pro	Phe	Val	Ser	Pro	Leu	Ser	Glu	Phe	Asp	Asn	Val	Ile
			275					280					285		
Leu	Thr	Pro	His	Ile	Gly	Gly	Ser	Thr	Glu	Glu	Ala	Gln	Glu	Asn	Ile
			290					295					300		
Gly	Ile	Glu	Val	Ala	Gly	Lys	Leu	Ala	Lys	Tyr	Ser	Asp	Asn	Gly	Ser
305					310						315				320
Thr	Leu	Ser	Ala	Val	Asn	Phe	Pro	Glu	Val	Ser	Leu	Pro	Met	His	Gly
				325						330					335
Ile	Ser	Ala	Ser	Arg	Leu	Leu	His	Ile	His	Glu	Asn	Arg	Pro	Gly	Val
			340					345						350	
Leu	Thr	Ala	Ile	Asn	Gln	Ile	Phe	Ala	Glu	Gln	Gly	Ile	Asn	Ile	Ala
			355					360					365		
Ala	Gln	Tyr	Leu	Gln	Thr	Ser	Pro	Met	Met	Gly	Tyr	Val	Val	Ile	Asp
			370					375					380		
Ile	Asp	Ala	Glu	His	Glu	Leu	Ala	Glu	Lys	Ala	Leu	Gln	Leu	Met	Lys
385					390						395				400
Ala	Ile	Pro	Gly	Thr	Ile	Arg	Ala	Arg	Leu	Leu	Tyr				
				405							410				

<210> 51

<211> 36

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P11

<400> 51

agctgagtcg acccccagga aaaattgggtt aataac

36

<210> 52

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P12

<400> 52

agctgagcat gcttccaact gcgctaata ga cgc

33

<210> 53

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P13

<400> 53

agctgatcta gaaaacagaa tttagcctggc ggc

33

<210> 54

<211> 33

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P14

<400> 54

agctgaggat ccaggaagag tttagtagaaa cgc

33

<210> 55

<211> 32

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P15

<400> 55
agctgagtcg acgtgttcgc tgaatacggg gt 32

<210> 56
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mỗi P16

<400> 56
agctgatcta gagaaagcat caggattgca gc 32

<210> 57
<211> 59
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mỗi P17

<220>
<221> dấu hiệu_misc
<222> (25)..(25)
<223> n là a, c, g, hoặc t

<220>
<221> dấu hiệu_misc
<222> (27)..(28)
<223> n là a, c, g, hoặc t

<220>
<221> dấu hiệu_misc
<222> (46)..(46)
<223> n là a, c, g, hoặc t

<220>
<221> dấu hiệu_misc
<222> (49)..(49)
<223> n là a, c, g, hoặc t

<220>
<221> dấu hiệu_misc
<222> (51)..(51)
<223> n là a, c, g, hoặc t

<400> 57
atcgtgaaga tcttttccag tgttnannag ggtgccttgc acggttatna ngtcactgg 59

<210> 58
<211> 32
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mỗi P18

<220>
<221> dấu hiệu_misc
<222> (16)..(20)
<223> n là a, c, g, hoặc t

<400> 58
tggaagaagat cttcannnnn cgtcgacctg cg 32

<210> 59
<211> 60
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mỗi P19

<400> 59
tcgctcagc atttttttca tcgctggtaa ggtcatttat cccccaggaa aaattgggta 60

<210> 60
<211> 63
<212> ADN
<213> Trình tự nhân tạo

<220>
<223> Đoạn mỗi P20

<400> 60
tttcacaccg ctcaaccgca gggcataacc ggcccttgaa gcctgctttt ttataactaag 60
ttg 63

<210> 61
<211> 20

<212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi P21

<400> 61
 ctttgtccct ttagtgaagg

20

<210> 62
 <211> 44
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi P22

<400> 62
 agctgatcta gaagctgact cgagttaatg gctcccaga cgac

44

<210> 63
 <211> 33
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi P23

<400> 63
 agctgagtcg acatggcaaa ggtatcactg gaa

33

<210> 64
 <211> 34
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> Đoạn mỗi P24

<400> 64
 gagaacgccc gggcgggctt cgtgaatatg cagc

34

<210> 65
 <211> 32
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P25

<400> 65

agctgatcta gacgtgggat cagtaaagca gg

32

<210> 66

<211> 22

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Đoạn mỗi P26

<400> 66

aaaaccgccc gggcggtctc ac

22

Fig. 1

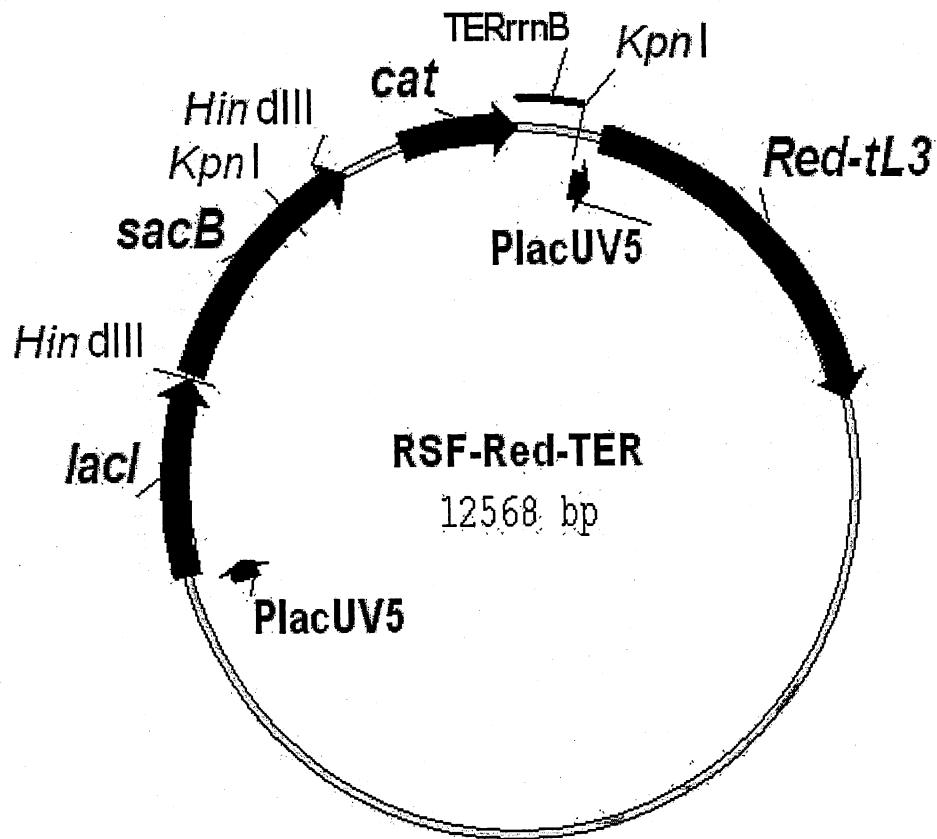


Fig. 2

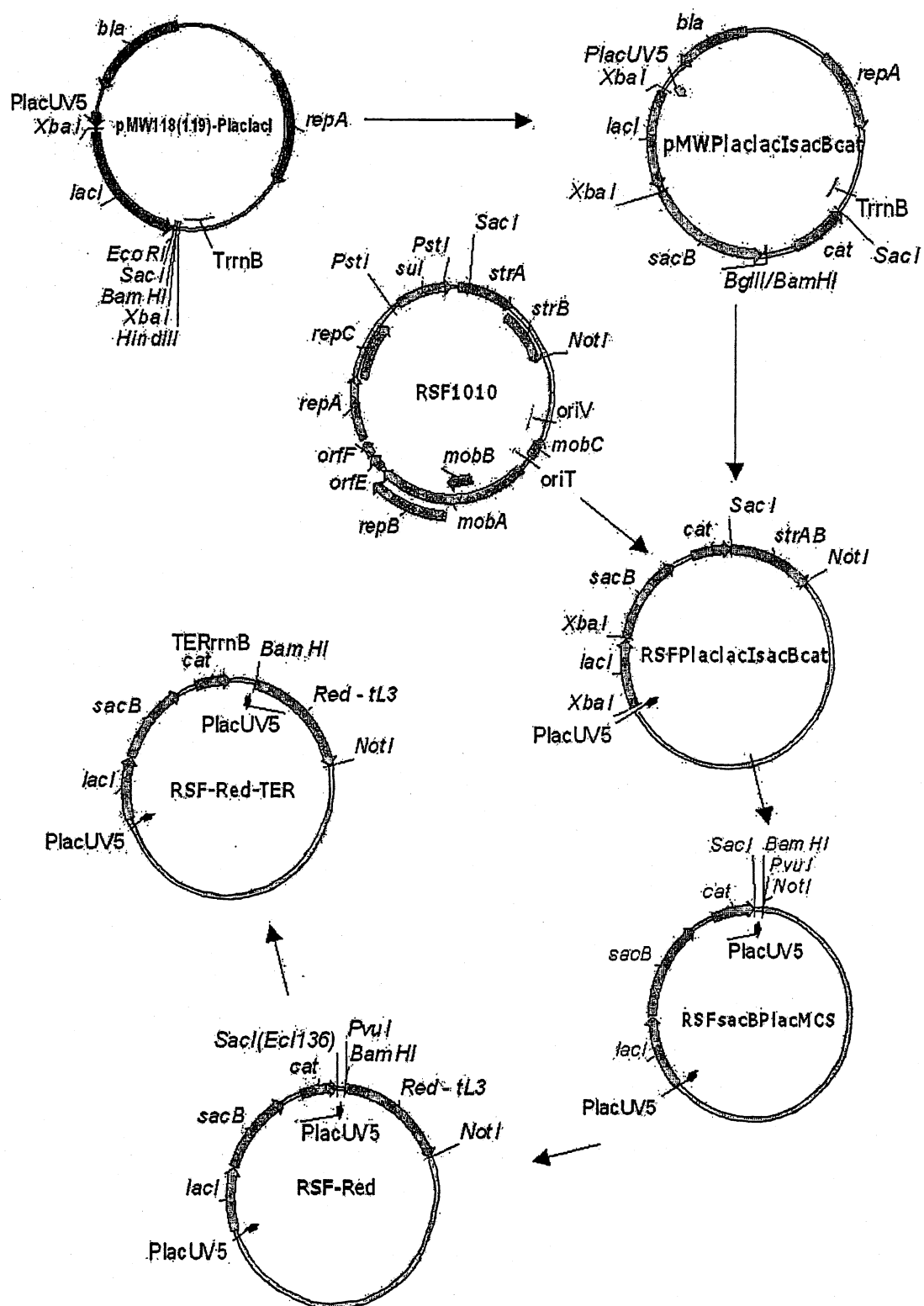


Fig. 3

Trình tự của đoạn khởi đầu nguyên gốc Pnlp

```

                                -35(Pnlp2)                                -10(Pnlp2)
aaaacgtgaggaaatacctggatttttcctgggtattttgccgcaggtcagcgtatcgtg
                                -35(Pnlp1)                                -10(Pnlp1) bắt đầu phiên mã
aagatcttttccagtgttcagtagggtgccttgcacggtaattatgtcactgggtattaa
                                M   S
ccaatttttcctgggggataaatgagc

```