



(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)



1-0039176

(51)¹⁹ C07K 14/34; C12P 19/40; C12N 15/77 (13) B

(21) 1-2019-06465

(22) 14/12/2018

(86) PCT/KR2018/015937 14/12/2018

(87) WO2019/117673 20/06/2019

(30) 10-2017-0173505 15/12/2017 KR

(45) 25/03/2024 432

(43) 25/06/2020 387

(73) CJ CHEILJEDANG CORPORATION (KR)

330, Dongho-ro, Jung-gu, Seoul 04560, Republic of Korea

(72) KWON, Jung Gun (KR); BAEK, Min Ji (KR); LEE, Ji Hye (KR); KWON, Nara (KR); KIM, Ju Jeong (KR); RHO, Jin Ah (KR); CHO, Jin Man (KR).

(74) Công ty TNHH Sáng chế ACTIP (ACTIP PATENT LIMITED)

(54) BIẾN THỂ PROTEIN CÓ HOẠT TÍNH SẢN SINH 5'-INOSIN MONOPHOSPHAT, POLYNUCLEOTIT, VECTƠ CHỨA POLYNUCLEOTIT, VI SINH VẬT CHỨA BIẾN THỂ PROTEIN, VÀ PHƯƠNG PHÁP ĐIỀU CHẾ 5'-INOSIN MONOPHOSPHAT

(57) Sáng chế đề cập đến biến thể protein, vectơ chứa polynucleotit, polypeptit mới có hoạt tính sản sinh 5'-inosin monophosphat, vi sinh vật bao gồm polypeptit này, phương pháp điều chế 5'-inosin monophosphat sử dụng polypeptit, và phương pháp tăng sản sinh 5'-inosin monophosphat.

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến biến thể protein mới có hoạt tính sản sinh ra 5'-inosin monophosphat (IMP), polynucleotit, vectơ chứa polynucleotit, vi sinh vật chứa protein này, phương pháp để điều chế IMP và phương pháp để làm tăng sự sản sinh IMP.

Tình trạng kỹ thuật của sáng chế

5' Inosin monophosphat (sau đây gọi là IMP), loại vật liệu có bản chất là axit nucleic là một dạng trung gian của quá trình chuyển hóa axit nucleic được sử dụng trong nhiều lĩnh vực như trong thực phẩm, y tế, các ứng dụng y tế khác nhau v.v.. Đặc biệt là, IMP được sử dụng rộng rãi như chất phụ gia cho gia vị thực phẩm hoặc thực phẩm cùng với monophosphat 5-guanin (sau đây gọi là GMP). Mặc dù bản thân IMP được biết đến để cung cấp gia vị có hương vị thịt bò, nó cũng được biết đến là chất tăng cường hương vị của axit glutamic dạng bột (bột ngọt) và vì thế được chú ý đến như loại gia vị axit nucleic làm tăng hương vị.

Ví dụ về các phương pháp sản sinh IMP bao gồm phương pháp phân hủy enzym axit ribonucleic chiết xuất từ các tế bào nấm men (Công bố patent Nhật Bản số 1614/1957), phương pháp để phosphoryl hóa inosin được sản xuất bằng cách lên men (Agri. Biol. Chem., 36, 1511, v.v.), phương pháp nuôi cấy vi sinh vật có thể trực tiếp sản sinh IMP và thu hồi IMP trong môi trường nuôi cấy, v.v.. Trong các phương pháp này, phương pháp được sử dụng thường xuyên nhất hiện nay là phương pháp sử dụng vi sinh vật có khả năng sản xuất trực tiếp IMP.

Trong khi đó, do các enzym không phải lúc nào cũng thể hiện các tính chất tối ưu trong tự nhiên liên quan đến hoạt tính, tính ổn định, tính đặc hiệu của chất đồng phân quang học, v.v. theo yêu cầu của các ứng dụng công nghiệp, nhiều nỗ lực đã được thực hiện để cải thiện enzym cho phù hợp với mục đích sử dụng bằng cách thay đổi trình tự axit amin của chúng v.v.. Trong số này, mặc dù các thiết kế hợp lý và gây đột biến điểm tại chỗ của các enzym đã được áp dụng để cải thiện chức năng enzym, nhưng trong

nhiều trường hợp, những nỗ lực này đã được chứng minh là bất lợi khi thông tin về cấu trúc của enzym đích là không phù hợp hoặc mối tương quan cấu trúc-chức năng không rõ ràng, do đó ngăn chặn hiệu quả khi ứng dụng chúng. Ngoài ra, phương pháp cải thiện hoạt động của enzym bằng cách cố gắng tăng cường enzym thông qua quá trình tác động trực tiếp, nhằm sàng lọc các enzym đáp ứng các đặc điểm mong muốn từ một loạt các enzym biến đổi tạo được thông qua đột biến gen ngẫu nhiên, đã được báo cáo trước đây.

Để sản sinh IMP với năng suất cao bằng cách sử dụng phương pháp sản xuất trực tiếp IMP thông qua quá trình lên men vi sinh vật, IMP đã được sản sinh một cách thuận lợi. Để thực hiện mục đích này, các tác giả của sáng chế đã phát hiện ra protein có liên quan đến hoạt tính sản sinh IMP và cũng đã thực hiện nhiều nỗ lực để tăng sản lượng IMP. Kết quả là, họ đã phát hiện ra các biến thể protein có hoạt tính sản sinh IMP, do đó hoàn thành sáng chế này.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất biến thể protein sản sinh IMP.

Mục đích khác của sáng chế là đề xuất polynucleotit mã hóa biến thể protein của sáng chế.

Mục đích khác của sáng chế là đề xuất vector chứa polynucleotit của sáng chế.

Mục đích khác của sáng chế là đề xuất vi sinh vật sản sinh IMP, chứa biến thể protein và vector của sáng chế.

Mục đích khác nữa của sáng chế là đề xuất phương pháp để điều chế IMP bao gồm nuôi cấy vi sinh vật của sáng chế trong môi trường.

Mục đích khác nữa của sáng chế là đề xuất phương pháp để tăng cường sản sinh IMP bao gồm tăng cường hoạt tính của biến thể protein của sáng chế, có hoạt tính sản sinh IMP.

Hiệu quả đạt được của sáng chế

IMP có thể được sản xuất với năng suất cao bằng cách nuôi cấy vi sinh vật có nguồn gốc thuộc chi *Corynebacterium* sản sinh IMP bằng cách sử dụng biến thể protein của sáng chế có khả năng sản sinh IMP.

Mô tả chi tiết sáng chế

Sáng chế sẽ được mô tả chi tiết dưới đây. Trong đó, mỗi phần mô tả và các phương án ví dụ thực hiện được bộc lộ ở đây có thể được áp dụng cho các phần mô tả tương ứng và các phương án ví dụ khác. Tức là, tất cả sự kết hợp của các yếu tố khác nhau được thể hiện ở đây thuộc về phạm vi của sáng chế. Ngoài ra, phạm vi của sáng chế sẽ không bị giới hạn bởi sáng chế được mô tả dưới đây.

Để đạt được các mục đích trên, khía cạnh của sáng chế cung cấp biến thể protein có hoạt tính sản sinh IMP.

Như được sử dụng ở đây, thuật ngữ “protein sản sinh 5'-inosin monophosphat (IMP)” dùng để chỉ protein liên quan đến sản sinh ngoại bào của IMP. Với mục đích của sáng chế, thuật ngữ này có thể được sử dụng thay thế cho protein có hoạt tính sản sinh IMP, protein sản sinh IMP, protein có hoạt động sản sinh 5'-inosin monophosphat, tạo 5'-inosin monophosphat protein, v.v.; cụ thể là, protein có thể được biểu thị dưới dạng ImpE, và cụ thể hơn là, có thể được biểu thị dưới dạng ImpE1 hoặc ImpE2, nhưng sáng chế không giới hạn ở đây. Ngoài ra, protein có thể được lấy từ vi sinh vật thuộc chi *Corynebacterium*, và cụ thể là từ *Corynebacterium stationis*, nhưng vi sinh vật không bị giới hạn ở chi này.

Ví dụ, protein có thể bao gồm trình tự axit amin được biểu thị bằng SEQ ID NO: 1 hoặc SEQ ID NO: 2, nhưng bất kỳ trình tự nào có cùng hoạt tính như protein đều có thể được đưa vào mà không bị giới hạn, và người có hiểu biết trung bình trong cùng lĩnh vực tương ứng có thể có được thông tin trình tự từ Ngân hàng Gen của Trung tâm Thông tin Công nghệ sinh học Quốc gia Hoa Kỳ (The National Center for Biotechnology Information: NCBI), cơ sở dữ liệu nổi tiếng. Ngoài ra, protein có thể bao gồm trình tự axit amin của SEQ ID NO: 1 hoặc SEQ ID NO: 2 hoặc trình tự axit amin có sự đồng nhất hoặc tương đồng với trình tự SEQ ID NO: 1 hoặc SEQ ID NO: 2 ít nhất 80%, ít nhất 90%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98%, hoặc ít nhất 99%. Hơn nữa, rõ ràng là bất kỳ protein nào có trình tự axit amin bị mất, sửa đổi, thay thế hoặc bổ sung vào một phần của trình tự cũng có thể thuộc phạm vi của sáng chế, miễn là trình tự axit amin có đồng nhất hoặc tương đồng như được mô tả ở trên và có tác dụng tương ứng với protein.

Điều đó có nghĩa là, mặc dù được mô tả là “biến thể protein có trình tự axit amin là SEQ ID NO cụ thể” hoặc “protein có trình tự axit amin là SEQ ID NO cụ thể” trong

sáng chế, protein có thể có hoạt tính giống hoặc tương tự với protein có trình tự axit amin là SEQ ID NO tương ứng. Trong trường hợp như vậy, rõ ràng là bất kỳ protein nào có trình tự axit amin bị mất, sửa đổi, thay thế, thay thế bảo thủ hoặc bổ sung vào một phần của trình tự cũng có thể được sử dụng trong sáng chế này. Ví dụ, trong trường hợp có hoạt tính giống hoặc tương tự với hoạt tính của protein bị sửa đổi, không loại trừ trình tự bổ sung ngược hoặc xuôi của trình tự axit amin, mà không làm thay đổi chức năng của protein, đột biến có thể xảy ra một cách tự nhiên, đột biến im lặng của nó hoặc thay thế bảo thủ và ngay cả khi bổ sung trình tự hoặc đột biến, nó rõ ràng thuộc vào phạm vi của sáng chế.

Như được sử dụng trong sáng chế, thuật ngữ “tương đồng” và “đồng nhất” liên quan đến mức độ phù hợp với hai trình tự axit amin hoặc trình tự nucleotit nhất định và có thể được biểu thị bằng phần trăm.

Các thuật ngữ “tương đồng” và “đồng nhất” thường có thể được sử dụng thay thế cho nhau.

Trình tự tương đồng và đồng nhất của các trình tự polynucleotit hoặc polypeptit bảo tồn có thể được xác định bằng các thuật toán căn chỉnh tiêu chuẩn và có thể được sử dụng với hàm phạt khoảng trống dịch chuyển mặc định được thiết lập bởi chương trình đang được sử dụng. Các trình tự tương đồng hoặc đồng nhất thường được dự kiến sẽ lai ghép ở mức độ trung bình hoặc cao, dọc theo toàn bộ chiều dài của trình tự hoặc ít nhất khoảng 50%, khoảng 60%, khoảng 70%, khoảng 80%, hoặc khoảng 90% toàn bộ chiều dài của trình tự. Polynucleotit có chứa các mã thoái biến thay vì các mã trong các polypeptit lai hóa cũng được xem xét.

Cho dù hai trình tự polynucleotit hoặc polypeptit có sự tương đồng, tương tự, hay đồng nhất có thể được xác định bằng thuật toán máy tính đã biết như chương trình “FASTA” (Pearson et al., (1988) [Proc. Natl. Acad. Sci. USA 85]: 2444: sử dụng tham số mặc định trong 2444). Ngoài ra, nó có thể được xác định bằng thuật toán Needman-Wunsch (Needman và Wunsch, 1970, J. Mol. Biol. 48: 443-453), được thực hiện trong chương trình Needman của gói Bộ phần mềm mở sinh học phân tử châu Âu ((EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Genet Trends. 16: 276-277) (phiên bản 5.0.0 hoặc cao hơn) (gói chương trình GCG (Devereux, J., et al., Nghiên cứu axit nucleic 12: 387 (1984)), BLASTP, BLASTN, FASTA (Atschul,

[S.] [F.,] [ET AL, J MOLEC BIOL 215]: 403 (1990); Hướng dẫn về máy tính khổng lồ, Martin J. Giám mục, [ED.,] Nhà xuất bản học thuật, San Diego, 1994 và [CARILLO ETA/.] (1988) SIAM J Toán ứng dụng 48 : 1073). Ví dụ, tính tương đồng, tương tự hoặc đồng nhất có thể được xác định bằng BLAST hoặc ClustalW của Trung tâm Thông tin Công nghệ sinh học Quốc gia (National Center for Biotechnology Information: NCBI).

Tính tương đồng, tương tự hoặc đồng nhất của các trình tự polynucleotit hoặc polypeptit có thể được xác định bằng cách so sánh thông tin trình tự sử dụng, ví dụ, chương trình máy tính GAP (ví dụ, Needleman et al., (1970), J Mol Biol.48: 443) như đã được công bố (ví dụ, Smith và Waterman, Adv. Appl. Math (1981) 2: 482). Tóm lại, chương trình GAP định nghĩa tương đồng, tương tự hoặc đồng nhất là giá trị thu được bằng cách chia số lượng các dấu hiệu được căn chỉnh tương tự (nghĩa là, nucleotit hoặc axit amin) vào tổng số dấu hiệu trong trình tự ngắn hơn của hai trình tự. Các tham số mặc định cho chương trình GAP có thể bao gồm (1) ma trận so sánh đơn nhất (chứa giá trị 1 cho giống nhau và 0 cho không giống nhau) và ma trận so sánh có trọng số của Gribskov et al. (1986), Hạt nhân. Axit Res. 14: 6745, như được thể hiện trong Schwartz và Dayhoff, biên tập, Atlas về trình tự và cấu trúc protein, Quỹ nghiên cứu y sinh quốc gia, trang 353-358, 1979; (2) điểm phạt là 3,0 cho mỗi khoảng trống dịch chuyển và điểm phạt bổ sung là 0,10 cho mỗi dấu hiệu trong mỗi khoảng trống dịch chuyển (hoặc điểm phạt mở khoảng trống dịch chuyển là 10 và điểm phạt mở rộng khoảng trống dịch chuyển là 0,5); và (3) không bị phạt đối với các khoảng trống dịch chuyển cuối. Theo đó, như được sử dụng trong tài liệu này, thuật ngữ “tương đồng” hay “đồng nhất” là đề cập đến sự liên quan giữa các trình tự. Cụ thể là, biến thể protein theo sáng chế này có hoạt tính sản sinh IMP có thể là một protein mà trong đó có ít nhất axit amin được chọn từ nhóm bao gồm axit amin thứ 164 trong trình tự axit amin của SEQ ID NO: 1, axit amino thứ 222 trong trình tự axit amin của SEQ ID NO: 1, axit amin thứ 2 trong trình tự axit amin của SEQ ID NO: 2, và axit amin thứ 64 trong trình tự axit amin của SEQ ID NO: 2 được thay thế bằng axit amin khác, nhưng sáng chế không giới hạn ở đây.

Ví dụ, trong biến thể protein có hoạt tính sản sinh IMP, axit amin thứ 164 trong trình tự axit amin của SEQ ID NO: 1 được thay thế bằng lysin, arginin, asparagin, glyxin, threonin hoặc prolin; axit amin thứ 2 trong trình tự axit amin của SEQ ID NO: 2 được thay thế bằng isoleuxin, phenylalanin, methionin, axit glutamic, histidin hoặc asparagin; hoặc axit amin thứ 64 trong trình tự axit amin của SEQ ID NO: 2 được thay thế bằng

axit aspartic, axit glutamic, asparagin, xystein, isoleuxin hoặc phenylalanin, nhưng sáng chế không bị giới hạn ở đây.

Một ví dụ thực hiện, biến thể protein có hoạt tính sản sinh IMP có thể là protein có trình tự axit amin bao gồm SEQ ID NO: 141, SEQ ID NO: 142, SEQ ID NO: 145, SEQ ID NO: 147, SEQ ID NO: 149 hoặc SEQ ID NO: 151, protein có trình tự axit amin được mã hóa bởi polynucleotit của SEQ ID NO: 153 hoặc SEQ ID NO: 154, hoặc protein có trình tự axit amin có sự tương đồng ít nhất 80%, ít nhất 90%, ít nhất 95%, ít nhất 96%, ít nhất 97%, ít nhất 98%, hoặc ít nhất 99%. Ngoài ra, rõ ràng là protein bị mất, sửa đổi, thay thế hoặc bổ sung một số trình tự có thể được sử dụng làm protein của sáng chế này miễn nó là protein có trình tự axit amin tương đồng ở trên và thể hiện hiệu quả tương tự với hiệu quả của protein.

Khía cạnh khác của sáng chế đề xuất polynucleotit mã hóa biến thể protein, hoặc vector bao gồm cả polynucleotit.

Như được sử dụng trong sáng chế, thuật ngữ “polynucleotit” có nghĩa là polyme của nucleotit được kéo dài thành chuỗi dài bởi các liên kết cộng hóa trị và có một chuỗi ADN hoặc một chuỗi ARN dài hơn một độ dài nhất định và cụ thể hơn là nói đến đoạn polynucleotit mã hóa biến thể protein.

Rõ ràng là polynucleotit có thể được dịch mã bởi đơn vị mã hóa có tính thoái hóa thành protein bao gồm trình tự axit amin của SEQ ID NO: 141, SEQ ID NO: 142, SEQ ID NO: 145, SEQ ID NO: 147, SEQ ID NO: 149 hoặc SEQ ID NO: 151, biến thể protein bao gồm trình tự axit amin được mã hóa bởi polynucleotit của SEQ ID NO: 153 hoặc SEQ ID NO: 154, hoặc thành protein có sự đồng nhất cũng có thể được tính như là polynucleotit của sáng chế. Ví dụ, polynucleotit của sáng chế có thể là polynucleotit có trình tự nucleotit của SEQ ID NO: 143, SEQ ID NO: 144, SEQ ID NO: 146, SEQ ID NO: 148, SEQ ID NO: 150, SEQ ID NO: 152, SEQ ID NO: 153 hoặc SEQ ID NO: 154, và cụ thể hơn, có thể là polynucleotit bao gồm trình tự nucleotit của SEQ ID NO: 143, SEQ ID NO: 144, SEQ ID NO: 146, SEQ ID NO: 148, SEQ ID NO: 150, SEQ ID NO: 152, SEQ ID NO: 153 hoặc SEQ ID NO: 154. Ngoài ra, trình tự polynucleotit mã hóa protein có hoạt tính của protein có trình tự axit amin của SEQ ID NO: 141, SEQ ID NO: 142, SEQ ID NO: 145, SEQ ID NO: 147, SEQ ID NO: 149 hoặc SEQ ID NO: 151 hoặc trình tự axit amin được mã hóa bởi polynucleotit của SEQ ID NO: 153 hoặc SEQ ID

NO: 154 bằng cách lai dưới các điều kiện nghiêm ngặt với đầu dò có thể được điều chế từ các trình tự gen đã biết, ví dụ, trình tự bổ sung cho tất cả hoặc một phần trình tự nucleotit, có thể được đưa vào và không giới hạn.

Thuật ngữ “điều kiện nghiêm ngặt” đề cập đến các điều kiện theo đó các phép lai cụ thể giữa các polynucleotit được thực hiện. Các điều kiện như vậy được mô tả cụ thể trong tài liệu tham khảo (ví dụ, J. Sambrook et al, *supra*). Ví dụ, các điều kiện có thể bao gồm thực hiện lai giữa các gen có tính tương đồng cao, tương đồng từ 40% trở lên, cụ thể là 90% hoặc cao hơn, cụ thể hơn là 95% hoặc cao hơn, cụ thể hơn là 97% hoặc cao hơn, và cụ thể nhất là 99% hoặc cao hơn, trong khi không thực hiện lai giữa các gen có tỷ lệ tương đồng thấp hơn các tỷ lệ tương đồng ở trên; hoặc thực hiện lai tạo một lần, cụ thể hai hoặc ba lần, trong các điều kiện rửa thông thường để lai Southern 60°C, $1 \times \text{SSC}$ (saline-sodium citrate: Salin natri xintrat) và 0,1% SDS (sodium dodecyl sulfate: natri dodecyl sulfat), cụ thể ở nồng độ muối và nhiệt độ tương ứng với 60°C, $0,1 \times \text{SSC}$, và 0,1% SDS, và cụ thể hơn là 68°C, $0,1 \times \text{SSC}$ và 0,1% SDS.

Phép lai đòi hỏi hai axit nucleic có trình tự bổ sung, mặc dù sự không phù hợp giữa các bazơ có thể phụ thuộc vào mức độ nghiêm ngặt của phép lai. Thuật ngữ “bổ sung” được sử dụng để thể hiện mối quan hệ giữa các bazơ nucleotit lai lẫn nhau. Ví dụ, liên quan đến ADN, adenosin bổ sung cho thymin và xytoxin được bổ sung cho guanin. Theo đó, sáng chế cũng có thể bao gồm các đoạn axit nucleic đã phân lập bổ sung cho toàn bộ trình tự cũng như các trình tự axit nucleic tương tự về cơ bản.

Cụ thể, các polynucleotit có tương đồng có thể được phát hiện ở giá trị T_m là 55°C bằng cách sử dụng các điều kiện lai bao gồm bước lai hóa và sử dụng các điều kiện được mô tả ở trên. Ngoài ra, giá trị T_m có thể là 60°C, 63°C hoặc 65°C, nhưng sáng chế không bị giới hạn ở các nhiệt độ này và có thể được điều chỉnh phù hợp bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng theo mục đích đã định.

Tính nghiêm ngặt phù hợp cho việc lai các polynucleotit phụ thuộc vào độ dài và tính bổ sung của các polynucleotit và các biến số liên quan được biết đến trong kỹ thuật lai (tham khảo Sambrook et al, *supra*, 9,50 đến 9,51 và 11,7 đến 11,8).

Như được sử dụng trong sáng chế, thuật ngữ “vector” đề cập đến cấu trúc ADN bao gồm trình tự nucleotit của polynucleotit mã hóa protein đích được liên kết hoạt động với trình tự kiểm soát phù hợp để protein đích có thể được biểu thị trong vật chủ thích hợp.

Trình tự kiểm soát có thể bao gồm vùng khởi động có khả năng khởi đầu phiên mã, bất kỳ trình tự chỉ huy để kiểm soát phiên mã, trình tự mã hóa cho vùng liên kết mARN-ribosom thích hợp, trình tự kiểm soát đầu cuối của phiên mã và dịch mã. Vector, sau khi được biến nạp vào vật chủ thích hợp, có thể được sao chép hoặc đồng nhất chức năng với bộ gen của vật chủ hoặc có thể được tích hợp vào chính bộ gen của vật chủ.

Vector được sử dụng theo sáng chế có thể không bị giới hạn đặc biệt nào miễn là vector có thể nhân rộng trong tế bào chủ và nó có thể được xây dựng bằng bất kỳ vector nào được biết trong kỹ thuật tương ứng. Ví dụ về vector có thể bao gồm các plasmid tự nhiên hoặc tái tổ hợp, các cosmid, virus và vi khuẩn. Ví dụ, như vector thể thực khuẩn hoặc vector cosmid, có thể được sử dụng pWE15, M13, MBL3, MBL4, IXII, ASHII, APII, t10, t11, Charon4A, Charon21A, v.v. và như vector plasmid, những loại dựa trên pBR, pUC, pBluescriptII, pGEM, pTZ, pCL, pET, v.v. có thể được sử dụng. Cụ thể, có thể sử dụng các vector pDZ, pACYC177, pACYC184, pCL, pECCG117, pUC19, pBR322, pMW118, pCC1BAC, v.v..

Theo phương án của sáng chế, polynucleotit mã hóa protein đích có thể được thay thế bằng polynucleotit biến đổi trong nhiễm sắc thể bằng cách sử dụng vector để đưa vào nhiễm sắc thể trong tế bào. Việc chèn polynucleotit vào nhiễm sắc thể có thể được thực hiện bằng phương pháp đã biết trong kỹ thuật tương ứng, ví dụ, bằng cách tái tổ hợp tương đồng, nhưng không giới hạn ở đó. Cụ thể, dấu hiệu lựa chọn để xác nhận sự chèn vào nhiễm sắc thể có thể được thêm vào. Dấu hiệu lựa chọn được sử dụng để chọn một tế bào đã biến nạp, nghĩa là, để xác nhận xem axit nucleic đích đã được chèn vào chưa, và các dấu hiệu có khả năng cung cấp các kiểu hình được lựa chọn như kháng thuốc, yêu cầu dinh dưỡng, kháng các tác nhân gây độc tế bào và biểu hiện protein bề mặt có thể được sử dụng. Trong trường hợp xử lý các tác nhân chọn lọc, chỉ có các tế bào có khả năng biểu thị các dấu hiệu chọn lọc có thể tồn tại hoặc biểu hiện các đặc điểm kiểu hình khác, và do đó có thể dễ dàng chọn lựa các tế bào biến nạp.

Khía cạnh khác nữa của sáng chế đề xuất vi sinh vật sản sinh IMP bao gồm biến thể protein theo sáng chế, polynucleotit theo sáng chế mã hóa biến thể protein hoặc vector theo sáng chế. Cụ thể, vi sinh vật bao gồm biến thể protein và/hoặc polynucleotit mã hóa biến thể protein có thể là vi sinh vật được điều chế bởi sự biến nạp bằng cách sử dụng vector chứa polynucleotit mã hóa biến thể protein, nhưng sáng chế không bị giới hạn ở vi sinh vật này.

Như được sử dụng trong sáng chế, thuật ngữ “biến nạp” đề cập đến quá trình đưa vector bao gồm polynucleotit mã hóa protein đích vào trong tế bào chủ, do đó cho phép biểu hiện protein được mã hóa bởi polynucleotit trong tế bào chủ. Đối với polynucleotit biến đổi, không quan trọng việc nó được đưa vào nhiễm sắc thể của tế bào chủ và nằm ở đó hay nằm bên ngoài nhiễm sắc thể, miễn là polynucleotit biến nạp có thể được biểu hiện trong tế bào chủ. Ngoài ra, polynucleotit bao gồm ADN và ARN mã hóa protein đích. Polynucleotit có thể được chèn vào dưới bất kỳ hình thức nào miễn là nó có thể được đưa vào tế bào chủ và được biểu hiện trong đó. Ví dụ, polynucleotit có thể được đưa vào tế bào chủ dưới dạng cát-xét biểu hiện, đây là cấu trúc gen bao gồm tất cả các yếu tố cần thiết để tự biểu hiện. Cát-xét biểu hiện thông thường có thể bao gồm vùng khởi động được liên kết hoạt động với polynucleotit, tín hiệu kết thúc phiên mã, miền liên kết ribosom và tín hiệu kết thúc dịch mã. Cát-xét biểu hiện có thể ở dạng vector biểu hiện tự sao chép. Ngoài ra, polynucleotit có thể được đưa vào tế bào chủ và được liên kết hoạt động với trình tự cần thiết cho biểu hiện của nó trong tế bào chủ, nhưng sáng chế không bị giới hạn ở đây.

Ngoài ra, như được sử dụng trong tài liệu này, thuật ngữ “liên kết hoạt động”, có nghĩa là liên kết chức năng giữa trình tự khởi động để khởi tạo và làm trung gian cho quá trình phiên mã của polynucleotit mã hóa protein đích, nghĩa là kết hợp của sáng chế và trình tự gen trên.

Như được sử dụng trong tài liệu này, thuật ngữ “vi sinh vật sản sinh IMP” đề cập đến vi sinh vật có khả năng tự nhiên sản sinh IMP; hoặc vi sinh vật đã được đưa vào có khả năng sản xuất hoặc sản sinh IMP mà chủng gốc không có khả năng sản xuất và/hoặc sản sinh IMP một cách tự nhiên. Trong sáng chế, vi sinh vật sản sinh IMP có thể được sử dụng thay thế cho vi sinh vật có hoạt tính sản sinh IMP.

Vi sinh vật sản sinh IMP là tế bào hoặc vi sinh vật bao gồm biến thể protein có hoạt tính sản sinh IMP hoặc polynucleotit mã hóa biến thể protein, hoặc được biến nạp bằng vector chứa polynucleotit mã hóa biến thể protein, và do đó có khả năng biểu hiện biến thể protein. Theo mục đích của sáng chế, tế bào chủ của vi sinh vật hoặc vi sinh vật sản sinh IMP có thể là bất kỳ vi sinh vật nào bao gồm cả biến thể protein do đó có khả năng sản sinh IMP. Ví dụ, vi sinh vật có thể là sinh vật thuộc chi *Escherichia*, vi sinh vật thuộc chi *Serratia*, vi sinh vật thuộc chi *Erwinia*, vi sinh vật thuộc chi *Enterobacteria*, vi sinh vật thuộc chi *Salmonella*, vi sinh vật thuộc chi *Pseudomonas*, vi

sinh vật thuộc chi *Brevibacterium*, vi sinh vật thuộc chi *Corynebacterium*, v.v., và cụ thể là, vi sinh vật thuộc chi *Corynebacterium*.

Như được sử dụng ở đây, thuật ngữ “vi sinh vật sản sinh IMP thuộc chi *Corynebacterium*” đề cập đến vi sinh vật thuộc chi *Corynebacterium* có khả năng sản sinh IMP tự nhiên hoặc có khả năng sản sinh IMP bằng cách biến đổi. Cụ thể là, như được sử dụng ở đây, vi sinh vật thuộc chi *Corynebacterium* có khả năng sản sinh IMP đề cập đến chủng vi sinh vật bản địa của chi *Corynebacterium* có khả năng sản sinh IMP; hoặc vi sinh vật thuộc chi *Corynebacterium* có khả năng tăng cường sản sinh IMP được điều chế bằng cách chèn gen liên quan đến sản sinh IMP hoặc bằng cách tăng cường hoặc làm suy giảm gen nội sinh liên quan đến sản sinh IMP.

Cụ thể hơn là, theo sáng chế, vi sinh vật thuộc chi *Corynebacterium* có khả năng sản sinh IMP đề cập đến vi sinh vật thuộc chi *Corynebacterium* đã cải thiện khả năng sản sinh IMP bao gồm chứa biến thể protein có hoạt tính sản sinh IMP hoặc polynucleotit mã hóa biến thể protein, hoặc bằng cách biến nạp với vector chứa polynucleotit mã hóa biến thể protein. “Vi sinh vật thuộc chi *Corynebacterium* với khả năng tăng cường sản sinh IMP” đề cập đến loại vi sinh vật thuộc chi *Corynebacterium* đã được tăng cường khả năng sản sinh IMP so với chủng vi khuẩn gốc trước khi biến nạp hoặc vi sinh vật chưa biến đổi của chi *Corynebacter*. Các “vi sinh vật chưa biến đổi thuộc chi *Corynebacterium*” đề cập đến vi sinh vật bản địa của chi *Corynebacterium*, vi sinh vật thuộc chi *Corynebacterium* không chứa biến thể protein có khả năng sản sinh IMP, hoặc vi sinh vật thuộc chi *Corynebacterium* không được biến nạp với vector chứa polynucleotit mã hóa biến thể protein có khả năng sản sinh IMP.

Theo phương án của sáng chế, vi sinh vật có thể là vi sinh vật thuộc chi *Corynebacterium*, trong đó hoạt động của adenylosuccin synthetaza và/hoặc IMP dehydrogenaza bị suy giảm thêm.

Theo sáng chế, “vi sinh vật thuộc chi *Corynebacterium*” đề cập cụ thể đến *Corynebacterium glutamicum*, *Corynebacterium ammoniagenes*, *Brevibacterium lactofermentum*, *Brevibacterium flavum*, *Corynebacterium thermoaminogenes*, *Corynebacterium efficiens*, *Corynebacterium stationis*, v.v., nhưng vi sinh vật không nhất thiết bị giới hạn ở đó.

Khía cạnh khác nữa của sáng chế đề xuất phương pháp để điều chế IMP bao gồm

nuôi cấy vi sinh vật của chi *Corynebacterium* trong môi trường nuôi cấy.

Cụ thể, phương pháp theo sáng chế có thể bao gồm cả bước thu hồi IMP từ vi sinh vật hoặc môi trường nuôi cấy.

Trong phương pháp nêu trên, việc nuôi cấy vi sinh vật có thể được thực hiện quy trình theo mẻ, quy trình liên tục, quy trình lên men bổ sung, v.v. được biết đến trong kỹ thuật tương ứng, nhưng quy trình nuôi cấy không bị giới hạn cụ thể. Đặc biệt là, liên quan đến điều kiện nuôi cấy, độ pH của môi trường nuôi cấy có thể được điều chỉnh với độ pH phù hợp (ví dụ, pH 5 đến 9, tốt hơn là pH 6 đến 8, và tốt nhất là với một hợp chất cơ bản thích hợp (ví dụ, natri hydroxit, kali hydroxit, hoặc amoniac) hoặc hợp chất axit (ví dụ, axit photphoric hoặc axit sunfuric) và điều kiện hiếu khí của nuôi cấy có thể được duy trì bằng cách đưa oxy hoặc hỗn hợp khí có chứa oxy vào môi trường nuôi cấy. Nhiệt độ nuôi cấy thường có thể nằm trong khoảng từ 20°C đến 45°C, và tốt hơn là 25°C đến 40°C trong khoảng 10 đến 160 giờ, nhưng điều kiện nuôi cấy không bị giới hạn ở đây. IMP được tạo ra bởi sự nuôi cấy nêu trên có thể được tiết vào môi trường nuôi cấy hoặc có thể được giữ lại trong các tế bào.

Ngoài ra, các ví dụ về nguồn cacbon được sử dụng trong môi trường nuôi cấy có thể bao gồm đường và cacbonhydrat (ví dụ, glucoza, sucroza, lactoza, fructoza, maltoza, mật, tinh bột và xenluloza); dầu và chất béo (ví dụ, dầu đậu nành, dầu hướng dương, dầu đậu phộng và dầu dừa); axit béo (ví dụ, axit palmitic, axit stearic và axit linoleic); rượu (ví dụ, glyxerol và etanol); và axit hữu cơ (ví dụ, axit axetic), nhưng sáng chế không giới hạn ở đây. Những nguồn cacbon này có thể được sử dụng riêng lẻ hoặc kết hợp, nhưng sáng chế không giới hạn ở đó. Ví dụ về các nguồn nitơ được sử dụng trong môi trường nuôi cấy có thể bao gồm các hợp chất hữu cơ chứa nitơ (ví dụ, pepton, chiết xuất men, nước thịt, chiết xuất mạch nha, rượu ngô, bột đậu nành và ure) hoặc các hợp chất vô cơ (ví dụ, amoni sunfat, amoni clorua, amoni photphat, amoni cacbonat và amoni nitrat), v.v.. Những nguồn nitơ này có thể được sử dụng riêng lẻ hoặc kết hợp, nhưng không giới hạn ở đây. Ví dụ về các nguồn phospho được sử dụng trong môi trường nuôi cấy có thể bao gồm kali dihydrogen phosphat, đipot kali hydro phosphat, muối có chứa natri tương ứng, v.v., nhưng sáng chế không bị giới hạn. Ngoài ra, muối kim loại (ví dụ, magie sunfat hoặc sắt sunfat), axit amin, vitamin, v.v., là những vật liệu thúc đẩy tăng trưởng thiết yếu, có thể được chứa trong môi trường nuôi cấy.

Trong sáng chế, phương pháp thu hồi IMP được tạo ra trong bước nuôi cấy có thể được thực hiện bằng cách thu gom IMP từ môi trường nuôi cấy bằng phương pháp thích hợp được biết đến trong kỹ thuật đã biết. Ví dụ, có thể sử dụng các phương pháp như ly tâm, lọc, sắc ký trao đổi anion, kết tinh, sắc ký lỏng hiệu năng cao (High Performance Liquid Chromatography: HPLC), v.v. và có thể thu hồi IMP mong muốn từ môi trường nuôi cấy hoặc vi sinh vật được nuôi cấy bằng phương pháp thích hợp được biết đến trong kỹ thuật này.

Hơn nữa, việc thu hồi có thể bao gồm quá trình lọc và có thể được thực hiện bằng phương pháp thích hợp được biết đến trong kỹ thuật tương ứng. Do đó, IMP được thu hồi có thể ở dạng tinh khiết hoặc môi trường lên men vi sinh vật có chứa IMP.

Khía cạnh khác của sáng chế đề xuất chế phẩm để sản sinh IMP bao gồm biến thể protein theo sáng chế, có hoạt tính sản sinh IMP, hoặc mã hóa polynucleotit của nó.

Chế phẩm theo sáng chế có thể còn bao gồm nhưng không giới hạn ở thành phần có khả năng điều khiển polynucleotit. Trong chế phẩm của sáng chế, polynucleotit có thể ở dạng bao gồm trong vector để biểu hiện gen liên kết hoạt động trong tế bào chủ mà nó được đưa vào.

Ngoài ra, chế phẩm có thể còn bao gồm bất kỳ tá dược nào phù hợp thường được sử dụng trong chế phẩm để sản sinh IMP. Các tá dược như vậy có thể là chất bảo quản, chất làm ẩm, dịch huyền phù, chất đệm, chất ổn định hoặc chất đồng vị, nhưng không bị giới hạn ở các chất này.

Khía cạnh khác nữa của sáng chế đề xuất phương pháp để tăng cường sản sinh IMP bao gồm tăng cường hoạt tính của biến thể protein có hoạt tính sản sinh IMP trong vi sinh vật thuộc chi *Corynebacterium*.

Các thuật ngữ “protein có hoạt tính sản sinh IMP”, “cải tiến”, và “vi sinh vật thuộc chi *Corynebacterium*” như được mô tả ở trên.

Các ví dụ của sáng chế

Sau đây, sáng chế sẽ được mô tả chi tiết thông qua các phương án ví dụ. Tuy nhiên, hiển nhiên là đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng rằng những phương án ví dụ này chỉ được đưa ra nhằm mục đích minh họa và không nhằm giới hạn phạm vi của sáng chế.

Ví dụ 1: Phát hiện về protein sản sinh IMP

Ngân hàng gen ADN của *Corynebacterium Stationis ATCC6872* đã được chuẩn bị để xác định protein màng của *Corynebacterium* liên quan đến việc sản sinh IMP. Sau đó, do chủng *Corynebacterium* hoang dại không thể tạo ra IMP, hoặc ngay cả khi nó tạo ra IMP, nó chỉ tạo ra một lượng nhỏ, trong khi đó, một chủng khác có tên là *CJI0323*, có khả năng sản sinh IMP, có nguồn gốc từ chủng *ATCC6872* đã được chuẩn bị cho việc xác định khả năng sản sinh IMP. Chủng *CJI0323* được chuẩn bị đã được sàng lọc các protein màng liên quan đến việc tổng hợp bằng cách sử dụng ngân hàng gen ADN của chủng *ATCC6872*. Chi tiết của thí nghiệm như sau.

Ví dụ 1-1: Lựa chọn chủng sản sinh IMP, CJI0323

Các tế bào *ATCC6872* được pha loãng trong dung dịch đệm phosphat (pH 7,0) hoặc đệm xitrat (pH 5,5) ở nồng độ 10^7 tế bào/mL đến 10^8 tế bào/mL để điều chế chủng sản sinh IMP có nguồn gốc *ATCC6872* và các tế bào đã được xử lý bằng tia cực tím UV để gây đột biến. Các tế bào thu được được rửa hai lần bằng dung dịch muối 0,85%, sau đó pha loãng và đặt lên đĩa trong môi trường nuôi cấy đã được điều chế bằng cách thêm nguyên liệu liệu cung cấp sức đề kháng ở nồng độ thích hợp vào môi trường tối thiểu chứa 1,7% aga, và các khuẩn lạc được lấy sau đó. Mỗi khuẩn lạc được nuôi cấy trong môi trường dinh dưỡng và nuôi cấy trong môi trường nuôi cấy giống trong 24 giờ. Sau khi nuôi cấy các khuẩn lạc trong 3 đến 4 ngày trong môi trường lên men, khuẩn lạc có khả năng sản sinh IMP cao nhất được tích lũy trong môi trường nuôi cấy đã được chọn. Trong quá trình điều chế chủng có khả năng sản sinh IMP ở nồng độ cao, để cung cấp khuyết dưỡng adenin, rò rỉ guanin, khả năng nhạy lizozim, kháng 3,4-đihydroprolin, kháng streptomycin, kháng axit azetidin cacboxylic, kháng thiaprolin, kháng azaserin, kháng sulfaguanidin, kháng norvalin và kháng trimethoprim, các quy trình trên được thực hiện tuần tự cho từng nguyên liệu. Kết quả là, *CJI0323* cho thấy khả năng kháng các nguyên liệu trên và khả năng tốt để sản sinh IMP. Mức độ kháng giữa *ATCC6872* và *CJI0323* đã được so sánh và kết quả được thể hiện trong Bảng 1 dưới đây.

Bảng 1

Đặc điểm	<i>ATCC6872</i>	<i>CJI0323</i>
Khuyết dưỡng Adenin	Không khuyết dưỡng	khuyết dưỡng
Rò rỉ Guanin	Không khuyết dưỡng	Khuyết dưỡng rò rỉ

Khả năng nhạy Lizozim	80 µg/mL	8 µg/mL
Kháng 3,4-đihydroprolin	1.000 µg/mL	3.500 µg/mL
Kháng Streptomycin	500 µg/mL	2.000 µg/mL
Kháng axit Azetidin cacboxylic	5 mg/mL	30 mg/mL
Kháng Thiaprolin	10 µg/mL	100 µg/mL
Kháng Azaserin	25 µg/mL	100 µg/mL
Kháng Sulfaguanidin	50 µg/mL	200 µg/mL
Kháng Norvalin	0,2 mg/mL	2 mg/mL
Kháng Trimethoprim	20 µg/mL	100 µg/mL

- Môi trường tối thiểu: 2% glucoza, 0,3% natri sulfat, 0,1% KH_2SO_4 , 0,3% K_2HPO_4 , 0,3% sulfat magie, canxi clorua (10 mg/L), sắt sulfat (10 mg/L), kẽm sulfat (1 mg/L), mangan clorua (3,6 mg/L), L-xystein (20 mg/L), canxi pantothenat (10 mg/L), thiamin hydrochlorit (5 mg/L), biotin (30 µg/L), adenin (20 mg/L), guanin (20 mg/L), độ pH 7,3

- Môi trường dinh dưỡng: 1% pepton, 1% nước thịt, 0,25% natri clorua, 1% chiết xuất men, 2% aga, độ pH 7,2

- Môi trường nuôi cấy giống: 1% glucoza, 1% pepton, 1% nước thịt, 1% chiết xuất men, 0,25% natri clorua, adenin (100 mg/L), guanin (100 mg/L), độ pH 7,5

- Môi trường lên men: 0,1% natri glutamat, 1% amoni clorua, 1,2% magie sunfat, 0,01% canxi clorua, sắt sunfat (20 mg/L), mangan sulfat (20 mg/L), kẽm sulfat (20 mg/L), đồng sunfat (5 mg/L), L-xystein (23 mg/L), alanin (24 mg/L), axit nicotinic (8 mg/L), biotin (45 µg/L), thiamin hydrochlorit (5 mg/L), adenin (30 mg/L), 1,9% axit photphoric (85%), 2,55% glucoza, 1,45% fructoza

Ví dụ 1-2: Các thí nghiệm về hàm lượng lên men của *CJI0323*

Môi trường nuôi cấy giống (2 mL) được pha vào các ống nghiệm (đường kính: 18 mm), sau đó được hấp khử trùng và lần lượt cấy chuyển với *ATCC6872* và *CJI0323*. Sau đó, các kết quả được nuôi cấy lắc ở 30°C trong 24 giờ và sau đó được sử dụng làm dung dịch nuôi cấy giống. Môi trường lên men (29 mL) được pha vào bình tam giác (250 mL) để lắc, hấp khử trùng ở 121°C trong 15 phút và dung dịch nuôi cấy giống (2 mL) được cấy vào đó và nuôi cấy trong 3 ngày. Các điều kiện nuôi cấy được đặt ở 170 vòng/phút, nhiệt độ 30°C và độ pH là 7,5.

Sau khi hoàn thành nuôi cấy, lượng IMP sản xuất được đo bằng HPLC

(SHIMADZU LC20A) và kết quả nuôi cấy được thể hiện trong Bảng 2 bên dưới.

Bảng 2

Chủng	IMP (g/L)
<i>ATCC6872</i>	0
<i>CJI0323</i>	9,52

Chủng *CJI0323* được định danh là *Corynebacterium Stationis* CN01-0323. Chủng này đã lưu giữ theo Hiệp ước Budapest tại Bảo tàng giống chuẩn Vi sinh vật Hàn Quốc (Korean Culture Center of Microorganisms: KCCM) vào ngày 7 tháng 11 năm 2017. Ngoài ra, chủng này được chỉ định với số lưu chủng là KCCM12151P.

Ví dụ 1-3: Phát hiện tổng hợp protein

Các điều kiện sàng lọc cho thấy sự ức chế tăng trưởng của chủng *CJI0323* đã được thiết lập bằng cách bổ sung IMP vào môi trường tối thiểu chứa 1,7% aga. Các plasmid của bảo tàng gen của chủng *ATCC6872* đã được biến nạp vào chủng *CJI0323* bằng phương pháp điện biến nạp (van der Rest *et al.* 1999), và những khuẩn lạc trong đó ức chế tăng trưởng được giải phóng trong điều kiện nuôi cấy được bổ sung lượng dư IMP được chọn. Plasmid thu được từ các khuẩn lạc được chọn và được phân tích bằng kỹ thuật giải trình tự. Kết quả là, một loại protein màng liên quan đến việc giải phóng sự ức chế tăng trưởng đã được xác định trong điều kiện bổ sung lượng dư IMP.

Một loại protein màng từ *Corynebacterium* đã được xác định dựa trên trình tự axit amin SEQ ID NO: 2 và trình tự nucleotit SEQ ID NO: 4 (NCBI GenBank: NZ_CP014279, WP_066795121, vận chuyển MFS). Protein màng đã được biết đến là protein vận chuyển MFS, nhưng chức năng cụ thể của nó chưa được xác nhận, và hơn nữa, chức năng của nó liên quan đến sản sinh IMP thì vẫn chưa được biết đến. Trong sáng chế này, protein màng được định danh là ImpE2 (WT).

Ví dụ 2: Xác định ImpE1 và ImpE2

Ví dụ 2-1: Xác nhận *impE1* và *impE2*

Để kiểm tra các chức năng của protein màng, ImpE2, cấu trúc gen của trình tự SEQ ID NO: 4 đã được xác nhận trong NCBI (NCBI GenBank: NZ_CP014279, WP_066795121, vận chuyển MFS). Do đó, đã xác nhận rằng phần khởi đầu 7 bp của ORF của SEQ ID NO: 4 (*impE2*) trùng lặp trong 7 bp với gen khác (NCBI GenBank:

NZ_CP014279, WP_066795119, điều hòa phiên mã), nằm ở xuôi dòng của *impE2*. Do chức năng của gen nằm ở ngược dòng của *impE2* và protein được mã hóa bởi gen chưa được xác nhận, nên trong sáng chế, protein được định danh là *ImpE1* (WT) (trình tự axit amin SEQ ID NO: 1 và trình tự nucleotit SEQ ID NO: 3).

Ví dụ 2-2: Điều chế vector thiếu *impE1*- hoặc *impE2*

Để xác nhận việc xóa *ImpE1* hoặc *ImpE2*, có liên quan đến việc giải phóng sự ức chế tăng trưởng gây ra bởi IMP như được xác định trong các ví dụ 1 và ví dụ 2-1, trong một chủng sản sinh IMP có thể làm giảm khả năng sản sinh IMP của nó, các nỗ lực đã được thực hiện để điều chế các vector thiếu trong mỗi gen.

Các đoạn gen để điều chế các vector đã thu được bằng cách phản ứng chuỗi polymeraza (Polymerase Chain Reaction: PCR) sử dụng ADN bộ gen của chủng *ATCC6872* làm khuôn mẫu.

Cụ thể, PCR cho *impE1* được thực hiện bằng cách sử dụng cặp mồi của SEQ ID NO: 5 và SEQ ID NO: 6 và cặp mồi của SEQ ID NO: 7 và SEQ ID NO: 8; và PCR cho *impE2* được thực hiện bằng cách sử dụng cặp mồi của SEQ ID NO: 9 và SEQ ID NO: 10 và cặp mồi của SEQ ID NO: 11 và SEQ ID NO: 12 (Bảng 3).

Bảng 3

SEQ ID NO	Mồi	Trình tự (5' đến 3')
5	<i>impE1</i> kop-1	GCTCTAGACGAGAAAGCTAAAGCCGGTGA
6	<i>impE1</i> kop-2	GTTTTTAGCTACCATTGTTACACCCCGTGCAA GTTT
7	<i>impE1</i> kop-3	GCACGGGGTGTAACAATGGTAGCTAAAACT CCACC
8	<i>impE1</i> kop-4	GCTCTAGAAATAGTTGGGGAAGTCCACTC
9	<i>impE2</i> kop-1	GCTCTAGACTTGGATGACCTGGTGGAAAA
10	<i>impE2</i> kop-2	CTTGGAGAAAATTTCTTACCATTCCAGTCCTTT CGT
11	<i>impE2</i> kop-3	GGACTGGAATGGTAGGAAATTTTCTCCAAGGG AAAT
12	<i>impE2</i> kop-4	GGACTAGTGGATTGTGTTGACGCACGATG
13	<i>impE1E2</i> kop-2	CTTGGAGAAAATTTCTGTTACACCCCGTGCAA GTTT
14	<i>impE1E2</i> kop-3	GCACGGGGTGTAACAGAAATTTTCTCCAAGGG AAAT

Cụ thể, các đoạn môi được sử dụng đã được điều chế dựa trên thông tin về gen của *Corynebacterium stationis* (ATCC6872) (Ngân hàng Gen NCBI: NZ_CP014279) được đăng ký tại Ngân hàng Gen NIH và các trình tự nucleotit liên kế sau đó.

PCR được thực hiện bằng cách biến tính ban đầu ở 94°C trong 5 phút; 25 chu kỳ bao gồm biến tính ở 94°C trong 30 giây, ủ ở 52°C trong 30 phút, và trùng hợp ở 72°C trong 1 phút; và trùng hợp cuối cùng ở 72°C trong 5 phút.

PCR chồng lắp được thực hiện bằng cách sử dụng hai đoạn gen *impE1*, được khuếch đại bằng cách sử dụng cặp môi của SEQ ID NO: 5 và SEQ ID NO: 6 và cặp môi của SEQ ID NO: 7 và SEQ ID NO: 8, làm mạch khuôn, và kết quả là thu được mẫu polynucleotit (1,8 kbp). Đoạn gen thu được được tách dòng thành vectơ pDZ được tuyển tính (patent Hàn Quốc số 10-0924065 và Công bố patent quốc tế số 2008-033001), được sắp đặt có thứ tự bằng enzym giới hạn (*XbaI*), và được sử dụng bằng T4 ligaza, và do đó vectơ *pDZ-ΔimpE1* đã được điều chế. Ngoài ra, phản ứng chuỗi polymeraza chồng lắp đã được thực hiện bằng cách sử dụng đoạn gen *impE2* được khuếch đại bằng cách sử dụng cặp môi của SEQ ID NO: 9 và SEQ ID NO: 10, và hai đoạn của gen *impE2* được khuếch đại bằng cách sử dụng cặp môi của SEQ ID NO: 11 và SEQ ID NO: 12, làm mạch khuôn, và kết quả là đã thu được mẫu polynucleotit (1,7 kbp). Đoạn gen thu được phân cắt với các enzym giới hạn *XbaI* và *SpeI*. Đoạn gen được tách dòng bằng cách sử dụng T4 ligaza vào vectơ *pDZ* được tuyển tính, đã phân cắt với enzym giới hạn (*XbaI*), và do đó vectơ *pDZ-ΔimpE2* đã được điều chế.

Ví dụ 2-3: Điều chế các vectơ thiếu tích hợp *impE1* và *impE2*

Vì các gen *impE1* và *impE2* mã hóa các protein liên quan đến việc giải phóng sự ức chế tăng trưởng gây ra bởi IMP bị chồng lắp, nên cần phải điều chỉnh đồng thời cả hai gen. Do đó, các nỗ lực đã được thực hiện để điều chế vectơ trong đó cả *impE1* và *impE2* đều thiếu.

Đối với phản ứng PCR gen *impE1* và *impE2*, cặp môi của SEQ ID NO: 5 và SEQ ID NO: 13 và cặp môi của SEQ ID NO: 14 và SEQ ID NO: 12 đã được sử dụng. Các đoạn môi được sử dụng đã được điều chế dựa trên thông tin về gen của *Corynebacterium Stationis* (ATCC6872) (ngân hàng Gen NCBI: NZ_CP014279) được đăng ký tại Ngân hàng Gen NIH và các trình tự nucleotit liên kế. PCR chồng lắp được thực hiện bằng cách sử dụng đoạn gen *impE1* được khuếch đại bằng cách sử dụng cặp môi của SEQ ID

NO: 5 và SEQ ID NO: 13, và hai đoạn của gen *impE2* được khuếch đại bằng cách sử dụng cặp mồi của SEQ ID NO: 14 và SEQ ID NO: 12 làm mạch khuôn, và kết quả là đã thu được mẫu polynucleotit (2,0 kbp). Các đoạn gen thu được phân cắt với *XbaI* và *SpeI*, tương ứng. Các đoạn gen được tách dòng bằng cách sử dụng T4 ligaza vào vectơ *pDZ* được tuyến tính, đã phân cắt với enzym giới hạn (*XbaI*), và do đó vectơ *pDZ-ΔimpE1E2* đã được điều chế.

Ví dụ 2-4: Điều chế các chủng thiếu *impE1* và *impE2*

Hai loại plasmit được điều chế trong ví dụ 2-2 và một loại plasmit được điều chế trong ví dụ 2-3 được chuyển vào trong chủng *CJI0323* bằng phương pháp điện biến nạp (sử dụng phương pháp biến đổi được thể hiện trong *Appl. Microbiol. Biotechnol.* (1999) 52: 541 đến 545). Các chủng trong đó vectơ được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp được lựa chọn đã trải qua lai chéo lần thứ hai. Sự thiếu hụt di truyền ở các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng các cặp mồi SEQ ID NO: 5 và SEQ ID NO: 8, SEQ ID NO: 9 và SEQ ID NO: 12, và SEQ ID NO: 5 và SEQ ID NO: 12.

Các chủng được chọn được định danh là *CJI0323_ΔimpE1*, *CJI0323_ΔimpE2* và *CJI0323_ΔimpE1E2*. Ngoài ra, khả năng sản sinh IMP của các chủng này đã được đánh giá.

Môi trường nuôi cấy giống (2 mL) đã được pha vào các ống nghiệm (đường kính: 18 mm), sau đó được hấp khử trùng, lần lượt được cấy chuyển với *CJI0323*, *CJI0323_ΔimpE1*, *CJI0323_ΔimpE2* và *CJI0323_ΔimpE1E2* nuôi cấy lắc ở nhiệt độ 30°C trong 24 giờ, và được sử dụng làm dung dịch nuôi cấy giống. Môi trường lên men (29 mL) được pha vào bình tam giác (250 mL) để lắc và hấp khử trùng ở nhiệt độ 121°C trong 15 phút. Sau đó, dung dịch nuôi cấy giống (2 mL) được bổ sung vào đó và kết quả được nuôi cấy trong 3 ngày. Các điều kiện nuôi cấy được đặt ở 170 vòng/phút, 30°C và độ pH là 7,5.

Sau khi hoàn thành nuôi cấy, lượng IMP sản sinh được đo bằng phương pháp đo sắc ký lỏng hiệu suất cao (High Performance Liquid Chromatography: HPLC) và kết quả nuôi cấy được thể hiện trong Bảng 4 dưới đây.

Bảng 4

Chủng	IMP (g/L)
<i>CJI0323</i>	9,52
<i>CJI0323_ΔimpE1</i>	1,92
<i>CJI0323_ΔimpE2</i>	1,88
<i>CJI0323_ΔimpE1E2</i>	1,80

Lượng IMP tích lũy trong mỗi chủng được so sánh với chủng mẹ, *Corynebacterium Stationis CJI0323*. Kết quả là, có thể thấy rằng, như thể hiện trong Bảng 4 ở trên, nồng độ IMP của các chủng *CJI0323_ΔimpE1*, *CJI0323_ΔimpE2* và *CJI0323_ΔimpE1E2* đã giảm khoảng 8 g/L trong cùng điều kiện so với chủng bố mẹ, *ImpE1* và *ImpE2* là các protein liên quan đến việc sản sinh IMP.

Ví dụ 3: Xác nhận trình tự nucleotit của *impE1* và *impE2* của chủng sản sinh IMP, *CJI0323*

Trong trường hợp chủng *CJI0323* tạo ra IMP ở nồng độ cao trong ví dụ 1, có thể chủng này có khả năng tạo ra IMP được cải thiện để sản sinh IMP ở nồng độ cao. Theo đó, nỗ lực đã được thực hiện để xác nhận sự hiện diện của bất kỳ đột biến nào trong *impE1* và *impE2* của chủng *CJI0323*.

ADN nhiễm sắc thể của chủng *CJI0323* được khuếch đại bằng phản ứng chuỗi polymeraza (sau đây gọi là “PCR”). Cụ thể, trước tiên, PCR được thực hiện bằng cách lặp lại 28 chu kỳ bao gồm biến tính ở 94°C trong 1 phút, ủ ở 58°C trong 30 giây, và trùng hợp ở 72°C trong 2 phút bằng ADN nhiễm sắc thể của chủng *CJI0323* làm mạch khuôn cùng với cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16 (Bảng 5), và do đó, đoạn khoảng 2,8 kbp đã được khuếch đại.

Bảng 5

SEQ ID NO	Mồi	Trình tự (5' to 3')
15	<i>impE1E2 seqF</i>	GAACGGAGTCATCTCCTTTGC
16	<i>impE1E2 seqR</i>	CCAAACGCTCTGCAAGAACTG

Sau khi phân tích trình tự nucleotit bằng cách sử dụng cùng đoạn mồi, người ta đã xác nhận rằng nucleotit thứ 490 của gen *impE1* (ví dụ, g) đã được thay thế bằng ‘a’, so với trình tự nucleotit của chủng loại tự nhiên, ATCC6872. Sự thay thế này chỉ ra rằng có sự điều chỉnh trong đó axit amin thứ 164 của protein *ImpE1* (tức là axit glutamic) đã được thay thế bằng lysin.

Ngoài ra, có thể xác nhận rằng nucleotit thứ 4 của gen *impE2* (nghĩa là g) đã được thay thế bằng 'a' (điều này có nghĩa là nucleotit thứ 666 của gen *impE1* (nghĩa là g) đã được thay thế bằng 'a') và nucleotit thứ 191 của gen *impE1* (nghĩa là g) đã được thay thế bằng 'a'. Những sự thay thế này chỉ ra rằng có những sửa đổi trong đó axit amin thứ 2 của protein *ImpE2* (tức là valin), tương ứng với axit amin thứ 222 của protein *ImpE1*, được thay thế bằng isoleuxin; và axit amin thứ 64 của protein *ImpE2* (tức là glyxin) đã được thay thế bằng axit glutamic.

Nucleotit *impE1* của chủng *CJI0323* được định danh là *impE1_CJI0323* (ID SEQ NO: 143) và protein của chúng được định danh là *ImpE1_CJI0323* (ID SEQ NO: 141), trong khi đó, nucleotit *impE2* của chủng *CJI0323* và protein của chúng được định danh là *ImpE2_CJI0323* (ID SEQ NO: 142).

Ví dụ 4: Thu hồi các sửa đổi trong *impE1* và *impE2*

Ví dụ 4-1: Điều chế các vector để thu hồi các biến đổi trong *impE1* hoặc *impE2*

Trong ví dụ 3, sự hiện diện của bất kỳ biến đổi nào trong *impE1* và *impE2* của chủng sản sinh IMP *CJI0323* đã được kiểm tra. Kết quả là, đã xác nhận rằng *impE1* có một biến đổi và *impE2* có hai biến đổi. Do chủng *CJI0323* tạo ra IMP ở nồng độ cao, nên rất có khả năng biến đổi là loại có thể cải thiện khả năng sản sinh IMP. Theo đó, sau khi thu hồi *impE1* và *impE2* bị đột biến thành *ImpE* kiểu hoang dã tự nhiên mà không biến đổi, thử nghiệm sau đây đã được thực hiện để xác nhận xem mỗi sửa đổi có thực sự mang lại khả năng sản sinh IMP hay không.

Để điều chế vector thu hồi, PCR đã được thực hiện bằng cách sử dụng *Corynebacterium Stationis ATCC6872* làm mạch khuôn.

Đoạn gen *impE1impE2* được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 17 và SEQ ID NO: 18 đã được xử lý bằng enzym giới hạn *XbaI*, và được tách dòng thành vị trí giới hạn *XbaI* trên vector *pDZ*, và do đó *pDZ-impE1E2 (WT)* đã được điều chế.

Ví dụ 4-2: Điều chế các vector với biến đổi duy nhất trong *impE1* hoặc *impE2*

Vector có biến đổi E164K duy nhất trong gen *ImpE1* đã được điều chế bằng cách sử dụng chủng loại hoang dã bản địa *Corynebacterium Stationis ATCC6872* làm mạch khuôn cùng với cặp mồi SEQ ID NO: 19 và SEQ ID NO: 20 và cặp mồi SEQ ID NO:

21 và SEQ ID NO: 22. PCR chồng lấp được thực hiện bằng cách sử dụng đoạn gen E164K-1 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 19 và SEQ ID NO: 20 và hai đoạn gen E164K-2 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 21 và SEQ ID NO: 22, và nhờ đó thu được mạch khuôn có polynucleotit 1,8 kbp. Các đoạn gen thu được đã được phân cắt với *XbaI* và được tách dòng thành vector *pDZ* được tuyến tính, đã được phân cắt với *XbaI* sử dụng T4 ligaza, và nhờ đó vector *pDZ-impE1 (E164K)* được điều chế.

Vector có biến đổi V2I duy nhất trong gen *ImpE2* đã được điều chế bằng cách sử dụng chủng *ATCC6872* làm mạch khuôn cùng với cặp mồi SEQ ID NO: 19 và SEQ ID NO: 23 và cặp mồi SEQ ID NO: 24 và SEQ ID NO: 22. PCR chồng lấp được thực hiện bằng cách sử dụng đoạn gen V2I-1 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 19 và SEQ ID NO: 23, và hai đoạn gen V2I-2 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 24 và SEQ ID NO: 22, và nhờ đó thu được mạch khuôn có polynucleotit 1,8 kbp. Các đoạn gen thu được đã được phân cắt với *XbaI* và được tách dòng thành vector *pDZ* được tuyến tính hóa, đã được phân cắt với *XbaI* sử dụng T4 ligaza, và nhờ đó vector *pDZ-impE2 (V2I)* được điều chế.

Vector có biến đổi G64E duy nhất trong gen *ImpE2* đã được điều chế bằng cách sử dụng chủng *ATCC6872* làm mạch khuôn cùng với cặp mồi SEQ ID NO: 19 và SEQ ID NO: 25, và cặp mồi SEQ ID NO: 26 và SEQ ID NO: 22. PCR chồng lấp được thực hiện bằng cách sử dụng đoạn gen G64E-1 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 19 và SEQ ID NO: 25, và hai đoạn gen G64E-2 được khuếch đại bằng cách sử dụng cặp mồi SEQ ID NO: 26 và SEQ ID NO: 22, và nhờ đó thu được mạch khuôn có polynucleotit 1,8 kbp. Các đoạn gen thu được đã được phân cắt với *XbaI* và được tách dòng thành vector *pDZ* được tuyến tính hóa, đã được phân cắt với *XbaI* sử dụng T4 ligaza, và nhờ đó vector *pDZ-impE2 (G64E)* được điều chế.

Bảng 6

SEQ ID NO	Đoạn mồi	Tình tự (5' đến 3')
17	impE1E2 WT F	GCTCTAGAGAACGGAGTCATCTCCTTTGC
18	impE1E2 WT R	GCTCTAGACCAAACGCTCTGCAAGAAACTG
19	impE1 164K-1	GCTCTAGACTTGGATGACCTGGTGGAAAA
20	impE1 164K-2	CTGGGGCGCGTGTGTTTTTCAGGATGCTCCCG AAGACG

21	impE1 164K-3	AACAACGCGCCCCAGAATTGG
22	impE1 164K-4	GCTCTAGAAATAGTTGGGGAAGTCCACTC
23	impE2 V2I-2	TGGAGTTTTTAGCTATCATTCAGTCCTTTC GTGTAA
24	impE2 V2I-3	TAGCTAAAACTCCACCCCAA
25	impE2 G64E-2	CCGAAAATCATCTGCTCCAAAGAGCTCATC AGCATGG
26	impE2 G64E-3	GCAGATGATTTTCGGTTCCGC

Ví dụ 4-3: Thu hồi các biến đổi *impE1*, *impE2* và điều chế các chủng có biến đổi duy nhất

Plasmit được điều chế trong ví dụ 4-1 đã được chuyển vào trong chủng CJI0323 bằng phương pháp điện biến nạp (sử dụng phương pháp biến đổi được thể hiện trong Appl. Microbiol. Biotechnol. (1999) 52: 541 đến 545). Các chủng trong đó vectơ được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp đã được lựa chọn trải qua lai chéo lần thứ hai. Việc thu hồi biến đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó phân tích trình tự nucleotit. Chủng đã điều chế được định danh là *CJI0323_impE1E2* (WT).

Ba loại plasmit được điều chế trong ví dụ 4-2 được chuyển vào chủng *CJI0323_impE1E2* (WT) bằng phương pháp điện biến nạp (sử dụng phương pháp biến đổi được thể hiện trong Appl. Microbiol. Biotechnol. (1999) 52: 541 đến 545). Các chủng trong đó vectơ được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp đã được lựa chọn trải qua lai chéo lần thứ hai. Việc đưa vào biến đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó là phân tích trình tự nucleotit. Các chủng được chọn được định danh là *CJI0323_impE1* (E164K), *CJI0323_impE2* (V2I) và *CJI0323_impE2* (G64E).

Các chủng *Corynebacterium Stationis* CJI0323_impE1 (E164K), *Corynebacterium Stationis* CJI0323_impE2 (V2I) và *Corynebacterium Stationis* CJI0323_impE2 (G64E) đã được lưu giữ theo Hiệp ước Budapest vào Bảo tàng giống chuẩn Vi sinh vật Hàn Quốc (KCCM) vào ngày 2 tháng 11 năm 2018. Ngoài ra, các

chúng đã được chỉ định với các số truy cập KCCM12359P, KCCM12360P và KCCM12361P, tương ứng.

Ví dụ 4-4: Điều chế các chủng biến đổi tích hợp *impE1*- và *impE2*

Các plasmid *pDZ-impE2 (V2I)* và *pDZ-impE2 (G64E)* được điều chế trong ví dụ 4-2 đã được biến nạp vào chủng *CJI0323_impE1 (E164K)* bằng phương pháp điện biến nạp (sử dụng phương pháp biến đổi được thể hiện trong Appl. Microbiol. Biotechnol. 54:1 đến 545). Các chủng trong đó các vector được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp được lựa chọn đã trải qua lai chéo lần thứ hai. Việc đưa vào biến đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó là phân tích trình tự nucleotit. Các chủng đã điều chế được định danh là *CJI0323_impE1 (E164K)_impE2 (V2I)* và *CJI0323_impE1 (E164K)_impE2 (G64E)*.

Plasmid *pDZ-impE2 (G64E)* đã được biến nạp vào chủng *CJI0323_impE2 (V2I)* bằng phương pháp điện biến nạp (sử dụng phương pháp biến đổi được thể hiện trong Appl. Microbiol. Biotechnol. (1999) 52: 541 đến 545). Các chủng trong đó vector được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp đã được lựa chọn trải qua lai chéo lần thứ hai. Việc đưa vào biến đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó phân tích trình tự nucleotit. Chủng được chọn được định danh là *CJI0323_impE2 (V2I) (G64E)*.

Ví dụ 4-5: Đánh giá các chủng có biến đổi *impE1*, *impE2*

Môi trường nuôi cấy giống (2 mL) đã được pha vào các ống nghiệm (đường kính: 18 mm), sau đó được khử trùng, mỗi mẫu được cấy với *CJI0323_impE1E2 (WT)*, *CJI0323_impE1E2 (WT)*, *CJI0323_impE1 (E164K)*, *CJI0323_impE1 (E164K)_impE2 (V2I)*, *CJI0323_impE1 (E164K)_impE2 (G64E)* và *CJI0323_impE2 (V2I) (G64E)*, nuôi cấy lắc ở 30°C trong 24 giờ và được sử dụng làm dung dịch nuôi cấy giống. Môi trường lên men (29 mL) được pha vào bình tam giác (250 mL) để lắc và hấp khử trùng ở 121°C trong 15 phút. Sau đó, các dung dịch nuôi cấy giống (2 mL) đã được chủng vào đó và kết quả được nuôi cấy trong 3 ngày. Các điều kiện nuôi cấy được đặt ở 170 vòng/phút,

nhiệt độ 30°C và độ pH là 7,5.

Sau khi hoàn thành nuôi cấy, lượng IMP sản sinh được đo bằng HPLC và kết quả nuôi cấy được thể hiện trong Bảng 7 dưới đây.

Bảng 7

Chủng	IMP (g/L)
<i>CJI0323</i>	9,52
<i>CJI0323_impE1E2(WT)</i>	2,32
<i>CJI0323_impE1(E164K)</i>	2,57
<i>CJI0323_impE2(V2I)</i>	3,11
<i>CJI0323_impE2(G64E)</i>	3,27
<i>CJI0323_impE1(E164K)_impE2(V2I)</i>	4,24
<i>CJI0323_impE1(E164K)_impE2(G64E)</i>	6,27
<i>CJI0323_impE2(V2I)(G64E)</i>	7,35

Như được thể hiện ở trên, có thể xác nhận rằng đối với từng vị trí biến đổi, một loại biến đổi, tích hợp hai loại biến đổi, và tích hợp ba loại biến đổi đều liên quan đến việc sản sinh IMP. Theo đó, trong vi sinh vật thuộc chi *Corynebacterium* không tạo ra IMP hoặc chỉ tạo ra một lượng nhỏ IMP, sự gia tăng số lượng IMP được tổng hợp do biến đổi protein *ImpE* theo sáng chế là rất có ý nghĩa.

Ví dụ 5: Thay thế các axit amin trong biến đổi *impE1*, *impE2* bằng các axit amin khác

Ví dụ 5-1: Điều chế các vector để chèn thay thế các axit amin trong các biến đổi *impE1*, *impE2*

Để xác nhận tầm quan trọng vị trí của ba loại biến đổi (ví dụ, *impE1* (E164K), *impE2* (V2I) và *impE2* (G64E)) với các khả năng nâng cao để sinh ra IMP như được xác định trong các kết quả nêu trên, vector được đưa vào các biến đổi (ví dụ, biến đổi thay thế axit amin thứ 164 trong trình tự axit amin của *impE1*, axit amin thứ 2 trong trình tự axit amin của *impE2*, và axit amin thứ 64 trong trình tự axit amin của *impE2* bằng axit amin khác) đã được điều chế.

Trước tiên, quy trình điều chế vector để đưa vào biến đổi *ImpE1* (E164K) như sau.

Dựa trên các trình tự polynucleotit được báo cáo, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập, thu được các đoạn gen bằng cách

thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* làm mạch khuôn cùng với các cặp mồi giữa đoạn mồi SEQ ID NO: 27 và mỗi đoạn mồi SEQ ID NO: 28 đến SEQ ID NO: 45. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm biến tính ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong 30 giây, và trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 0,7 kbp. Sau đó, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập, và thu được các đoạn gen bằng cách thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* làm mạch khuôn cùng với các cặp mồi giữa đoạn mồi SEQ ID NO: 46 và mỗi đoạn mồi SEQ ID NO: 47 đến SEQ ID NO: 64. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm biến tính ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong 30 giây và trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 0,7 kbp.

PCR chồng lắp được thực hiện bằng cách sử dụng hai đoạn thu được từ các kết quả bên trên dưới dạng mạch khuôn và do đó thu được 18 loại polynucleotit 1,4 kbp được sử dụng làm mạch khuôn. Các đoạn gen thu được được phân cắt với enzym giới hạn *SpeI*, được gắn vào vector *pDZ* đã được tuyến tính đã được phân cắt với enzym hạn chế *XbaI*, được biến nạp vào trong *E. coli* DH5 α , và các biến nạp được cấy trải trên môi trường LB đã đông đặc có chứa kanamycin (25 mg/L).

Thông tin trình tự về các đoạn mồi được sử dụng để điều chế vector được thể hiện trong Bảng 8 dưới đây.

Bảng 8

SEQ ID NO	Đoạn mồi	Trình tự (5' đến 3')
27	<i>SpeI</i> -impE1 164 1F	GGGACTAGTGATTCCGGCCAACTGTCTG
28	impE1 164-R 1R	TGGGGCGCGTTGGCGTTCAGGATGCTC
29	impE1164-H 1R	TGGGGCGCGTTGGTGTTCAGGATGCTC
30	impE1 164-D 1R	TGGGGCGCGTTGATCTTCAGGATGCTC
31	impE1 164-S 1R	TGGGGCGCGTTGGGATTCAGGATGCTC
32	impE1 164-T 1R	TGGGGCGCGTTGGGTTTCAGGATGCTC
33	impE1 164-N 1R	TGGGGCGCGTTGGTTTTTCAGGATGCTC

34	impE1 164-Q 1R	TGGGGCGCGTTGCTGTTCAGGATGCTC
35	impE1 164-C 1R	TGGGGCGCGTTGGCATTTCAGGATGCTC
36	impE1 164-G 1R	TGGGGCGCGTTGGCCTTCAGGATGCTC
37	impE1 164-P 1R	TGGGGCGCGTTGCGGTTCAGGATGCTC
38	impE1 164-A 1R	TGGGGCGCGTTGGGCTTCAGGATGCTC
39	impE1 164-V 1R	TGGGGCGCGTTGGACTTCAGGATGCTC
40	impE1 164-I 1R	TGGGGCGCGTTGGATTTCAGGATGCTC
41	impE1 164-L 1R	TGGGGCGCGTTGCAGTTCAGGATGCTC
42	impE1 164-M 1R	TGGGGCGCGTTGCATTTCAGGATGCTC
43	impE1 164-F 1R	TGGGGCGCGTTGGAATTCAGGATGCTC
44	impE1 164-Y 1R	TGGGGCGCGTTGGTATTCAGGATGCTC
45	impE1 164-W 1R	TGGGGCGCGTTGCCATTCAGGATGCTC
46	Spe1-impE1 164 2R	GGGACTAGTCATGAACTTGCCGCGCTC
47	impE1 164-R 2F	GAGCATCCTGAACGCCAACGCGCCCCA
48	impE1 164-H 2F	GAGCATCCTGAACACCAACGCGCCCCA
49	impE1 164-D 2F	GAGCATCCTGAAGATCAACGCGCCCCA
50	impE1 164-S 2F	GAGCATCCTGAATCCCAACGCGCCCCA
51	impE1 164-T 2F	GAGCATCCTGAAACCCAACGCGCCCCA
52	impE1 164-N 2F	GAGCATCCTGAAAACCAACGCGCCCCA
53	impE1 164-Q 2F	GAGCATCCTGAACAGCAACGCGCCCCA
54	impE1 164-C 2F	GAGCATCCTGAATGCCAACGCGCCCCA
55	impE1 164-G 2F	GAGCATCCTGAAGGCCAACGCGCCCCA
56	impE1 164-P 2F	GAGCATCCTGAACCGCAACGCGCCCCA
57	impE1 164-A 2F	GAGCATCCTGAAGCCCAACGCGCCCCA
58	impE1 164-V 2F	GAGCATCCTGAAGTCCAACGCGCCCCA
59	impE1 164-I 2F	GAGCATCCTGAAATCCAACGCGCCCCA
60	impE1 164-L 2F	GAGCATCCTGAACTGCAACGCGCCCCA
61	impE1 164-M 2F	GAGCATCCTGAAATGCAACGCGCCCCA
62	impE1 164-F 2F	GAGCATCCTGAATTCCAACGCGCCCCA
63	impE1 164-Y 2F	GAGCATCCTGAATACCAACGCGCCCCA
64	impE1 164-W 2F	GAGCATCCTGAATGGCAACGCGCCCCA

Sau khi chọn lọc bằng phương pháp PCR các khuẩn lạc được biến nạp với vector mà gen đích được chèn vào, các plasmit thu được bằng phương pháp tách chiết plasmit thông thường. Thông tin về các plasmit thu được được hiển thị trong Bảng 9 bên dưới.

Bảng 9

Số	Plasmit
1	<i>pDZ-impE1 164R</i>
2	<i>pDZ-impE1 164H</i>
3	<i>pDZ-impE1 164D</i>
4	<i>pDZ-impE1 164S</i>
5	<i>pDZ-impE1 164T</i>
6	<i>pDZ-impE1 164N</i>
7	<i>pDZ-impE1 164Q</i>
8	<i>pDZ-impE1 164C</i>
9	<i>pDZ-impE1 164G</i>
10	<i>pDZ-impE1 164P</i>
11	<i>pDZ-impE1 164A</i>
12	<i>pDZ-impE1 164V</i>
13	<i>pDZ-impE1 164I</i>
14	<i>pDZ-impE1 164L</i>
15	<i>pDZ-impE1 164M</i>
16	<i>pDZ-impE1 164F</i>
17	<i>pDZ-impE1 164Y</i>
18	<i>pDZ-impE1 164W</i>

Thứ hai, quy trình điều chế vector để đưa vào ImpE2 (V2I) như sau.

Dựa trên các trình tự polynucleotit được báo cáo, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập và thu được các đoạn gen bằng cách thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* làm mạch khuôn cùng với cặp mồi của đoạn mồi SEQ ID NO: 65 và mỗi đoạn mồi SEQ ID NO: 66 đến SEQ ID NO: 83. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm biến tính ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong 30 giây và trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 0,7 kbp. Sau đó, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập, và thu được các đoạn gen bằng cách thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* làm mạch khuôn cùng với cặp mồi giữa đoạn mồi SEQ ID NO: 84 và mỗi đoạn mồi SEQ ID NO: 85 đến SEQ ID NO: 102. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm biến tính ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong 30 giây và trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 0,7 kbp.

PCR chồng lắp được thực hiện bằng cách sử dụng hai đoạn thu được từ các kết quả bên trên làm mạch khuôn mẫu và do đó 18 loại polynucleotit 1,4 kbp thu được được sử dụng làm mạch khuôn. Các đoạn gen thu được được phân cắt với enzym giới hạn *XbaI*, được gắn vào vector *pDZ* được tuyển tính, đã được phân cắt với enzym giới hạn *XbaI*, biến nạp vào *E. coli* DH5 α , sau đó được cấy trên môi trường LB đông đặc có chứa kanamycin (25 mg/L).

Thông tin trình tự về các đoạn mồi được sử dụng để điều chế vector được thể hiện trong Bảng 10 dưới đây.

Bảng 10

SEQ ID NO	Đoạn mồi	Trình tự (5' đến 3')
65	XbaI-impE2 2 1F	GGGTCTAGATTGCATGCTGTGCAAGA
66	impE2 2-R 1R	GGAGTTTTTTAGCGCGCATTCCAGTCCT
67	impE2 2-H 1R	GGAGTTTTTTAGCGTGCATTCCAGTCCT
68	impE2 2-K 1R	GGAGTTTTTTAGCCTTCATTCCAGTCCT
69	impE2 2-D 1R	GGAGTTTTTTAGCGTCCATTCCAGTCCT
70	impE2 2-E 1R	GGAGTTTTTTAGCTTCCATTCCAGTCCT
71	impE2 2-S 1R	GGAGTTTTTTAGCGGACATTCCAGTCCT
72	impE2 2-T 1R	GGAGTTTTTTAGCGGTCATTCCAGTCCT
73	impE2 2-N 1R	GGAGTTTTTTAGCGTTCATTCCAGTCCT
74	impE2 2-Q 1R	GGAGTTTTTTAGCCTGCATTCCAGTCCT
75	impE2 2-C 1R	GGAGTTTTTTAGCGCACATTCCAGTCCT
76	impE2 2-G 1R	GGAGTTTTTTAGCGCCCATTCAGTCCT
77	impE2 2-P 1R	GGAGTTTTTTAGCTGGCATTCCAGTCCT
78	impE2 2-A 1R	GGAGTTTTTTAGCAGCCATTCCAGTCCT
79	impE2 2-L 1R	GGAGTTTTTTAGCCAGCATTCCAGTCCT
80	impE2 2-M 1R	GGAGTTTTTTAGCCATCATTCCAGTCCT
81	impE2 2-F 1R	GGAGTTTTTTAGCGAACATTCCAGTCCT
82	impE2 2-Y 1R	GGAGTTTTTTAGCGTACATTCCAGTCCT
83	impE2 2-W 1R	GGAGTTTTTTAGCCCACATTCCAGTCCT
84	XbaI-impE2 2 2R	GGGTCTAGATTGCTCGCCACGCGCA
85	impE2 2-R 2F	AGGACTGGAATGCGCGCTAAAAACTCC
86	impE2 2-H 2F	AGGACTGGAATGCACGCTAAAAACTCC
87	impE2 2-K 2F	AGGACTGGAATGAAGGCTAAAAACTCC
88	impE2 2-D 2F	AGGACTGGAATGGACGCTAAAAACTCC

89	impE2 2-E 2F	AGGACTGGAATGGAAGCTAAAAACTCC
90	impE2 2-S 2F	AGGACTGGAATGTCCGCTAAAAACTCC
91	impE2 2-T 2F	AGGACTGGAATGACCGCTAAAAACTCC
92	impE2 2-N 2F	AGGACTGGAATGAACGCTAAAAACTCC
93	impE2 2-Q 2F	AGGACTGGAATGCAGGCTAAAAACTCC
94	impE2 2-C 2F	AGGACTGGAATGTGCGCTAAAAACTCC
95	impE2 2-G 2F	AGGACTGGAATGGGCGCTAAAAACTCC
96	impE2 2-P 2F	AGGACTGGAATGCCAGCTAAAAACTCC
97	impE2 2-A 2F	AGGACTGGAATGGCTGCTAAAAACTCC
98	impE2 2-L 2F	AGGACTGGAATGCTGGCTAAAAACTCC
99	impE2 2-M 2F	AGGACTGGAATGATGGCTAAAAACTCC
100	impE2 2-F 2F	AGGACTGGAATGTTCGCTAAAAACTCC
101	impE2 2-Y 2F	AGGACTGGAATGTACGCTAAAAACTCC
102	impE2 2-W 2F	AGGACTGGAATGTGGGCTAAAAACTCC

Sau khi chọn bằng phương pháp PCR các khuẩn lạc được biến đổi với vector mà gen đích được chèn vào, các plasmit thu được bằng phương pháp tách chiết plasmit thông thường. Thông tin về các plasmit thu được được thể hiện trong Bảng 11 dưới đây.

Bảng 11

Số	Plasmit
1	<i>pDZ-impE2 2R</i>
2	<i>pDZ-impE2 2H</i>
3	<i>pDZ-impE2 2K</i>
4	<i>pDZ-impE2 2D</i>
5	<i>pDZ-impE2 2E</i>
6	<i>pDZ-impE2 2S</i>
7	<i>pDZ-impE2 2T</i>
8	<i>pDZ-impE2 2N</i>
9	<i>pDZ-impE2 2Q</i>
10	<i>pDZ-impE2 2C</i>
11	<i>pDZ-impE2 2G</i>
12	<i>pDZ-impE2 2P</i>
13	<i>pDZ-impE2 2A</i>
14	<i>pDZ-impE2 2L</i>
15	<i>pDZ-impE2 2M</i>
16	<i>pDZ-impE2 2F</i>

17	<i>pDZ-impE2 2Y</i>
18	<i>pDZ-impE2 2W</i>

Cuối cùng, quy trình điều chế vector để đưa vào ImpE2 (G64E) như sau.

Dựa trên các trình tự polynucleotit được báo cáo, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập, và thu được các đoạn gen bằng cách thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* là biến tính mạch khuôn ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong vòng 30 giây và cùng với các cặp mồi giữa đoạn mồi SEQ ID NO: 103 và mỗi đoạn mồi SEQ ID NO: 104 đến SEQ ID NO: 121. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm biến tính ở nhiệt độ 94°C trong 30 giây, ủ ở nhiệt độ 55°C trong 30 giây và trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 1 kbp. Sau đó, các gen nhiễm sắc thể của *Corynebacterium Stationis CJI0323* đã được phân lập, và thu được các đoạn gen bằng cách thực hiện PCR sử dụng ADN nhiễm sắc thể của *Corynebacterium Stationis CJI0323* làm mạch khuôn cùng với cặp mồi giữa đoạn mồi SEQ ID NO: 84 và mỗi đoạn mồi SEQ ID NO: 85 đến SEQ ID NO: 102. PCR được thực hiện bằng cách biến tính ban đầu ở nhiệt độ 94°C trong 5 phút; 20 chu kỳ bao gồm trùng hợp ở nhiệt độ 72°C trong 1 phút; và trùng hợp cuối cùng ở nhiệt độ 72°C trong 5 phút. Kết quả là thu được 18 loại polynucleotit 1 kbp.

PCR chồng lắp được thực hiện bằng cách sử dụng hai đoạn thu được từ các kết quả bên trên làm mạch khuôn, và nhờ đó thu được 18 loại polynucleotit 2 kbp được sử dụng làm mạch khuôn. Các đoạn gen thu được được phân cắt với enzym giới hạn *XbaI*, được gắn vào vector pDZ được tuyến tính đã được phân cắt với enzym giới hạn *XbaI*, biến nạp vào trong *E. coli DH5α*, và sau đó được cấy trải trên môi trường LB đông đặc có chứa kanamycin (25 mg/L).

Thông tin trình tự về các đoạn mồi được sử dụng để điều chế vector được thể hiện trong Bảng 12 dưới đây.

Bảng 12

SEQ ID NO	Đoạn mồi	Trình tự (5' đến 3')
103	XbaI-impE2 64 1F	GGGTCTAGAAAAGAGCTTAAGGCAGCT GCT

104	impE2 64-R 1R	GAAAATCATCTGGCGCAAAGAGCTCAT
105	impE2 64-H 1R	GAAAATCATCTGGTGCAAAGAGCTCAT
106	impE2 64-D 1R	GAAAATCATCTGGTCCAAAGAGCTCAT
107	impE2 64-K 1R	GAAAATCATCTGCTTCAAAGAGCTCAT
108	impE2 64-S 1R	GAAAATCATCTGGGACAAAGAGCTCAT
109	impE2 64-T 1R	GAAAATCATCTGGGTCAAAGAGCTCAT
110	impE2 64-N 1R	GAAAATCATCTGGTTCAAAGAGCTCAT
111	impE2 64-Q 1R	GAAAATCATCTGCTGCAAAGAGCTCAT
112	impE2 64-C 1R	GAAAATCATCTGGCACAAAGAGCTCAT
113	impE2 64-P 1R	GAAAATCATCTGTGGCAAAGAGCTCAT
114	impE2 64-A 1R	GAAAATCATCTGAGCCAAAGAGCTCAT
115	impE2 64-V 1R	GAAAATCATCTGGACCAAAGAGCTCAT
116	impE2 64-I 1R	GAAAATCATCTGGATCAAAGAGCTCAT
117	impE2 64-L 1R	GAAAATCATCTGCAGCAAAGAGCTCAT
118	impE2 64-M 1R	GAAAATCATCTGCATCAAAGAGCTCAT
119	impE2 64-F 1R	GAAAATCATCTGGAACAAAGAGCTCAT
120	impE2 64-Y 1R	GAAAATCATCTGGTACAAAGAGCTCAT
121	impE2 64-W 1R	GAAAATCATCTGCCACAAAGAGCTCAT
122	XbaI-impE2 64 2R	GGGTCTAGACGGTCAATGAAGTCTCAAC GG
123	impE2 64-R 2F	ATGAGCTCTTTGCGCCAGATGATTTTC
124	impE2 64-H 2F	ATGAGCTCTTTGCACCAGATGATTTTC
125	impE2 64-D 2F	ATGAGCTCTTTGGACCAGATGATTTTC
126	impE2 64-K 2F	ATGAGCTCTTTGAAGCAGATGATTTTC
127	impE2 64-S 2F	ATGAGCTCTTTGTCCCAGATGATTTTC
128	impE2 64-T 2F	ATGAGCTCTTTGACCCAGATGATTTTC
129	impE2 64-N 2F	ATGAGCTCTTTGAACCAGATGATTTTC
130	impE2 64-Q 2F	ATGAGCTCTTTGCAGCAGATGATTTTC
131	impE2 64-C 2F	ATGAGCTCTTTGTGCCAGATGATTTTC
132	impE2 64-P 2F	ATGAGCTCTTTGCCACAGATGATTTTC
133	impE2 64-A 2F	ATGAGCTCTTTGGCTCAGATGATTTTC
134	impE2 64-V 2F	ATGAGCTCTTTGGTCCAGATGATTTTC
135	impE2 64-I 2F	ATGAGCTCTTTGATCCAGATGATTTTC
136	impE2 64-L 2F	ATGAGCTCTTTGCTGCAGATGATTTTC
137	impE2 64-M 2F	ATGAGCTCTTTGATGCAGATGATTTTC
138	impE2 64-F 2F	ATGAGCTCTTTGTTCCAGATGATTTTC
139	impE2 64-Y 2F	ATGAGCTCTTTGTACCAGATGATTTTC

140	impE2 64-W 2F	ATGAGCTCTTTGTGGCAGATGATTTTC
-----	---------------	-----------------------------

Sau khi chọn bằng phương pháp PCR các khuẩn lạc biến đổi với vector mà gen đích được đưa vào, các plasmit thu được bằng phương pháp tách chiết plasmit thông thường. Thông tin về các plasmit thu được được thể hiện trong Bảng 13 dưới đây.

Bảng 13

Số	Plasmit
1	<i>pDZ-impE2 64R</i>
2	<i>pDZ-impE2 64H</i>
3	<i>pDZ-impE2 64D</i>
4	<i>pDZ-impE2 64K</i>
5	<i>pDZ-impE2 64S</i>
6	<i>pDZ-impE2 64T</i>
7	<i>pDZ-impE2 64N</i>
8	<i>pDZ-impE2 64Q</i>
9	<i>pDZ-impE2 64C</i>
10	<i>pDZ-impE2 64P</i>
11	<i>pDZ-impE2 64A</i>
12	<i>pDZ-impE2 64V</i>
13	<i>pDZ-impE2 64I</i>
14	<i>pDZ-impE2 64L</i>
15	<i>pDZ-impE2 64M</i>
16	<i>pDZ-impE2 64F</i>
17	<i>pDZ-impE2 64Y</i>
18	<i>pDZ-impE2 64W</i>

Ví dụ 5-2: Điều chế các chủng trong đó axit amin tại vị trí của các sản phẩm biến đổi (*ImpE1*, *ImpE2*) được thay thế bằng các axit amin khác và so sánh khả năng sản sinh IMP 54 loại plasmit được điều chế trong ví dụ 5-1 đã được biến nạp vào chủng *CJI0323*. Các chủng trong đó vector được chèn vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn lọc trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp được lựa chọn đã trải qua lai chéo lần thứ hai. Việc đưa vào đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó là phân tích trình tự nucleotit. Tên các chủng theo các biến đổi được chèn được thể hiện trong Bảng 14 bên dưới.

Bảng 14

Số	Chủng
1	<i>CJI0323::impE1(E164R)</i>
2	<i>CJI0323::impE1(E164H)</i>
3	<i>CJI0323::impE1(E164D)</i>
4	<i>CJI0323::impE1(E164S)</i>
5	<i>CJI0323::impE1(E164T)</i>
6	<i>CJI0323::impE1(E164N)</i>
7	<i>CJI0323::impE1(E164Q)</i>
8	<i>CJI0323::impE1(E164C)</i>
9	<i>CJI0323::impE1(E164G)</i>
10	<i>CJI0323::impE1(E164P)</i>
11	<i>CJI0323::impE1(E164A)</i>
12	<i>CJI0323::impE1(E164V)</i>
13	<i>CJI0323::impE1(E164I)</i>
14	<i>CJI0323::impE1(E164L)</i>
15	<i>CJI0323::impE1(E164M)</i>
16	<i>CJI0323::impE1(E164F)</i>
17	<i>CJI0323::impE1(E164Y)</i>
18	<i>CJI0323::impE1(E164W)</i>
19	<i>CJI0323::impE2(V2R)</i>
20	<i>CJI0323::impE2(V2H)</i>
21	<i>CJI0323::impE2(V2K)</i>
22	<i>CJI0323::impE2(V2D)</i>
23	<i>CJI0323::impE2(V2E)</i>
24	<i>CJI0323::impE2(V2S)</i>
25	<i>CJI0323::impE2(V2T)</i>
26	<i>CJI0323::impE2(V2N)</i>
27	<i>CJI0323::impE2(V2Q)</i>
28	<i>CJI0323::impE2(V2C)</i>
29	<i>CJI0323::impE2(V2G)</i>
30	<i>CJI0323::impE2(V2P)</i>
31	<i>CJI0323::impE2(V2A)</i>
32	<i>CJI0323::impE2(V2L)</i>
33	<i>CJI0323::impE2(V2M)</i>
34	<i>CJI0323::impE2(V2F)</i>

35	<i>CJI0323::impE2(V2Y)</i>
36	<i>CJI0323::impE2(V2W)</i>
37	<i>CJI0323::impE2(G64R)</i>
38	<i>CJI0323::impE2(G64H)</i>
39	<i>CJI0323::impE2(G64D)</i>
40	<i>CJI0323::impE2(G64K)</i>
41	<i>CJI0323::impE2(G64S)</i>
42	<i>CJI0323::impE2(G64T)</i>
43	<i>CJI0323::impE2(G64N)</i>
44	<i>CJI0323::impE2(G64Q)</i>
45	<i>CJI0323::impE2(G64C)</i>
46	<i>CJI0323::impE2(G64P)</i>
47	<i>CJI0323::impE2(G64A)</i>
48	<i>CJI0323::impE2(G64V)</i>
49	<i>CJI0323::impE2(G64I)</i>
50	<i>CJI0323::impE2(G64L)</i>
51	<i>CJI0323::impE2(G64M)</i>
52	<i>CJI0323::impE2(G64F)</i>
53	<i>CJI0323::impE2(G64Y)</i>
54	<i>CJI0323::impE2(G64W)</i>

Việc nuôi cấy được thực hiện theo cách tương tự như trong ví dụ 1 và nồng độ của sản phẩm IMP được phân tích (Bảng 15).

Bảng 15

Nồng độ (g/L) sản sinh IMP trong các chủng có sự kết hợp được đưa vào của các biến đổi <i>impE1</i> , <i>impE2</i>		
Chủng		IMP trung bình
Đối chứng	<i>CJI0323_impE1E2(WT)</i>	2,32
1	<i>CJI0323::impE1(E164R)</i>	9,42
2	<i>CJI0323::impE1(E164H)</i>	8,47
3	<i>CJI0323::impE1(E164D)</i>	7,37
4	<i>CJI0323::impE1(E164S)</i>	8,56
5	<i>CJI0323::impE1(E164T)</i>	8,85
6	<i>CJI0323::impE1(E164N)</i>	9,13
7	<i>CJI0323::impE1(E164Q)</i>	7,45
8	<i>CJI0323::impE1(E164C)</i>	7,37

9	<i>CJI0323::impE1(E164G)</i>	9,13
10	<i>CJI0323::impE1(E164P)</i>	9,43
11	<i>CJI0323::impE1(E164A)</i>	7,44
12	<i>CJI0323::impE1(E164V)</i>	8,18
13	<i>CJI0323::impE1(E164I)</i>	8,09
14	<i>CJI0323::impE1(E164L)</i>	7,85
15	<i>CJI0323::impE1(E164M)</i>	7,39
16	<i>CJI0323::impE1(E164F)</i>	7,56
17	<i>CJI0323::impE1(E164Y)</i>	7,60
18	<i>CJI0323::impE1(E164W)</i>	8,56
19	<i>CJI0323::impE2(V2R)</i>	7,99
20	<i>CJI0323::impE2(V2H)</i>	8,75
21	<i>CJI0323::impE2(V2K)</i>	8,66
22	<i>CJI0323::impE2(V2D)</i>	8,28
23	<i>CJI0323::impE2(V2E)</i>	9,32
24	<i>CJI0323::impE2(V2S)</i>	6,37
25	<i>CJI0323::impE2(V2T)</i>	8,37
26	<i>CJI0323::impE2(V2N)</i>	9,80
27	<i>CJI0323::impE2(V2Q)</i>	7,04
28	<i>CJI0323::impE2(V2C)</i>	7,23
29	<i>CJI0323::impE2(V2G)</i>	7,71
30	<i>CJI0323::impE2(V2P)</i>	7,80
31	<i>CJI0323::impE2(V2A)</i>	6,57
32	<i>CJI0323::impE2(V2L)</i>	6,42
33	<i>CJI0323::impE2(V2M)</i>	9,20
34	<i>CJI0323::impE2(V2F)</i>	9,43
35	<i>CJI0323::impE2(V2Y)</i>	8,37
36	<i>CJI0323::impE2(V2W)</i>	7,22
37	<i>CJI0323::impE2(G64R)</i>	4,42
38	<i>CJI0323::impE2(G64H)</i>	5,14
39	<i>CJI0323::impE2(G64D)</i>	11,53
40	<i>CJI0323::impE2(G64K)</i>	4,8
41	<i>CJI0323::impE2(G64S)</i>	5,7
42	<i>CJI0323::impE2(G64T)</i>	5,52
43	<i>CJI0323::impE2(G64N)</i>	5,9
44	<i>CJI0323::impE2(G64Q)</i>	4,8

45	<i>CJI0323::impE2(G64C)</i>	5,9
46	<i>CJI0323::impE2(G64P)</i>	4,75
47	<i>CJI0323::impE2(G64A)</i>	4,58
48	<i>CJI0323::impE2(G64V)</i>	4,56
49	<i>CJI0323::impE2(G64I)</i>	5,89
50	<i>CJI0323::impE2(G64L)</i>	5,6
51	<i>CJI0323::impE2(G64M)</i>	4,3
52	<i>CJI0323::impE2(G64F)</i>	5,89
53	<i>CJI0323::impE2(G64Y)</i>	4,6
54	<i>CJI0323::impE2(G64W)</i>	4,76

Như được thể hiện ở trên, tất cả các chủng biến đổi cho thấy khả năng sản sinh IMP tăng so với từng chủng đối chứng, và do đó, có thể xác nhận rằng ba vị trí biến đổi là những vị trí quan trọng có ảnh hưởng đáng kể đến sự gia tăng của khả năng của protein ImpE liên quan đến sản sinh IMP.

Ví dụ 6: Đưa vào các biến đổi *impE1*, *impE2* dựa trên các chủng sản sinh IMP

Ví dụ 6-1: Điều chế các chủng có biến đổi *impE1*, *impE2* dựa trên các chủng sản sinh IMP

Để xác nhận hiệu quả của việc đưa vào các biến đổi *impE1* và *impE2*, chủng sản sinh IMP đã được điều chế trong đó các hoạt tính của adenylosuccin synthetaza và IMP dehydrogenaza tương ứng với quá trình thoái hóa của IMP trong chủng *ATCC6872* đã bị suy giảm. Các mã khởi đầu đã được thay đổi bằng cách thay đổi bazơ đầu tiên từ “a” thành “t” trong mỗi trình tự nucleotit của hai gen *purA* và *guaB*, mã hóa hai enzym. Chủng trong đó biểu hiện của hai gen bị suy giảm trong chủng *ATCC6872* được định danh là *CJI9088*. Các vector *pDZ-impE1* (*E164K*), *pDZ-impE2* (*V2I*) và *pDZ-impE2* (*G64E*) được điều chế trong ví dụ 4-2 đã được biến nạp vào chủng *CJI9088* bằng cách điện biến nạp, và vector *pDZ-impE2* (*G64D*) được điều chế trong ví dụ 5-1 đã được biến nạp vào chủng *CJI9088_impE1* (*E164K*) *_impE2* (*V2I*) bằng phương pháp điện biến nạp. Các chủng trong đó các vector được đưa vào nhiễm sắc thể bằng cách tái tổ hợp các trình tự tương đồng đã được chọn lọc trên môi trường chứa kanamycin (25 mg/L). Các chủng sơ cấp được lựa chọn đã trải qua lai chéo lần thứ hai. Việc đưa vào các biến đổi trong các chủng biến đổi cuối cùng đã được xác nhận bằng cách thực hiện PCR bằng cặp mồi SEQ ID NO: 15 và SEQ ID NO: 16, sau đó là phân tích trình tự nucleotit.

Khả năng của các chủng đã điều chế (tức là, *CJI9088*, *CJI9088_impE1 (E164K)*, *CJI9088_impE2 (V2I)*, *CJI9088_impE2 (G64E)* và *CJI9088_impE1 (E164K)*. Sau khi hoàn thành nuôi cấy, lượng IMP tổng hợp được đo bằng HPLC và kết quả được thể hiện trong Bảng 16 dưới đây.

Bảng 16

Chủng	IMP (g/L)
<i>CJI9088</i>	0,52
<i>CJI9088_impE1(E164K)</i>	0,84
<i>CJI9088_impE2(V2I)</i>	0,93
<i>CJI9088_impE2(G64E)</i>	1,73
<i>CJI9088_impE1(E164K)_impE2(V2I)(G64D)</i>	4,30

Sau khi xác nhận lượng IMP tích lũy trong môi trường nuôi cấy, có thể xác nhận rằng các chủng này cho thấy lượng IMP tổng hợp được gia tăng ít nhất là 61%, và tăng tối đa 727% so với chủng mẹ *CJ9088*. Theo đó, sự gia tăng số lượng IMP tổng hợp được do biến đổi protein *ImpE* theo sáng chế có thể được hiểu là rất có ý nghĩa.

Từ phần mô tả ở trên, người có hiểu biết trung bình trong lĩnh vực kỹ thuật tương ứng sẽ có thể hiểu rằng sáng chế có thể được thể hiện dưới các hình thức cụ thể khác mà không sửa đổi các nguyên lý kỹ thuật hoặc đặc điểm thiết yếu của sáng chế. Về vấn đề này, các phương án ưu tiên được thực hiện ở đây chỉ nhằm mục đích minh họa và không nên được hiểu là giới hạn phạm vi của sáng chế. Ngược lại, sáng chế không chỉ bao gồm các phương án ưu tiên mà còn bao gồm nhiều phương án sửa đổi, tương đương và các phương án khác có thể được đưa ra mà không tách rời khỏi nguyên lý và phạm vi bảo hộ của sáng chế như được xác định bởi các yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Biến thể protein sản sinh 5'-inosin monophosphat, trong đó axit amin thứ 164 trong trình tự axit amin SEQ ID NO: 1 được thay thế bằng axit amin khác.
2. Biến thể protein theo điểm 1, trong đó axit amin thứ 164 trong trình tự axit amin SEQ ID NO: 1 được thay thế bằng axit amin được lựa chọn từ nhóm bao gồm lysin, arginin, asparagin, glyxin, threonin và prolin.
3. Polynucleotit mã hóa biến thể protein theo điểm 1 hoặc 2.
4. Vector bao gồm polynucleotit theo điểm 3.
5. Vi sinh vật thuộc chi *Corynebacterium* sản sinh 5'-inosin monophosphat, trong đó vi sinh vật chứa biến thể protein theo điểm 1 hoặc 2, hoặc vector bao gồm polynucleotit mã hóa biến thể protein theo điểm 1 hoặc 2.
6. Vi sinh vật theo điểm 5, trong đó vi sinh vật thuộc chi *Corynebacterium* là *Corynebacterium stationis*.
7. Phương pháp điều chế 5'-inosin monophosphat bao gồm nuôi cấy vi sinh vật thuộc chi *Corynebacterium* theo điểm 5 trong môi trường nuôi cấy, và thu hồi 5'-inosin monophosphat từ vi sinh vật hoặc môi trường nuôi cấy.
8. Phương pháp theo điểm 7, trong đó vi sinh vật thuộc chi *Corynebacterium* là *Corynebacterium Stationis*.

DANH MỤC TRÌNH TỰ

<110> CJ CheilJedang Corporation

<120> BIẾN THỂ PROTEIN SẢN SINH 5'-INOSIN MONOPHOSPHAT, POLYNUCLEOTIT, VECTO CHỨA POLYNUCLEOTIT, VI SINH VẬT CHỨA BIẾN THỂ PROTEIN, VÀ PHƯƠNG PHÁP ĐIỀU CHẾ 5'-INOSIN MONOPHOSPHAT

<130> OPA18597

<150> KR10-2017-0173505

<151> 2017-12-15

<160> 154

<170> KoPatentIn 3.0

<210> 1

<211> 222

<212> PRT

<213> Chưa biết

<220>

<223> ImpE1

<400> 1

Leu	His	Ala	Val	Gln	Glu	Val	Asn	Asp	Asn	Glu	Glu	Asp	Ser	Leu	Pro
1				5					10					15	
Gly	Ser	Asp	Leu	Gly	Leu	Arg	Glu	Gln	Lys	Arg	Leu	Ala	Thr	Lys	His
			20					25					30		
Arg	Ile	Glu	Asp	Ala	Ala	Thr	Arg	Leu	Val	Asp	Glu	Ser	Ser	Phe	Asp
		35					40					45			
Lys	Val	Thr	Ile	Glu	Glu	Ile	Cys	Glu	Ala	Ala	Gly	Ile	Ser	Arg	Arg
	50					55					60				
Thr	Phe	Phe	Asn	Tyr	Phe	Ser	Thr	Lys	Glu	Ser	Ala	Val	Ile	Gly	Ala
65					70					75					80
Ser	Ser	Glu	Pro	Leu	Thr	Glu	Lys	Gln	Arg	Asn	Asp	Phe	Leu	Asn	Ala
				85					90					95	
Asp	Ala	Ser	Asn	Leu	Leu	Gln	Leu	Met	Val	Glu	Gln	Ile	Lys	Gln	His
			100					105					110		
Leu	Glu	Ser	Ser	His	Gln	Ser	Gln	Ala	Ile	His	Asp	Arg	Arg	Gln	Arg
		115					120					125			
Ile	Phe	Ala	Asp	Pro	Asp	Val	Ala	Val	Arg	Ala	Met	Ala	Phe	Arg	Lys
	130						135				140				
Glu	Arg	Ser	Arg	Glu	Thr	Met	Glu	Leu	Ile	Ala	Gln	Arg	Leu	Arg	Glu
145					150					155					160

39176

His Pro Glu Glu Gln Arg Ala Pro Glu Leu Asp Pro Glu Thr Glu Ala
165 170 175
Met Leu Leu Ser Gly Phe Ile Arg Glu Ala Thr Trp Met Ala Ile Ser
180 185 190
Arg Pro Asp Arg Asp Cys Ala Leu Pro Val Gly Asp Arg Ile Tyr Arg
195 200 205
Ala Met Glu Leu Val Lys Asn Tyr Thr Lys Gly Leu Glu Trp
210 215 220

<210> 2
<211> 549
<212> PRT
<213> Chưa biết

<220>
<223> ImpE2

<400> 2
Met Val Ala Lys Asn Ser Thr Pro Ser Thr Ala Gly His Ala Ser Ala
1 5 10 15
His Thr Ala Glu Glu Phe Pro Val Ala Asn Ala Glu Met Ala Thr Pro
20 25 30
Ser Ala Ile Asp Pro Asn His Gly Lys Lys Thr Ala Asp Asn Val Gly
35 40 45
Ile Ile Phe Ala Ala Leu Met Leu Thr Met Leu Met Ser Ser Leu Gly
50 55 60
Gln Met Ile Phe Gly Ser Ala Leu Pro Thr Ile Val Gly Glu Leu Gly
65 70 75 80
Gly Val Asp Gln Met Ser Trp Val Ile Ser Ala Phe Met Val Thr Met
85 90 95
Thr Ile Ala Met Pro Leu Ala Gly Gln Leu Gly Asp Arg Met Gly Arg
100 105 110
Lys Trp Val Tyr Ile Ser Gly Ile Ser Ile Phe Val Ile Gly Ser Thr
115 120 125
Leu Gly Gly Phe Ala Asn Gly Met Gly Met Leu Ile Thr Gly Arg Ala
130 135 140
Ile Gln Gly Phe Gly Ala Gly Ile Met Met Ile Ser Ser Gln Ser Ile
145 150 155 160
Val Ala Glu Val Val Ser Ala Arg Glu Arg Gly Lys Phe Met Gly Ile
165 170 175
Met Gly Gly Val Phe Gly Val Ser Ser Val Leu Gly Pro Val Leu Gly
180 185 190

Gly	Trp	Phe	Thr	Asp	Gly	Pro	Gly	Trp	Arg	Trp	Gly	Leu	Trp	Ile	Asn
		195					200					205			
Ile	Pro	Leu	Gly	Leu	Leu	Ala	Ile	Ile	Val	Cys	Ala	Phe	Val	Leu	Lys
	210					215					220				
Leu	Arg	Val	Gly	Glu	Gln	Gly	Phe	Lys	Gly	Phe	Asp	Trp	Met	Gly	Phe
225					230					235					240
Ala	Ala	Ile	Ala	Ile	Thr	Thr	Ser	Thr	Leu	Ile	Leu	Leu	Thr	Thr	Trp
				245					250					255	
Gly	Gly	Ser	Glu	Tyr	Glu	Trp	Thr	Ser	Pro	Thr	Ile	Leu	Ser	Met	Ala
			260					265					270		
Ala	Val	Val	Ile	Val	Gly	Ala	Leu	Leu	Thr	Val	Phe	Ile	Glu	Ser	Arg
		275					280					285			
Ala	Ser	Gln	Pro	Leu	Ile	Pro	Val	Gln	Leu	Phe	Lys	Asn	Arg	Asn	Met
	290					295					300				
Val	Leu	Thr	Thr	Leu	Ala	Gly	Thr	Val	Leu	Gly	Leu	Ala	Met	Met	Gly
305					310					315					320
Val	Leu	Gly	Tyr	Met	Pro	Thr	Tyr	Leu	Gln	Met	Val	His	Thr	Leu	Thr
				325					330					335	
Pro	Thr	Glu	Ala	Gly	Leu	Met	Met	Ile	Pro	Met	Met	Val	Gly	Met	Ile
			340					345					350		
Gly	Val	Ser	Thr	Gly	Val	Gly	Phe	Ile	Ile	Ala	Lys	Thr	Gly	Asn	Tyr
		355					360					365			
Lys	Tyr	Tyr	Pro	Ile	Ala	Gly	Leu	Ala	Ile	Thr	Ala	Phe	Ala	Leu	Trp
	370					375					380				
Trp	Met	Ser	Gln	Met	Thr	Val	Glu	Thr	Ser	Leu	Thr	Gly	Ile	Gly	Val
385					390					395					400
Arg	Phe	Leu	Val	Phe	Gly	Val	Gly	Leu	Gly	Phe	Val	Met	Gln	Val	Leu
				405					410					415	
Val	Leu	Ile	Val	Gln	Asn	Ser	Phe	Pro	Val	Ser	Gln	Val	Gly	Thr	Ala
			420					425					430		
Thr	Ala	Ala	Asn	Asn	Phe	Phe	Arg	Gln	Ile	Gly	Ser	Ala	Leu	Gly	Ala
		435					440					445			
Ser	Ile	Val	Gly	Ser	Met	Phe	Ile	His	Asn	Met	Gln	Asn	Glu	Met	Ala
	450					455					460				
Thr	Arg	Leu	Pro	Asp	Ala	Leu	Ala	Ser	Leu	Gly	Lys	Glu	Gly	Ala	Ala
465					470					475					480
Ile	Ser	Gln	Gln	Phe	Gln	Gly	Ala	Asp	Ala	Ala	Asn	Ser	Leu	Thr	Pro
				485					490					495	

39176

His Ala Val Ala Glu Leu Pro Asp Val Leu Arg Asp Ala Ile Leu Asn
500 505 510

Ser Tyr Asn Asp Gly Leu Thr Pro Val Ile Gly Met Met Val Pro Leu
515 520 525

Ala Ile Val Ala Met Leu Ile Leu Phe Pro Leu Arg Gln Glu Arg Leu
530 535 540

Lys Glu Thr Ile Glu
545

<210> 3
<211> 669
<212> ADN
<213> Chưa biết

<220>
<223> ImpE1

<400> 3
ttgcatgctg tgcaagaagt taatgacaat gaagaagact ccctccctgg cagtgcacctc 60
gggttaaggg agcagaagcg attggcaacc aagcatcgca tcgaagacgc cgcgacacgg 120
ttggttgatg aatcgagctt tgacaaagta acaattgaag aaatttgcgga agccgccggg 180
atttcccgac gcaccttttt taattatttc agcacgaaag aaagcgccgt tattggcgcg 240
tcctcggaac cggtgacgga aaagcaacgc aatgacttct tgaatgctga cgccagcaat 300
ctcctgcagc tgatggttga gcagatcaaa caacacttgg agtctttctca ccagagtcaa 360
gcgattcacg accgtcgtca gcgaatcttt gcggatccgg atgtcgcggg acgtgcaatg 420
gcgtttcgca aggaacgctc acgggaaacc atggagctaa tcgctcaacg tcttcggggag 480
catcctgaag aacaacgcgc cccagaattg gatccggaaa cagaggcgat gctgctgagc 540
ggattcattc gcgaagccac ctggatggct atctcacgac ccgatcgtga ttgtgcactg 600
ccagtgggtg accgcatcta tcgcgcgatg gaattggtaa agaattacac gaaaggactg 660
gaatggtag 669

<210> 4
<211> 1650
<212> ADN
<213> Chưa biết

<220>
<223> ImpE2

<400> 4

```

atggtagcta aaaactccac cccaagcacg gccggccacg ccagtgtca cactgcgga 60
gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggt 120
aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgtcac catgtgatg 180
agctcttttg ggcagatgat tttcggttcc gctctgccaa ccatcgtcgg cgagctcggc 240
ggcgtggacc agatgagctg ggtaatttca gcatttatgg tcacatgac cattgctatg 300
ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctctacat ctcaggtatc 360
tccatthttc ttattggctc gacgctcggg ggctttgcca atggcatggg catgtgatc 420
accggacgtg caatccaggg cttcggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
tttggcgtct cctccgtact ggggccagtt ctcgggtggc ggttcaccga tggccccggt 600
tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgct 660
ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatgggtttt 720
gcggccatcg caatcacgac cagcaccctg attctgtca ccacttgggg cggaagcgaa 780
tacgagtgga cttccccaac tattttgtcc atggctgccg tagtcatcgt cggcgcgctg 840
ctcaccgtgt tcattgagtc gcgtgcatcc cagccgctga tcccggttca gctatttaag 900
aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
gtgctcggct acatgccaac ctacctgcag atgggtgcaca ccctgacgcc aactgaagca 1020
ggcttgatga tgatcccgat gatggtcggc atgatcggtg tctccactgg tgttggttc 1080
atcatcgcta agaccggcaa ctacaagtac taccatcg cgggcctggc catcacggcg 1140
tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcgagtt 1200
cgcttccttg tattcgggtg cggcttgggc tttgtcatgc aggtactggg gctgattgtt 1260
caaaactcct tccctgtatc gcaggtcggg actgccacgg cggctaataa cttcttcgc 1320
cagattgggt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
aatgagatgg ctaccggtt gcctgatgcc cttgcatcgt tgggcaagga aggcgcgct 1440
atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500
gagcttcccg atgtcctccg tgacgtatc ttaaattcct acaatgacgg tctgacccc 1560
gtgattggca tgatggtgcc actggccatt gttgcaatgc tgattttgtt cccactgcgc 1620
caagagcgct tgaaggaaac catcgaataa 1650

```

<211> 29
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 kop-1

<400> 5
 gctctagacg agaaagctaa agccggtga 29

<210> 6
 <211> 36
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 kop-2

<400> 6
 gtttttagct accattgtta caccctgtgc aagttt 36

<210> 7
 <211> 36
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 kop-3

<400> 7
 gcacggggtg taacaatggt agctaaaaac tccacc 36

<210> 8
 <211> 29
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 kop-4

<400> 8
 gctctagaaa tagttgggga agtccactc 29

<210> 9
 <211> 29
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 kop-1

<400> 9
 gctctagact tggatgacct ggtggaaaa 29

<210> 10
 <211> 36
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 kop-2

<400> 10
 cttggagaaa atttcctacc attccagtcc tttcgt 36

<210> 11
 <211> 36
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 kop-3

<400> 11
 ggactggaat ggtaggaaat tttctccaag ggaaat 36

<210> 12
 <211> 29
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 kop-4

<400> 12
 ggactagtgg attgtgttga cgcacgatg 29

<210> 13
 <211> 36
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1E2 kop-2

<400> 13
 cttggagaaa atttctgtta caccctgtgc aagttt 36

<210> 14

<211>	36	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE1E2 kop-3	
<400>	14	
	gcacggggtg taacagaaat tttctccaag ggaaat	36
<210>	15	
<211>	21	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE1E2 seqF	
<400>	15	
	gaacggagtc atctcctttg c	21
<210>	16	
<211>	22	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE1E2 seqR	
<400>	16	
	ccaaacgctc tgcaagaaac tg	22
<210>	17	
<211>	29	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE1E2 WT F	
<400>	17	
	gctctagaga acggagtcac ctcctttgc	29
<210>	18	
<211>	30	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE1E2 WT R	

<400>	18		
		gctctagacc aaacgctctg caagaaactg	30
<210>	19		
<211>	29		
<212>	ADN		
<213>	Tình tự nhân tạo		
<220>			
<223>	impE1 164K-1		
<400>	19		
		gctctagact tggatgacct ggtggaaaa	29
<210>	20		
<211>	37		
<212>	ADN		
<213>	Tình tự nhân tạo		
<220>			
<223>	impE1 164K-2		
<400>	20		
		ctggggcgcg ttgtttttca ggatgctccc gaagacg	37
<210>	21		
<211>	21		
<212>	ADN		
<213>	Tình tự nhân tạo		
<220>			
<223>	impE1 164K-3		
<400>	21		
		aacaacgcgc cccagaattg g	21
<210>	22		
<211>	29		
<212>	ADN		
<213>	Tình tự nhân tạo		
<220>			
<223>	impE1 164K-4		
<400>	22		
		gctctagaaa tagttgggga agtccactc	29

<210> 23
 <211> 37
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 V2I-2

<400> 23
 tggagttttt agctatcatt ccagtccttt cgtgtaa

37

<210> 24
 <211> 21
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 V2I-3

<400> 24
 tagctaaaaa ctccacccca a

21

<210> 25
 <211> 37
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 G64E-2

<400> 25
 ccgaaaatca tctgctccaa agagctcatc agcatgg

37

<210> 26
 <211> 21
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 G64E-3

<400> 26
 gcagatgatt ttcggttccg c

21

<210> 27
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>

<223> Spel-impE1 164 1F

<400> 27

gggactagtg attccggcca actgtcg

27

<210> 28

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-R 1R

<400> 28

tggggcgcggt tggcgttcag gatgctc

27

<210> 29

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-H 1R

<400> 29

tggggcgcggt tgggtgttcag gatgctc

27

<210> 30

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-D 1R

<400> 30

tggggcgcggt tgatcttcag gatgctc

27

<210> 31

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-S 1R

<400> 31

tggggcgcggt tgggattcag gatgctc

27

<210> 32
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-T 1R

<400> 32
 tggggcgcggt tgggttttcag gatgctc

27

<210> 33
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-N 1R

<400> 33
 tggggcgcggt tgggttttcag gatgctc

27

<210> 34
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-Q 1R

<400> 34
 tggggcgcggt tgctgttcag gatgctc

27

<210> 35
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-C 1R

<400> 35
 tggggcgcggt tggcatttcag gatgctc

27

<210> 36
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>

<223> impE1 164-G 1R

<400> 36

tggggcgcggt tggccttcag gatgctc

27

<210> 37

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-P 1R

<400> 37

tggggcgcggt tgcggttcag gatgctc

27

<210> 38

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-A 1R

<400> 38

tggggcgcggt tgggcttcag gatgctc

27

<210> 39

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-V 1R

<400> 39

tggggcgcggt tggacttcag gatgctc

27

<210> 40

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-I 1R

<400> 40

tggggcgcggt tggatttcag gatgctc

27

<210> 41
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-L 1R

<400> 41
 tggggcgcggt tgcagttcag gatgctc

27

<210> 42
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-M 1R

<400> 42
 tggggcgcggt tgcatttcag gatgctc

27

<210> 43
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-F 1R

<400> 43
 tggggcgcggt tggaattcag gatgctc

27

<210> 44
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-Y 1R

<400> 44
 tggggcgcggt tggatttcag gatgctc

27

<210> 45
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-W 1R

<400> 45
 tggggcgcggt tgccattcag gatgctc

27

<210> 46
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> Spe1-impE1 164 2R

<400> 46
 gggactagtc atgaacttgc cgcgctc

27

<210> 47
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-R 2F

<400> 47
 gagcatcctg aacgccaacg cgcccca

27

<210> 48
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-H 2F

<400> 48
 gagcatcctg aacaccaacg cgcccca

27

<210> 49
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-D 2F

<400> 49

gagcatcctg aagatcaacg cgcccca

27

<210> 50
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-S 2F

<400> 50
 gagcatcctg aatcccaacg cgcccca

27

<210> 51
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-T 2F

<400> 51
 gagcatcctg aaaccaacg cgcccca

27

<210> 52
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-N 2F

<400> 52
 gagcatcctg aaaaccaacg cgcccca

27

<210> 53
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-Q 2F

<400> 53
 gagcatcctg aacagcaacg cgcccca

27

<210> 54
 <211> 27
 <212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-C 2F

<400> 54

gagcatcctg aatgcccaacg cgcccca

27

<210> 55

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-G 2F

<400> 55

gagcatcctg aaggccaacg cgcccca

27

<210> 56

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-P 2F

<400> 56

gagcatcctg aaccgcaacg cgcccca

27

<210> 57

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-A 2F

<400> 57

gagcatcctg aagcccaacg cgcccca

27

<210> 58

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE1 164-V 2F

<400> 58
 gagcatcctg aagtccaacg cgcccca 27

<210> 59
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-I 2F

<400> 59
 gagcatcctg aaatccaacg cgcccca 27

<210> 60
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-L 2F

<400> 60
 gagcatcctg aactgcaacg cgcccca 27

<210> 61
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-M 2F

<400> 61
 gagcatcctg aaatgcaacg cgcccca 27

<210> 62
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-F 2F

<400> 62
 gagcatcctg aattccaacg cgcccca 27

<210> 63
 <211> 27

<212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-Y 2F

<400> 63
 gagcatcctg aataccaacg cgcccca

27

<210> 64
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE1 164-W 2F

<400> 64
 gagcatcctg aatggcaacg cgcccca

27

<210> 65
 <211> 26
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> XbaI-impE2 2 1F

<400> 65
 gggctctagat tgcattgctgt gcaaga

26

<210> 66
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-R 1R

<400> 66
 ggagttttta gcgcgcattc cagtcct

27

<210> 67
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-H 1R

<400>	67	
ggagtttttta	gcgtgcattc	cagtcct
		27
<210>	68	
<211>	27	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE2 2-K	1R
<400>	68	
ggagtttttta	gccttcattc	cagtcct
		27
<210>	69	
<211>	27	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE2 2-D	1R
<400>	69	
ggagtttttta	gcgtccattc	cagtcct
		27
<210>	70	
<211>	27	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE2 2-E	1R
<400>	70	
ggagtttttta	gcttccattc	cagtcct
		27
<210>	71	
<211>	27	
<212>	ADN	
<213>	Tình tự nhân tạo	
<220>		
<223>	impE2 2-S	1R
<400>	71	
ggagtttttta	gcggacattc	cagtcct
		27
<210>	72	

<211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-T 1R

<400> 72
 ggagtttttta gcggtcattc cagtcct

27

<210> 73
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-N 1R

<400> 73
 ggagtttttta gcgttcattc cagtcct

27

<210> 74
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-Q 1R

<400> 74
 ggagtttttta gcctgcattc cagtcct

27

<210> 75
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-C 1R

<400> 75
 ggagtttttta gcgcacattc cagtcct

27

<210> 76
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-G 1R

<400> 76
 ggagttttta gcgccattc cagtcct 27

<210> 77
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo
 <220>
 <223> impE2 2-P 1R

<400> 77
 ggagttttta gctggcattc cagtcct 27

<210> 78
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo
 <220>
 <223> impE2 2-A 1R

<400> 78
 ggagttttta gcagccattc cagtcct 27

<210> 79
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo
 <220>
 <223> impE2 2-L 1R

<400> 79
 ggagttttta gccagcattc cagtcct 27

<210> 80
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo
 <220>
 <223> impE2 2-M 1R

<400> 80
 ggagttttta gccatcattc cagtcct 27

<210> 81
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-F 1R

<400> 81
 ggagttttta gcgaacattc cagtcct

27

<210> 82
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-Y 1R

<400> 82
 ggagttttta gcgtacattc cagtcct

27

<210> 83
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-W 1R

<400> 83
 ggagttttta gcccacattc cagtcct

27

<210> 84
 <211> 26
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> XbaI-impE2 2 2R

<400> 84
 ggggtctagat tgctcgccca cgcgca

26

<210> 85
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>

<223> impE2 2-R 2F

<400> 85
aggactggaa tgcgcgctaa aaactcc

27

<210> 86
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 2-H 2F

<400> 86
aggactggaa tgcacgctaa aaactcc

27

<210> 87
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 2-K 2F

<400> 87
aggactggaa tgaaggctaa aaactcc

27

<210> 88
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 2-D 2F

<400> 88
aggactggaa tggacgctaa aaactcc

27

<210> 89
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 2-E 2F

<400> 89
aggactggaa tggaagctaa aaactcc

27

<210> 90
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-S 2F

<400> 90
 aggactggaa tgtccgctaa aaactcc

27

<210> 91
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-T 2F

<400> 91
 aggactggaa tgaccgctaa aaactcc

27

<210> 92
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-N 2F

<400> 92
 aggactggaa tgaacgctaa aaactcc

27

<210> 93
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-Q 2F

<400> 93
 aggactggaa tgcaggctaa aaactcc

27

<210> 94
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-C 2F

<400> 94
 aggactggaa tgtgcgctaa aaactcc 27

<210> 95
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-G 2F

<400> 95
 aggactggaa tgggcgctaa aaactcc 27

<210> 96
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-P 2F

<400> 96
 aggactggaa tgccagctaa aaactcc 27

<210> 97
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-A 2F

<400> 97
 aggactggaa tggctgctaa aaactcc 27

<210> 98
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-L 2F

<400> 98
 aggactggaa tgctggctaa aaactcc 27

<210> 99
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-M 2F

<400> 99
 aggactggaa tgatggctaa aaactcc

27

<210> 100
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-F 2F

<400> 100
 aggactggaa tgttcgctaa aaactcc

27

<210> 101
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-Y 2F

<400> 101
 aggactggaa tgtacgctaa aaactcc

27

<210> 102
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 2-W 2F

<400> 102
 aggactggaa tgtgggctaa aaactcc

27

<210> 103
 <211> 30
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> XbaI-impE2 64 1F

<400> 103
 gggctctagaa aagagcttaa ggcagctgct 30

<210> 104
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-R 1R

<400> 104
 gaaaatcatc tggcgcaaag agctcat 27

<210> 105
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-H 1R

<400> 105
 gaaaatcatc tgggtgcaaag agctcat 27

<210> 106
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-D 1R

<400> 106
 gaaaatcatc tgggtccaaag agctcat 27

<210> 107
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-K 1R

<400> 107

gaaaatcatc tgcttcaaag agctcat

27

<210> 108
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-S 1R

<400> 108
 gaaaatcatc tgggacaaag agctcat

27

<210> 109
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-T 1R

<400> 109
 gaaaatcatc tgggtcaaag agctcat

27

<210> 110
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-N 1R

<400> 110
 gaaaatcatc tgggtcaaag agctcat

27

<210> 111
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-Q 1R

<400> 111
 gaaaatcatc tgctgcaaag agctcat

27

<210> 112
 <211> 27
 <212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE2 64-C 1R

<400> 112

gaaaatcatc tggcacaaag agctcat

27

<210> 113

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE2 64-P 1R

<400> 113

gaaaatcatc tgtggcaaag agctcat

27

<210> 114

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE2 64-A 1R

<400> 114

gaaaatcatc tgagccaaag agctcat

27

<210> 115

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE2 64-V 1R

<400> 115

gaaaatcatc tggaccaaag agctcat

27

<210> 116

<211> 27

<212> ADN

<213> Tình tự nhân tạo

<220>

<223> impE2 64-I 1R

<400> 116
 gaaaatcatc tggatcaaag agctcat 27

<210> 117
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-L 1R

<400> 117
 gaaaatcatc tgcagcaaag agctcat 27

<210> 118
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-M 1R

<400> 118
 gaaaatcatc tgcatacaaag agctcat 27

<210> 119
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-F 1R

<400> 119
 gaaaatcatc tggaacaaaag agctcat 27

<210> 120
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-Y 1R

<400> 120
 gaaaatcatc tggatacaaag agctcat 27

<210> 121
 <211> 27

<212> ADN
 <213> Tình tự nhân tạo

 <220>
 <223> impE2 64-W 1R

 <400> 121
 gaaaatcatc tgccacaaag agctcat 27

 <210> 122
 <211> 30
 <212> ADN
 <213> Tình tự nhân tạo

 <220>
 <223> XbaI-impE2 64 2R

 <400> 122
 gggctctagac ggtcaatgaa gtctcaacgg 30

 <210> 123
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

 <220>
 <223> impE2 64-R 2F

 <400> 123
 atgagctctt tgcgccagat gatttttc 27

 <210> 124
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

 <220>
 <223> impE2 64-H 2F

 <400> 124
 atgagctctt tgcaccagat gatttttc 27

 <210> 125
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

 <220>
 <223> impE2 64-D 2F

<400> 125
 atgagctctt tggaccagat gatttttc 27

<210> 126
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-K 2F

<400> 126
 atgagctctt tgaagcagat gatttttc 27

<210> 127
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-S 2F

<400> 127
 atgagctctt tgtcccagat gatttttc 27

<210> 128
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-T 2F

<400> 128
 atgagctctt tgaccagat gatttttc 27

<210> 129
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-N 2F

<400> 129
 atgagctctt tgaaccagat gatttttc 27

<210> 130

<211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-Q 2F

<400> 130
 atgagctctt tgcagcagat gatttttc

27

<210> 131
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-C 2F

<400> 131
 atgagctctt tgtgccagat gatttttc

27

<210> 132
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-P 2F

<400> 132
 atgagctctt tgccacagat gatttttc

27

<210> 133
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-A 2F

<400> 133
 atgagctctt tggctcagat gatttttc

27

<210> 134
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-V 2F

<400> 134
atgagctctt tgggtccagat gatttttc

27

<210> 135
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 64-I 2F

<400> 135
atgagctctt tgatccagat gatttttc

27

<210> 136
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 64-L 2F

<400> 136
atgagctctt tgctgcagat gatttttc

27

<210> 137
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 64-M 2F

<400> 137
atgagctctt tgatgcagat gatttttc

27

<210> 138
<211> 27
<212> ADN
<213> Tình tự nhân tạo

<220>
<223> impE2 64-F 2F

<400> 138
atgagctctt tgggtccagat gatttttc

27

<210> 139
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-Y 2F

<400> 139
 atgagctctt tgtaccagat gatttttc

27

<210> 140
 <211> 27
 <212> ADN
 <213> Tình tự nhân tạo

<220>
 <223> impE2 64-W 2F

<400> 140
 atgagctctt tgtggcagat gatttttc

27

<210> 141
 <211> 221
 <212> PRT
 <213> Chưa biết

<220>
 <223> ImpE1-CJI0323

<400> 141
 Leu His Ala Val Gln Glu Val Asn Asp Asn Glu Glu Asp Ser Leu Pro
 1 5 10 15
 Gly Ser Asp Leu Gly Leu Arg Glu Gln Lys Arg Leu Ala Thr Lys His
 20 25 30
 Arg Ile Glu Asp Ala Ala Thr Arg Leu Val Asp Glu Ser Ser Phe Asp
 35 40 45
 Lys Val Thr Ile Glu Glu Ile Cys Glu Ala Ala Gly Ile Ser Arg Arg
 50 55 60
 Thr Phe Phe Asn Tyr Phe Ser Thr Lys Glu Ser Ala Val Ile Gly Ala
 65 70 75 80
 Ser Ser Glu Pro Leu Thr Glu Lys Gln Arg Asn Asp Phe Leu Asn Ala
 85 90 95
 Asp Ala Ser Asn Leu Leu Gln Leu Met Val Glu Gln Ile Lys Gln His
 100 105 110
 Leu Glu Ser Ser His Gln Ser Gln Ala Ile His Asp Arg Arg Gln Arg

39176

115	120	125
Ile Phe Ala Asp Pro Asp Val Ala Val Arg Ala Met Ala Phe Arg Lys		
130	135	140
Glu Arg Ser Arg Glu Thr Met Glu Leu Ile Ala Gln Arg Leu Arg Glu		
145	150	155
His Pro Glu Lys Gln Arg Ala Pro Glu Leu Asp Pro Glu Thr Glu Ala		
165	170	175
Met Leu Leu Ser Gly Phe Ile Arg Glu Ala Thr Trp Met Ala Ile Ser		
180	185	190
Arg Pro Asp Arg Asp Cys Ala Leu Pro Val Gly Asp Arg Ile Tyr Arg		
195	200	205
Ala Met Glu Leu Val Lys Asn Tyr Thr Lys Gly Leu Glu		
210	215	220

<210> 142
 <211> 549
 <212> PRT
 <213> Chưa biết

<220>
 <223> ImpE2-CJI0323

<400> 142

Met Ile Ala Lys Asn Ser Thr Pro Ser Thr Ala Gly His Ala Ser Ala
1 5 10 15
His Thr Ala Glu Glu Phe Pro Val Ala Asn Ala Glu Met Ala Thr Pro
20 25 30
Ser Ala Ile Asp Pro Asn His Gly Lys Lys Thr Ala Asp Asn Val Gly
35 40 45
Ile Ile Phe Ala Ala Leu Met Leu Thr Met Leu Met Ser Ser Leu Glu
50 55 60
Gln Met Ile Phe Gly Ser Ala Leu Pro Thr Ile Val Gly Glu Leu Gly
65 70 75 80
Gly Val Asp Gln Met Ser Trp Val Ile Ser Ala Phe Met Val Thr Met
85 90 95
Thr Ile Ala Met Pro Leu Ala Gly Gln Leu Gly Asp Arg Met Gly Arg
100 105 110
Lys Trp Val Tyr Ile Ser Gly Ile Ser Ile Phe Val Ile Gly Ser Thr
115 120 125
Leu Gly Gly Phe Ala Asn Gly Met Gly Met Leu Ile Thr Gly Arg Ala
130 135 140

Ile	Gln	Gly	Phe	Gly	Ala	Gly	Ile	Met	Met	Ile	Ser	Ser	Gln	Ser	Ile	145	150	155	160
Val	Ala	Glu	Val	Val	Ser	Ala	Arg	Glu	Arg	Gly	Lys	Phe	Met	Gly	Ile	165	170	175	
Met	Gly	Gly	Val	Phe	Gly	Val	Ser	Ser	Val	Leu	Gly	Pro	Val	Leu	Gly	180	185	190	
Gly	Trp	Phe	Thr	Asp	Gly	Pro	Gly	Trp	Arg	Trp	Gly	Leu	Trp	Ile	Asn	195	200	205	
Ile	Pro	Leu	Gly	Leu	Leu	Ala	Ile	Ile	Val	Cys	Ala	Phe	Val	Leu	Lys	210	215	220	
Leu	Arg	Val	Gly	Glu	Gln	Gly	Phe	Lys	Gly	Phe	Asp	Trp	Met	Gly	Phe	225	230	235	240
Ala	Ala	Ile	Ala	Ile	Thr	Thr	Ser	Thr	Leu	Ile	Leu	Leu	Thr	Thr	Trp	245	250	255	
Gly	Gly	Ser	Glu	Tyr	Glu	Trp	Thr	Ser	Pro	Thr	Ile	Leu	Ser	Met	Ala	260	265	270	
Ala	Val	Val	Ile	Val	Gly	Ala	Leu	Leu	Thr	Val	Phe	Ile	Glu	Ser	Arg	275	280	285	
Ala	Ser	Gln	Pro	Leu	Ile	Pro	Val	Gln	Leu	Phe	Lys	Asn	Arg	Asn	Met	290	295	300	
Val	Leu	Thr	Thr	Leu	Ala	Gly	Thr	Val	Leu	Gly	Leu	Ala	Met	Met	Gly	305	310	315	320
Val	Leu	Gly	Tyr	Met	Pro	Thr	Tyr	Leu	Gln	Met	Val	His	Thr	Leu	Thr	325	330	335	
Pro	Thr	Glu	Ala	Gly	Leu	Met	Met	Ile	Pro	Met	Met	Val	Gly	Met	Ile	340	345	350	
Gly	Val	Ser	Thr	Gly	Val	Gly	Phe	Ile	Ile	Ala	Lys	Thr	Gly	Asn	Tyr	355	360	365	
Lys	Tyr	Tyr	Pro	Ile	Ala	Gly	Leu	Ala	Ile	Thr	Ala	Phe	Ala	Leu	Trp	370	375	380	
Trp	Met	Ser	Gln	Met	Thr	Val	Glu	Thr	Ser	Leu	Thr	Gly	Ile	Gly	Val	385	390	395	400
Arg	Phe	Leu	Val	Phe	Gly	Val	Gly	Leu	Gly	Phe	Val	Met	Gln	Val	Leu	405	410	415	
Val	Leu	Ile	Val	Gln	Asn	Ser	Phe	Pro	Val	Ser	Gln	Val	Gly	Thr	Ala	420	425	430	
Thr	Ala	Ala	Asn	Asn	Phe	Phe	Arg	Gln	Ile	Gly	Ser	Ala	Leu	Gly	Ala	435	440	445	
Ser	Ile	Val	Gly	Ser	Met	Phe	Ile	His	Asn	Met	Gln	Asn	Glu	Met	Ala				

450		455		460
Thr Arg Leu Pro Asp Ala Leu Ala Ser Leu Gly Lys Glu Gly Ala Ala				
465		470		475 480
Ile Ser Gln Gln Phe Gln Gly Ala Asp Ala Ala Asn Ser Leu Thr Pro				
	485		490	495
His Ala Val Ala Glu Leu Pro Asp Val Leu Arg Asp Ala Ile Leu Asn				
	500		505	510
Ser Tyr Asn Asp Gly Leu Thr Pro Val Ile Gly Met Met Val Pro Leu				
	515		520	525
Ala Ile Val Ala Met Leu Ile Leu Phe Pro Leu Arg Gln Glu Arg Leu				
	530		535	540
Lys Glu Thr Ile Glu				
545				

<210> 143
 <211> 669
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE1 NT - CJI0323

<400> 143
 ttgcatgctg tgcaagaagt taatgacaat gaagaagact ccctccctgg cagtgcctc 60
 gggttaaggg agcagaagcg attggcaacc aagcatcgca tcgaagacgc cgcgacacgg 120
 ttggttgatg aatcgagctt tgacaaagta acaattgaag aaatttgcca agccgccggg 180
 atttcccgac gcaccttttt taattatttc agcacgaaag aaagcgccgt tattggcgcg 240
 tcctcggaac cgttgacgga aaagcaacgc aatgacttct tgaatgctga cgccagcaat 300
 ctctgcagc tgatggttga gcagatcaaa caacacttgg agtcttctca ccagagtcaa 360
 gcgattcacg accgtcgtca gcgaatcttt gcggatccgg atgtcgcggt acgtgcaatg 420
 gcgtttcgca aggaacgctc acgggaaacc atggagctaa tcgctcaacg tcttcggggag 480
 catcctgaaa aacaacgcgc cccagaattg gatccggaaa cagaggcgat gctgctgagc 540
 ggattcattc gcgaagccac ctggatggct atctcacgac ccgatcgtga ttgtgcactg 600
 ccagtgggtg accgcatcta tcgcgcgatg gaattggtaa agaattacac gaaaggactg 660
 gaatgatag 669

<210> 144
 <211> 1650

<212> ADN
 <213> Chưa biết

<220>
 <223> ImpE2 NT - CJI0323

<400> 144
 atgatagcta aaaactccac cccaagcacg gccggccacg ccagtgtca cactgcggaa 60
 gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggt 120
 aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgtcac catgctgatg 180
 agctcttttg agcagatgat tttcggttcc gctctgcaa ccatcgtcgg cgagctcggc 240
 ggcgtggacc agatgagctg ggtaatttca gcatttatgg tcacatgac cattgctatg 300
 ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctacat ctcaggtatc 360
 tccattttcg ttattggctc gacgctcggg ggctttgcca atggcatggg catgctgatc 420
 accggacgtg caatccaggg cttcggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
 gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
 tttggcgtct cctccgtact gggccagtt ctcgggtggc ggttcaccga tgggtccggg 600
 tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgct 660
 ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatgggtttt 720
 gcggccatcg caatcacgac cagcaccctg attctgtca ccacttgggg cggaagcgaa 780
 tacgagtgga cttccccaac tattttgtcc atggctgccg tagtcatcgt cggcgcgctg 840
 ctcaccgtgt tcattgagtc gctgtcatcc cagccgctga tcccggttca gctatttaag 900
 aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
 gtgctcggct acatgccaac ctacctgcag atggtgcaca ccctgacgcc aactgaagca 1020
 ggcttgatga tgatcccgat gatggctggc atgatcgggt tctccactgg tgttggttc 1080
 atcatcgcta agaccggcaa ctacaagtac taccatcg cgggcctggc catcacggcg 1140
 tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcgagtt 1200
 cgcttccttg tattcgggtg cggcttgggc tttgtcatgc aggtactggg gctgattgtt 1260
 caaaactcct tccctgtatc gcaggctggg actgccacgg cggctaataa cttcttccgc 1320
 cagattgggt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
 aatgagatgg ctaccggttt gcctgatgcc cttgcatcgt tgggcaagga aggcgccgct 1440
 atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500

gagcttcccg atgtcctccg tgacgctatc ttaaattcct acaatgacgg tctgaccccc 1560
 gtgattggca tgatggtgcc actggccatt gttgcaatgc tgattttgtt cccactgcgc 1620
 caagagcgct tgaaggaaac catcgaataa 1650

<210> 145
 <211> 222
 <212> PRT
 <213> Chưa biết

<220>
 <223> ImpE1-164K

<400> 145
 Leu His Ala Val Gln Glu Val Asn Asp Asn Glu Glu Asp Ser Leu Pro
 1 5 10 15
 Gly Ser Asp Leu Gly Leu Arg Glu Gln Lys Arg Leu Ala Thr Lys His
 20 25 30
 Arg Ile Glu Asp Ala Ala Thr Arg Leu Val Asp Glu Ser Ser Phe Asp
 35 40 45
 Lys Val Thr Ile Glu Glu Ile Cys Glu Ala Ala Gly Ile Ser Arg Arg
 50 55 60
 Thr Phe Phe Asn Tyr Phe Ser Thr Lys Glu Ser Ala Val Ile Gly Ala
 65 70 75 80
 Ser Ser Glu Pro Leu Thr Glu Lys Gln Arg Asn Asp Phe Leu Asn Ala
 85 90 95
 Asp Ala Ser Asn Leu Leu Gln Leu Met Val Glu Gln Ile Lys Gln His
 100 105 110
 Leu Glu Ser Ser His Gln Ser Gln Ala Ile His Asp Arg Arg Gln Arg
 115 120 125
 Ile Phe Ala Asp Pro Asp Val Ala Val Arg Ala Met Ala Phe Arg Lys
 130 135 140
 Glu Arg Ser Arg Glu Thr Met Glu Leu Ile Ala Gln Arg Leu Arg Glu
 145 150 155 160
 His Pro Glu Lys Gln Arg Ala Pro Glu Leu Asp Pro Glu Thr Glu Ala
 165 170 175
 Met Leu Leu Ser Gly Phe Ile Arg Glu Ala Thr Trp Met Ala Ile Ser
 180 185 190
 Arg Pro Asp Arg Asp Cys Ala Leu Pro Val Gly Asp Arg Ile Tyr Arg
 195 200 205
 Ala Met Glu Leu Val Lys Asn Tyr Thr Lys Gly Leu Glu Trp
 210 215 220

<210> 146
 <211> 669
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE1 NT-164K

<400> 146
 ttgcatgctg tgcaagaagt taatgacaat gaagaagact ccctccctgg cagtgcacctc 60
 ggggttaaggg agcagaagcg attggcaacc aagcatcgca tcgaagacgc cgcgacacgg 120
 ttggttgatg aatcgagctt tgacaaagta acaattgaag aaatttgca agccgccggg 180
 atttcccgac gcaccttttt taattatttc agcacgaaag aaagcgccgt tattggcgcg 240
 tcctcggaac cgttgacgga aaagcaacgc aatgacttct tgaatgctga cgccagcaat 300
 ctctgcagc tgatggttga gcagatcaaa caacacttgg agtcttctca ccagagtcaa 360
 gcgattcacg accgtcgtca gcgaatcttt gcggatccgg atgtcgcggg acgtgcaatg 420
 gcgttttcgca aggaacgctc acgggaaacc atggagctaa tcgctcaacg tcttcgggag 480
 catcctgaaa aacaacgcgc cccagaattg gatccgaaa cagaggcgat gctgctgagc 540
 ggattcattc gcgaagccac ctggatggct atctcacgac ccgatcgtga ttgtgactg 600
 ccagtgggtg accgcatcta tcgcgcgatg gaattggtaa agaattacac gaaaggactg 660
 gaatggtag 669

<210> 147
 <211> 549
 <212> PRT
 <213> Chưa biết

<220>
 <223> ImpE2 - V2I

<400> 147
 Met Ile Ala Lys Asn Ser Thr Pro Ser Thr Ala Gly His Ala Ser Ala
 1 5 10 15
 His Thr Ala Glu Glu Phe Pro Val Ala Asn Ala Glu Met Ala Thr Pro
 20 25 30
 Ser Ala Ile Asp Pro Asn His Gly Lys Lys Thr Ala Asp Asn Val Gly
 35 40 45

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

355 360 365
 Lys Tyr Tyr Pro Ile Ala Gly Leu Ala Ile Thr Ala Phe Ala Leu Trp
 370 375 380
 Trp Met Ser Gln Met Thr Val Glu Thr Ser Leu Thr Gly Ile Gly Val
 385 390 395 400
 Arg Phe Leu Val Phe Gly Val Gly Leu Gly Phe Val Met Gln Val Leu
 405 410 415
 Val Leu Ile Val Gln Asn Ser Phe Pro Val Ser Gln Val Gly Thr Ala
 420 425 430
 Thr Ala Ala Asn Asn Phe Phe Arg Gln Ile Gly Ser Ala Leu Gly Ala
 435 440 445
 Ser Ile Val Gly Ser Met Phe Ile His Asn Met Gln Asn Glu Met Ala
 450 455 460
 Thr Arg Leu Pro Asp Ala Leu Ala Ser Leu Gly Lys Glu Gly Ala Ala
 465 470 475 480
 Ile Ser Gln Gln Phe Gln Gly Ala Asp Ala Ala Asn Ser Leu Thr Pro
 485 490 495
 His Ala Val Ala Glu Leu Pro Asp Val Leu Arg Asp Ala Ile Leu Asn
 500 505 510
 Ser Tyr Asn Asp Gly Leu Thr Pro Val Ile Gly Met Met Val Pro Leu
 515 520 525
 Ala Ile Val Ala Met Leu Ile Leu Phe Pro Leu Arg Gln Glu Arg Leu
 530 535 540
 Lys Glu Thr Ile Glu
 545

<210> 148
 <211> 1650
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE2 NT - V2I

<400> 148
 atgatagcta aaaactccac cccaagcacg gccggccacg ccagtgtca cactgcggaa 60
 gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggt 120
 aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgctcac catgctgatg 180
 agctctttgg ggcagatgat tttcggttcc gctctgccaa ccatcgtcgg cgagctcggc 240
 ggcgtggacc agatgagctg ggtaatttca gcatttatgg tcaccatgac cattgctatg 300

```

ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctctacat ctcaggtatc 360
tccatttttcg ttattggctc gacgctcggg ggctttgcca atggcatggg catgctgac 420
accggacgtg caatccaggg cttcgggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
tttggcgtct cctccgtact ggggccagtt ctcgggtggc ggttcaccga tgggtccggg 600
tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgct 660
ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatgggtttt 720
gcggccatcg caatcacgac cagcaccctg attctgctca ccacttgggg cggaagcgaa 780
tacgagtgga cttccccaac tattttgtcc atggctgccg tagtcatcgt cggcgcgctg 840
ctcaccgtgt tcattgagtc gcgtgcatcc cagccgctga tcccggttca gctatttaag 900
aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
gtgctcggct acatgccaac ctacctgcag atgggtgcaca cctgacgcc aactgaagca 1020
ggcttgatga tgatcccgat gatggtcggc atgatcggtg tctccactgg tgttggcttc 1080
atcatcgcta agaccggcaa ctacaagtac taccatcg cgggcctggc catcacggcg 1140
tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcggagtt 1200
cgcttccttg tattcgggtgt cggcttgggc tttgtcatgc aggtactggg gctgattgtt 1260
caaaactcct tccctgtatc gcaggtcggg actgccacgg cggctaataa cttcttccgc 1320
cagattgggt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
aatgagatgg ctaccggtt gcctgatgcc cttgcatcgt tgggcaagga aggcgccgct 1440
atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500
gagcttcccg atgtcctccg tgacgctatc ttaaattcct acaatgacgg tctgaccccc 1560
gtgattggca tgatggtgcc actggccatt gttgcaatgc tgattttgtt cccactgcgc 1620
caagagcgct tgaaggaaac catcgaataa 1650

```

```

<210>      149
<211>      549
<212>      PRT
<213>      Chưa biết

<220>
<223>      ImpE2 - G64E

<400>      149

```

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

39176

305		310		315		320									
Val	Leu	Gly	Tyr	Met	Pro	Thr	Tyr	Leu	Gln	Met	Val	His	Thr	Leu	Thr
				325					330					335	
Pro	Thr	Glu	Ala	Gly	Leu	Met	Met	Ile	Pro	Met	Met	Val	Gly	Met	Ile
			340					345					350		
Gly	Val	Ser	Thr	Gly	Val	Gly	Phe	Ile	Ile	Ala	Lys	Thr	Gly	Asn	Tyr
		355					360					365			
Lys	Tyr	Tyr	Pro	Ile	Ala	Gly	Leu	Ala	Ile	Thr	Ala	Phe	Ala	Leu	Trp
	370					375					380				
Trp	Met	Ser	Gln	Met	Thr	Val	Glu	Thr	Ser	Leu	Thr	Gly	Ile	Gly	Val
385					390					395					400
Arg	Phe	Leu	Val	Phe	Gly	Val	Gly	Leu	Gly	Phe	Val	Met	Gln	Val	Leu
				405					410					415	
Val	Leu	Ile	Val	Gln	Asn	Ser	Phe	Pro	Val	Ser	Gln	Val	Gly	Thr	Ala
			420					425					430		
Thr	Ala	Ala	Asn	Asn	Phe	Phe	Arg	Gln	Ile	Gly	Ser	Ala	Leu	Gly	Ala
		435					440					445			
Ser	Ile	Val	Gly	Ser	Met	Phe	Ile	His	Asn	Met	Gln	Asn	Glu	Met	Ala
	450					455					460				
Thr	Arg	Leu	Pro	Asp	Ala	Leu	Ala	Ser	Leu	Gly	Lys	Glu	Gly	Ala	Ala
465					470					475					480
Ile	Ser	Gln	Gln	Phe	Gln	Gly	Ala	Asp	Ala	Ala	Asn	Ser	Leu	Thr	Pro
				485					490					495	
His	Ala	Val	Ala	Glu	Leu	Pro	Asp	Val	Leu	Arg	Asp	Ala	Ile	Leu	Asn
			500					505					510		
Ser	Tyr	Asn	Asp	Gly	Leu	Thr	Pro	Val	Ile	Gly	Met	Met	Val	Pro	Leu
		515					520					525			
Ala	Ile	Val	Ala	Met	Leu	Ile	Leu	Phe	Pro	Leu	Arg	Gln	Glu	Arg	Leu
	530					535					540				
Lys	Glu	Thr	Ile	Glu											
545															

<210> 150
 <211> 1650
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE2 NT - G64E

<400> 150

atggtagcta aaaactccac cccaagcacg gccgggccacg ccagtgetca cactgcggaa 60
 gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggt 120
 aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgctcac catgctgatg 180
 agctcttttg agcagatgat tttcggttcc gctctgccaa ccatcgctcg cgagctcggc 240
 ggctgtggacc agatgagctg ggtaatttca gcatttatgg tcaccatgac cattgctatg 300
 ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctctacat ctgaggtatc 360
 tccattttcg ttattggctc gacgctcggg ggctttgccca atggcatggg catgctgata 420
 accggacgtg caatccaggg cttcgggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
 gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
 tttggcgtct cctccgtact ggggtccagtt ctcggtggct gggttcaccga tgggtcccgg 600
 tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgtc 660
 ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatgggtttt 720
 gcggccatcg caatcacgac cagcaccctg attctgctca ccacttgggg cggaagcgaa 780
 tacgagtggg cttccccaac tattttgtcc atggctgccg tagtcatcgt cggcgcgctg 840
 ctaccgtgt tcattgagtc gcgtgcatcc cagccgctga tcccggttca gctatttaag 900
 aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
 gtgctcggct acatgccaac ctacctgcag atgggtgcaca ccctgacgcc aactgaagca 1020
 ggcttgatga tgatcccgat gatggctggc atgatcgggtg tctccactgg tgttggttc 1080
 atcatcgcta agaccggcaa ctacaagtac taccatcg cggtcctggc catcacggcg 1140
 tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcgagatt 1200
 cgcttccttg tattcgggtgt cggcttgggc tttgtcatgc aggtactggg gctgattgtt 1260
 caaaactcct tccctgtatc gcaggtcggg actgccacgg cggctaataa cttcttccgc 1320
 cagattgggt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
 aatgagatgg ctaccggtt gcctgatgcc cttgcatcgt tgggcaagga aggcgccgct 1440
 atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500
 gagcttcccg atgtcctccg tgacgctatc ttaaattcct acaatgacgg tctgaccccc 1560
 gtgattggca tgatgggtgcc actggccatt gttgcaatgc tgattttgtt cccactgcgc 1620
 caagagcgct tgaaggaaac catcgaataa 1650

<210> 151

<211> 549
 <212> PRT
 <213> Trình tự nhân tạo

<220>
 <223> ImpE2 - G64D

<400> 151

```

Met Val Ala Lys Asn Ser Thr Pro Ser Thr Ala Gly His Ala Ser Ala
 1              5              10              15

His Thr Ala Glu Glu Phe Pro Val Ala Asn Ala Glu Met Ala Thr Pro
          20              25              30

Ser Ala Ile Asp Pro Asn His Gly Lys Lys Thr Ala Asp Asn Val Gly
          35              40              45

Ile Ile Phe Ala Ala Leu Met Leu Thr Met Leu Met Ser Ser Leu Asp
 50              55              60

Gln Met Ile Phe Gly Ser Ala Leu Pro Thr Ile Val Gly Glu Leu Gly
 65              70              75              80

Gly Val Asp Gln Met Ser Trp Val Ile Ser Ala Phe Met Val Thr Met
          85              90              95

Thr Ile Ala Met Pro Leu Ala Gly Gln Leu Gly Asp Arg Met Gly Arg
          100              105              110

Lys Trp Val Tyr Ile Ser Gly Ile Ser Ile Phe Val Ile Gly Ser Thr
          115              120              125

Leu Gly Gly Phe Ala Asn Gly Met Gly Met Leu Ile Thr Gly Arg Ala
 130              135              140

Ile Gln Gly Phe Gly Ala Gly Ile Met Met Ile Ser Ser Gln Ser Ile
 145              150              155              160

Val Ala Glu Val Val Ser Ala Arg Glu Arg Gly Lys Phe Met Gly Ile
          165              170              175

Met Gly Gly Val Phe Gly Val Ser Ser Val Leu Gly Pro Val Leu Gly
          180              185              190

Gly Trp Phe Thr Asp Gly Pro Gly Trp Arg Trp Gly Leu Trp Ile Asn
          195              200              205

Ile Pro Leu Gly Leu Leu Ala Ile Ile Val Cys Ala Phe Val Leu Lys
 210              215              220

Leu Arg Val Gly Glu Gln Gly Phe Lys Gly Phe Asp Trp Met Gly Phe
 225              230              235              240

Ala Ala Ile Ala Ile Thr Thr Ser Thr Leu Ile Leu Leu Thr Thr Trp
          245              250              255

Gly Gly Ser Glu Tyr Glu Trp Thr Ser Pro Thr Ile Leu Ser Met Ala

```

260					265					270					
Ala	Val	Val	Ile	Val	Gly	Ala	Leu	Leu	Thr	Val	Phe	Ile	Glu	Ser	Arg
		275					280					285			
Ala	Ser	Gln	Pro	Leu	Ile	Pro	Val	Gln	Leu	Phe	Lys	Asn	Arg	Asn	Met
	290					295					300				
Val	Leu	Thr	Thr	Leu	Ala	Gly	Thr	Val	Leu	Gly	Leu	Ala	Met	Met	Gly
305					310					315					320
Val	Leu	Gly	Tyr	Met	Pro	Thr	Tyr	Leu	Gln	Met	Val	His	Thr	Leu	Thr
				325					330					335	
Pro	Thr	Glu	Ala	Gly	Leu	Met	Met	Ile	Pro	Met	Met	Val	Gly	Met	Ile
			340					345					350		
Gly	Val	Ser	Thr	Gly	Val	Gly	Phe	Ile	Ile	Ala	Lys	Thr	Gly	Asn	Tyr
		355					360					365			
Lys	Tyr	Tyr	Pro	Ile	Ala	Gly	Leu	Ala	Ile	Thr	Ala	Phe	Ala	Leu	Trp
	370					375					380				
Trp	Met	Ser	Gln	Met	Thr	Val	Glu	Thr	Ser	Leu	Thr	Gly	Ile	Gly	Val
385					390					395					400
Arg	Phe	Leu	Val	Phe	Gly	Val	Gly	Leu	Gly	Phe	Val	Met	Gln	Val	Leu
				405					410					415	
Val	Leu	Ile	Val	Gln	Asn	Ser	Phe	Pro	Val	Ser	Gln	Val	Gly	Thr	Ala
			420					425					430		
Thr	Ala	Ala	Asn	Asn	Phe	Phe	Arg	Gln	Ile	Gly	Ser	Ala	Leu	Gly	Ala
		435					440					445			
Ser	Ile	Val	Gly	Ser	Met	Phe	Ile	His	Asn	Met	Gln	Asn	Glu	Met	Ala
	450					455					460				
Thr	Arg	Leu	Pro	Asp	Ala	Leu	Ala	Ser	Leu	Gly	Lys	Glu	Gly	Ala	Ala
465					470					475					480
Ile	Ser	Gln	Gln	Phe	Gln	Gly	Ala	Asp	Ala	Ala	Asn	Ser	Leu	Thr	Pro
				485				490						495	
His	Ala	Val	Ala	Glu	Leu	Pro	Asp	Val	Leu	Arg	Asp	Ala	Ile	Leu	Asn
			500					505					510		
Ser	Tyr	Asn	Asp	Gly	Leu	Thr	Pro	Val	Ile	Gly	Met	Met	Val	Pro	Leu
		515					520					525			
Ala	Ile	Val	Ala	Met	Leu	Ile	Leu	Phe	Pro	Leu	Arg	Gln	Glu	Arg	Leu
	530					535					540				
Lys	Glu	Thr	Ile	Glu											
545															

<210> 152

<211> 1650
 <212> ADN
 <213> Trình tự nhân tạo

<220>
 <223> ImpE2 NT - G64D

<400> 152
 atggtagcta aaaactccac cccaagcacg gccggccacg ccagtgtctca cactgcggaa 60
 gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggg 120
 aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgctcac catgctgatg 180
 agctcttttg accagatgat tttcggttcc gctctgcaa ccatcgctcg cgagctcggc 240
 ggcgtggacc agatgagctg ggtaatttca gcatttatgg tcaccatgac cattgctatg 300
 ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctctacat ctcaggtatc 360
 tccatttttcg ttattggctc gacgctcggg ggctttgcca atggcatggg catgctgatc 420
 accggacgtg caatccaggg cttcgggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
 gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
 tttggcgtct cctccgtact ggggtccagtt ctcggtggct gggtcaccga tgggtccggg 600
 tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgct 660
 ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatggggttt 720
 gcggccatcg caatcacgac cagcaccctg attctgtctca ccacttgggg cggaagcgaa 780
 tacgagtgga ctcccccaac tattttgtcc atggctgccg tagtcatcgt cggcgcgctg 840
 ctcaccgtgt tcattgagtc gcgtgcatcc cagccgctga tcccggttca gctatttaag 900
 aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
 gtgctcggct acatgccaac ctacctgcag atgggtgcaca ccctgacgcc aactgaagca 1020
 ggcttgatga tgatcccgat gatggctcggc atgatcgggtg tctccactgg tgttggcttc 1080
 atcatcgcta agaccggcaa ctacaagtac taccocatcg cgggcctggc catcacggcg 1140
 tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcggagtt 1200
 cgcttccttg tattcgggtg cggtttgggc tttgtcatgc aggtactggg gctgattgtt 1260
 caaaactcct tccctgtatc gcaggtcggg actgccacgg cggctaataa cttcttccgc 1320
 cagattgggtt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
 aatgagatgg ctaccggttt gcctgatgcc cttgcatcgt tgggcaagga aggcgccgct 1440
 atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500

gagcttcccg atgtcctccg tgacgctatc ttaaattcct acaatgacgg tctgaccccc 1560
 gtgattggca tgatggtgcc actggccatt gttgcaatgc tgattttgtt cccactgcgc 1620
 caagagcgct tgaaggaaac catcgaataa 1650

<210> 153
 <211> 1650
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE2 NT - V2I G64E

<400> 153
 atgatagcta aaaactccac cccaagcacg gccggccacg ccagtgtctca cactgcggaa 60
 gaattcccag tggccaatgc tgaaatggca acgccttcag caatcgaccc aaaccacggg 120
 aaaaagaccg cggataacgt cggcattatc ttcgctgcct tgatgctcac catgctgatg 180
 agctcttttg agcagatgat tttcggttcc gctctgccaa ccatcgctcg cgagctcggc 240
 ggcgtggacc agatgagctg ggtaatttca gcatttatgg tcaccatgac cattgctatg 300
 ccactagccg gtcagctcgg tgaccgcatg ggccgcaagt gggctacat ctcaggtatc 360
 tccattttcg ttattggctc gacgctcggg ggctttgcca atggcatggg catgctgatc 420
 accggacgtg caatccaggg cttcggtgcc ggcatcatga tgatttcctc gcagtcgatt 480
 gtggctgagg ttgtctccgc acgtgagcgc ggcaagttca tgggtattat gggcggcgtc 540
 tttggcgtct cctccgtact ggggccagtt ctcggtgggt gggtcaccga tgggtccggg 600
 tggcgttggg gcctgtggat caacattcca ctgggtctgc tggcaattat tgtctgcgct 660
 ttcgtactga agctgcgcgt gggcgagcaa ggctttaagg gctttgactg gatgggtttt 720
 gcggccatcg caatcacgac cagcaccctg attctgctca ccacttgggg cggaagcgaa 780
 tacgagtgga ctcccccaac tattttgtcc atggctgccg tagtcacgt cggcgcgctg 840
 ctcaccgtgt tcattgagtc gcgtgcatcc cagccgctga tcccggttca gctatttaag 900
 aaccgcaaca tggttttgac caccctcgcc ggtactgttt tgggtctggc catgatgggc 960
 gtgctcgggt acatgccaac ctacctgcag atgggtgcaca ccctgacgcc aactgaagca 1020
 ggcttgatga tgatcccgat gatggtcggc atgatcgggt tctccactgg tgttggttc 1080
 atcatcgcta agaccggcaa ctacaagtac taccatcg cgggcctggc catcacggcg 1140
 tttgctttgt ggtggatgtc ccagatgacc gttgagactt cattgaccgg tatcggagtt 1200

cgcttccttg tattcggtgt cggcttgggc tttgtcatgc aggtactggg gctgattggt 1260
 caaaactcct tcctgtatc gcaggtcggg actgccacgg cggctaataa cttcttccgc 1320
 cagattgggt cggcattggg tgcttccatc gtgggttcga tgttcattca caatatgcag 1380
 aatgagatgg ctaccggttt gctgatgcc cttgcatcgt tgggcaagga aggcgccgct 1440
 atttcgcagc agttccaagg tgcagatgcc gccaaactcct tgactccgca cgcagtcgca 1500
 gagcttcccg atgtcctccg tgacgctatc ttaaattcct acaatgacgg tctgaccccc 1560
 gtgattggca tgatggtgcc actggccatt gttgcaatgc tgattttggt cccactgcgc 1620
 caagagcgct tgaaggaaac catcgaataa 1650

<210> 154
 <211> 2312
 <212> ADN
 <213> Chưa biết

<220>
 <223> ImpE NT - CJI0323

<400> 154
 ttgcatgctg tgcaagaagt taatgacaat gaagaagact ccctccctgg cagtgcctc 60
 gggttaaggg agcagaagcg attggcaacc aagcatcgca tcgaagacgc cgcgacacgg 120
 ttggttgatg aatcgagctt tgacaaagta acaattgaag aaatttgca agccgccggg 180
 atttcccgac gcaccttttt taattatttc agcacgaaag aaagcgccgt tattggcgcg 240
 tcctcggaac cgttgacgga aaagcaacgc aatgacttct tgaatgctga cgccagcaat 300
 ctctgcagc tgatggttga gcagatcaaa caacacttgg agtcttctca ccagagtcaa 360
 gcgattcacg accgtcgtca gcgaatcttt gcggatccgg atgtcgcggg acgtgcaatg 420
 gcgtttcgca aggaacgctc acgggaaacc atggagctaa tcgctcaacg tcttcgggag 480
 catcctgaaa aacaacgcgc ccagaattg gatccgaaa cagaggcgat gctgctgagc 540
 ggattcattc gcgaagccac ctggatggct atctcacgac ccgatcgtga ttgtgactg 600
 ccagtgggtg accgcatcta tcgcgcgatg gaattggtaa agaattacac gaaaggactg 660
 gaatgatagc taaaaactcc accccaagca cggccggcca cgccagtgt cacactgcgg 720
 aagaattccc agtggccaat gctgaaatgg caacgccttc agcaatcgac ccaaaccacg 780
 gtaaaaagac cgcggataac gtcggcatta tcttcgctgc cttgatgctc accatgctga 840
 tgagctcttt ggagcagatg attttcgggt ccgctctgcc aaccatcgtc ggcgagctcg 900
 gcggcgtgga ccagatgagc tgggtaattt cagcatttat ggtcaccatg accattgcta 960

tgccactagc cggtcagctc ggtgaccgca tgggccgcaa gtgggtctac atctcaggta 1020
 tctccatttt cgttattggc tcgacgctcg gtggctttgc caatggcatg ggcattgctga 1080
 tcaccggacg tgcaatccag ggcttcggtg ccggcatcat gatgatttcc tcgcagtcga 1140
 ttgtggctga ggttgtctcc gcacgtgagc gcggcaagtt catgggtatt atgggcggcg 1200
 tctttggcgt ctctccgta ctgggtccag ttctcgggtg ctggttcacc gatgggtccc 1260
 gttggcgctt gggcctgtgg atcaacattc cactgggtct gctggcaatt attgtctgcg 1320
 ctttcgtact gaagctgcgc gtgggcgagc aaggctttaa gggctttgac tggatgggtt 1380
 ttgcggccat cgcaatcacg accagcacc cagattctgct caccattggg ggcggaagcg 1440
 aatacgagtg gacttcccca actattttgt ccatggctgc cgtagtcacg gtcggcgcg 1500
 tgctcaccgt gttcattgag tcgcgtgcat ccagccgct gatcccggtt cagctattta 1560
 agaaccgcaa catgggtttt accaccctcg ccggtactgt tttgggtctg gccatgatgg 1620
 gcgtgctcgg ctacatgcca acctacctgc agatgggtgca caccctgacg ccaactgaag 1680
 caggcttgat gatgatccc atgatggctg gcatgatcgg tgtctccact ggtgttggct 1740
 tcatcatcgc taagaccggc aactacaagt actaccccat cgcgggcctg gccatcacgg 1800
 cgtttgcttt gtgggtggatg tcccagatga ccgttgagac ttcatgacc ggtatcggag 1860
 ttcgcttcct tgtattcggg gtcggcttgg gctttgtcat gcagggtactg gtgctgattg 1920
 ttcaaaactc ctccctgta tcgcaggctg gtactgccac ggcggttaat aacttcttcc 1980
 gccagattgg ttcggcattg ggtgcttcca tcgtgggttc gatgttcatt cacaatatgc 2040
 agaatgagat ggctaccctg ttgcctgatg cccttgcatc gttgggcaag gaaggcgccg 2100
 ctatttcgca gcagttccaa ggtgcagatg ccgccaactc cttgactccg cacgcagtcg 2160
 cagagcttcc cgatgtctc cgtgacgcta tcttaaattc ctacaatgac ggtctgacct 2220
 ccgtgattgg catgatgggt ccaactggcca ttgttgcaat gctgattttg ttcccactgc 2280
 gccagagcg cttgaaggaa accatcgaat aa 2312