



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2017년07월10일  
(11) 등록번호 10-1756338  
(24) 등록일자 2017년07월04일

(51) 국제특허분류(Int. Cl.)  
C12N 15/77 (2006.01) C07K 14/34 (2006.01)  
C12N 15/52 (2006.01) C12N 9/10 (2006.01)  
C12P 13/12 (2006.01)  
(52) CPC특허분류  
C12N 15/77 (2013.01)  
C07K 14/34 (2013.01)  
(21) 출원번호 10-2016-0005378  
(22) 출원일자 2016년01월15일  
심사청구일자 2016년01월15일  
(56) 선행기술조사문헌  
US20020086373 A1\*  
GenBank Accession Number CP005959  
(2014.01.30.)\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
고려대학교 산학협력단  
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)  
(72) 발명자  
한성욱  
서울특별시 성북구 길음로 119, 209동 1503호 (길음동, 길음뉴타운(대우푸르지오아파트))  
주영철  
서울특별시 성북구 인촌로 46, 304호 (안암동3가, 에스엠오피스텔)  
(뒷면에 계속)  
(74) 대리인  
이처영, 장제환

전체 청구항 수 : 총 2 항

심사관 : 최준호

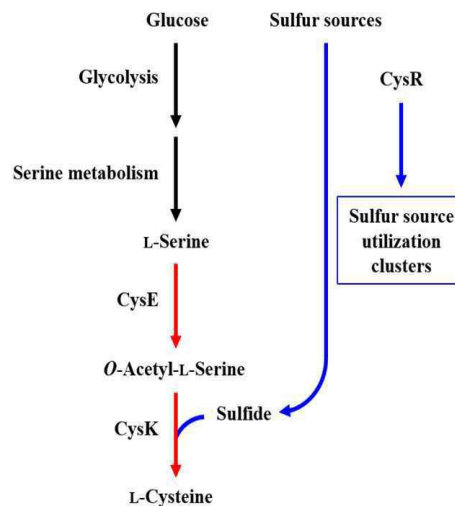
(54) 발명의 명칭 L-시스테인 생산용 변이미생물 및 이를 이용한 L-시스테인의 제조방법

(57) 요약

본 발명은 L-시스테인 대사경로 관련 효소를 코딩하는 유전자가 도입된 L-시스테인 생산용 변이미생물에 관한 것으로, 더욱 자세하게는 L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 *cysE*, *cysK* 및 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물 및 상기 변이 미생물을 이용한 L-시스테인의 제조방법에 관한 것이다.

본 발명에 따르면, 변이 미생물의 L-시스테인 대사경로와 관련된 대사흐름 조절과 L-시스테인에 필수적인 황원 공급 시스템을 조절하여, 고효율로 L-시스테인을 제조할 수 있다.

대표도 - 도1



(52) CPC특허분류

- C12N 15/52 (2013.01)
- C12N 9/1029 (2013.01)
- C12N 9/1085 (2013.01)
- C12P 13/12 (2013.01)
- C12Y 203/0103 (2013.01)
- C12Y 205/01047 (2013.01)

**유승규**

전라북도 전주시 덕진구 중상보로 10, 102동 504호  
(우아동2가, 롯데아파트)

(72) 발명자

**현정은**

서울특별시 성북구 인촌로13길 18-21, 1층 (안암동2가)

이 발명을 지원한 국가연구개발사업

과제고유번호 2014R1A2A1A11049949

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 (이공)중견연구자지원\_도약연구

연구과제명 [1차년도] 친환경 고분자 기질을 이용하는 코리네박테리움의 지방족 및 함황 아미노산 생  
합성 경로 연구 및 유용물질 합성

기여율 1/1

주관기관 고려대학교 산학협력단

연구기간 2014.11.01 ~ 2015.10.31

---

**명세서**

**청구범위**

**청구항 1**

코리네박테리움 글루타미쿰에 L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 서열번호 1의 염기서열로 표시되는 *cysE*, 서열번호 2의 염기서열로 표시되는 *cysK* 및 서열번호 3의 염기서열로 표시되는 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물.

**청구항 2**

삭제

**청구항 3**

삭제

**청구항 4**

삭제

**청구항 5**

삭제

**청구항 6**

다음 단계를 포함하는 L-시스테인의 제조방법:

- (a) 제1항의 변이미생물을 배양하여 L-시스테인을 생성시키는 단계; 및
- (b) 생성된 L-시스테인을 수득하는 단계.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 L-시스테인 생성능을 가지는 미생물에 L-시스테인 대사경로 관련 효소를 코딩하는 유전자가 도입된 L-시스테인 생산용 변이미생물에 관한 것으로, 더욱 자세하게는 L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 *cysE*, *cysK* 및 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물 및 상기 변이 미생물을 이용한 L-시스테인의 제조방법에 관한 것이다.

**배경 기술**

[0002] 최근 원재료소서 생물합성 유래의 L-시스테인에 대한 식품, 의약품, 화장품 등 다양한 산업분야에서 수요가 증가하고 친환경 생물학적 소재에 대한 소비자의 욕구가 증가하고 있는 추세이다.

[0003] L-시스테인은 비필수 황함 아미노산의 일종으로, 기존에는 산 가수분해 및 전기환원 방법을 통하여 인모 또는 동물의 털로부터 추출하여 왔으나, 이러한 종래의 방법은 산업적으로 폐기물의 처리량과 생산비용이 증가하는 문제가 발생하며 사회적으로는 극심한 환경오염과 동물유래원료사용에 대한 문제를 초래하였다.

[0004] 현재 세계시장에서 L-시스테인은 중국에서 전량을 생산하여 해외로 수출하며 국내에서는 L-시스테인에 대한 수요와 수입량이 증가하고 있는 추세이나 국내 L-시스테인 생산기업이 없는 실정이다.

[0005] 국내의 L-시스테인 생산 기술이 전무한 상황에서 중국의 물리화학적 처리법에 의한 L-시스테인을 대체할 친환경 생물학적 소재 L-시스테인 기술 개발로 인하여 다양한 산업분야의 원재료 공급이 가능할 것으로 예상되어, 기술

수요가 증가하고 있다.

[0006] 이에, 본 발명자들은 미생물 발효에 의하여 L-시스테인을 대량생산할 수 있는 기술을 개발하고자 예의 노력한 결과, L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 *cysE*, *cysK* 및 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물을 이용할 경우, L-시스테인의 합성에 필수적인 황원 공급이 미생물 내에서 지속적으로 가능하여, L-시스테인 생성능을 크게 향상시킬 수 있다는 것을 확인하고, 본 발명을 완성하게 되었다.

### 발명의 내용

#### 해결하려는 과제

[0007] 본 발명의 목적은 L-시스테인의 생성능을 향상시킨 변이 미생물을 제공하는데 있다.

[0008] 본 발명의 다른 목적은 상기 변이 미생물을 이용한 L-시스테인의 제조방법을 제공하는데 있다.

#### 과제의 해결 수단

[0009] 상기 목적을 달성하기 위하여, 본 발명은 L-시스테인 생성능을 가지는 미생물에 L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 *cysE*, *cysK* 및 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물을 제공한다.

[0010] 본 발명은 또한, (a) 상기 변이미생물을 배양하여 L-시스테인을 생성시키는 단계; 및 (b) 생성된 L-시스테인을 수득하는 단계를 포함하는 L-시스테인의 제조방법을 제공한다.

[0011]

#### 발명의 효과

[0012] 본 발명에 따르면, 변이 미생물의 L-시스테인 대사경로와 관련된 대사흐름 조절과 L-시스테인의 합성에 필수적인 황원 공급 시스템을 조절하여, 고효율로 L-시스테인을 제조할 수 있다.

#### 도면의 간단한 설명

[0013] 도 1은 코리네박테리움 글루타미쿰에서 지속 가능한 L-시스테인 생산을 위한 과발현과 황원흡수시스템의 조절에 관한 대사경로를 나타낸 것이다.

도 2는 코리네박테리움 글루타미쿰 항시성 과발현 재조합 벡터(pMT1-s)의 구조를 나타낸 것이다.

도 3은 코리네박테리움 글루타미쿰 과발현 재조합 벡터(pMT-tac)의 구조를 나타낸 것이다.

도 4는 코리네박테리움 글루타미쿰 CysE 과발현 재조합 벡터 pMT-tac::cysE의 구조를 나타낸 것이다.

도 5는 코리네박테리움 글루타미쿰 CysK 과발현 재조합 벡터 pMT-tac::cysK의 구조를 나타낸 것이다.

도 6은 코리네박테리움 글루타미쿰 CysR 과발현 재조합 벡터 pMT-tac::cysR의 구조를 나타낸 것이다.

도 7은 코리네박테리움 글루타미쿰 CysEK 과발현 재조합 벡터 pMT-tac::cysEK의 구조를 나타낸 것이다.

도 8은 코리네박테리움 글루타미쿰 CysER 과발현 재조합 벡터 pMT-tac::cysER의 구조를 나타낸 것이다.

도 9는 코리네박테리움 글루타미쿰 CysKR 과발현 재조합 벡터 pMT-tac::cysKR의 구조를 나타낸 것이다.

도 10은 코리네박테리움 글루타미쿰 CysEKR 과발현 재조합 벡터 pMT-tac::cysEKR의 구조를 나타낸 것이다.

도 11은 코리네박테리움 글루타미쿰에서 야생형과 단일 재조합 균주(pMT-tac::cysE, pMT-tac::cysK, pMT-tac::cysR)의 L-시스테인 생합성 활성을 나타낸 것이다.

도 12는 코리네박테리움 글루타미쿰에서 야생형과 이중 또는 삼중 재조합 균주(pMT-tac::cysEK, pMT-tac::cysER, pMT-tac::cysKR, pMT-tac::cysEKR)의 L-시스테인 생합성 활성을 나타낸 것이다.

도 13은 코리네박테리움 글루타미쿰에서 야생형과 단일, 이중 또는 삼중 재조합 균주(pMT-tac::cysR, pMT-tac::cysEK, pMT-tac::cysER, pMT-tac::cysKR, pMT-tac::cysEKR)의 세포 내 셀파이드 축적능을 나타낸 것이다.

#### 발명을 실시하기 위한 구체적인 내용

- [0014] 달리 정의되지 않는 한, 본 명세서에서 사용된 모든 기술적 및 과학적 용어들은 본 발명이 속하는 기술분야에서 숙련된 전문가에 의해서 통상적으로 이해되는 것과 동일한 의미를 갖는다. 일반적으로, 본 명세서에서 사용된 명명법 및 이하에 기술하는 실험 방법은 본 기술분야에서 잘 알려져 있고 통상적으로 사용되는 것이다.
- [0015] 기존의 미생물 발효에 의한 L-시스테인의 생산 기술은 아미노산 및 황 대사경로에 존재하는 피드백 억제 조절로 인하여 L-시스테인의 합성이 제한되어, L-시스테인을 생물학적으로 생산하는 기술에는 낮은 생산성 등과 같은 한계가 있었다.
- [0016] 본 발명에서는 미생물 발효에서 L-시스테인의 생산 수율, 생산 속도, 및 생산성을 획기적으로 증진시키기 위해 L-시스테인 및 황 유입 대사경로의 피드백 억제 조절에 따르지 않는 L-시스테인 합성 관련 효소 과발현과 황 유입 경로 전사활성화인자에 의한 조절 해제된 시스템 대사경로를 구축하였다.
- [0017] 따라서, 본 발명은 일 관점에서, L-시스테인 생성능을 가지는 미생물에 L-시스테인 대사경로 관련 효소를 코딩하는 유전자인 *cysE*, *cysK* 및 *cysR*가 도입되어 있는 L-시스테인 생산용 변이미생물에 관한 것이다.
- [0018] 본 발명의 일 양태에서는, 코리네텡테리움 글루타미쿰 균주로부터 유래한 세린아세틸트렌스포레이즈, 아세틸세린 설파이드릴레이즈, 및 황 유입 유전자군 활성화 트랜스크립션레귤레이터 유전자를 강력한 전사를 유도하는 잡종 촉진유전자(*lacUV5* 촉진유전자와 *trp* 촉진유전자를 결합)인 *tac* 촉진유전자를 에 제작하여 코리네텡테리움 글루타미쿰 균주에 도입함으로써 각 효소의 과발현을 유도시켰다. 높은 수준으로 과발현된 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*CysR*)는 코리네텡테리움 글루타미쿰 균주 세포 외부의 황원을 세포 내부로 유입하고 유입된 황원을 설파이드로 전환하였으며, 이후, 높은 수준으로 과발현된 세린아세틸트렌스포레이즈(*CysE*)는 코리네텡테리움 글루타미쿰 균주가 생산한 세포내 세린으로부터 아세틸세린으로 전환하고, 이후, 높은 수준으로 과발현된 아세틸세린 설파이드릴레이즈는 세린아세틸트렌스포레이즈(*CysK*)에 의해 생합성된 아세틸세린을 L-시스테인으로 전환하게 된다.
- [0019] 본 발명의 상기 *cysE*는 서열번호 1의 염기서열을 가지는 것을 특징으로 할 수 있으며, 상기 *cysK*는 서열번호 2의 염기서열을 가지는 것을 특징으로 할 수 있으며, 상기 *cysR*은 서열번호 3의 염기서열을 가지는 것을 특징으로 할 수 있다.
- [0020] 본 발명의 L-시스테인 생산능을 가지는 미생물은 코리네텡테리움 글루타미쿰을 사용하는 것이 바람직하나, 이에 한정되는 것은 아니다.
- [0021] 본 발명의 L-시스테인 생산 변이 미생물은 L-시스테인을 전구체로 하는 Co-A, 타우린, N-아세틸시스테인, 메티오닌과 같은 산업적으로 유용한 물질을 생산하는 공정에도 매우 유용하게 사용될 것으로 판단된다.
- [0022] 본 발명의 일양태에서 제조된 코리네텡테리움 글루타미쿰 *CysEKR*은 2015년 12월 23일자로 한국생명공학연구원 수탁번호 : KCTC12970BP로 기탁하였다.
- [0023] 다른 관점에서, 본 발명은 (a) 상기 변이미생물을 배양하여 L-시스테인을 생성시키는 단계; 및 (b) 생성된 L-시스테인을 수득하는 단계를 포함하는 L-시스테인의 제조방법에 관한 것이다.
- [0024] 본 발명의 일 양태에서는 세린아세틸트렌스포레이즈 유전자, 아세틸세린 설파이드릴레이즈 유전자, 및 황 유입 유전자군 활성화 트랜스크립션레귤레이터 유전자를 도입시킨 코리네텡테리움 균주가 야생형 및 상기 유전자를 중 하나 혹은 둘만을 도입한 균주보다 높은 L-시스테인 생산능을 나타내는 것을 확인하였다(도 11 및 도 12).
- [0025] 본 발명에서, 용어 "벡터 (vector)"는 적합한 숙주 내에서 DNA를 발현시킬 수 있는 적합한 조절 서열에 작동가능하게 연결된 DNA 서열을 함유하는 DNA 제조물을 의미한다. 벡터는 플라스미드, 파지 입자, 또는 간단하게 잠재적 게놈 삽입물일 수 있다. 적당한 숙주로 형질전환 되면, 벡터는 숙주 게놈과 무관하게 복제하고 기능할 수 있거나, 또는 일부 경우에 게놈 그 자체에 통합될 수 있다. 플라스미드가 현재 벡터의 가장 통상적으로 사용되는 형태이므로, 본 발명의 명세서에서 "플라스미드 (plasmid)" 및 "벡터 (vector)"는 때로 상호 교환적으로 사용된다. 그러나, 본 발명은 당업계에 알려진 또는 알려지게 되는 바와 동등한 기능을 갖는 벡터의 다른 형태를 포함한다. 포유동물 세포 배양물 발현을 위한 전형적인 발현 벡터는 예를 들면 pRK5 (EP 307,247호), pSV16B (WO 91/08291호) 및 pVL1392 (Pharmingen)을 기초로 한다.
- [0026] 본 발명에서 "증폭"이란 해당 유전자의 일부 염기를 변이, 치환, 또는 삭제시키거나, 일부 염기를 도입시키거나, 또는 동일한 효소를 코딩하는 다른 미생물 유래의 유전자를 도입시켜 대응하는 효소의 활성을 증가시키는 것을 포괄하는 개념이다.

- [0027] 본 발명에서 “발현 조절 서열 (expression control sequence)” 이라는 표현은 특정한 숙주 생물에서 작동가능하게 연결된 코딩 서열의 발현에 필수적인 DNA 서열을 의미한다. 그러한 조절 서열은 전사를 실시하기 위한 프로모터, 그러한 전사를 조절하기 위한 임의의 오퍼레이터 서열, 적합한 mRNA 리보솜 결합 부위를 코딩하는 서열 및 전사 및 해독의 종결을 조절하는 서열을 포함한다. 예를 들면, 원핵생물에 적합한 조절 서열은 프로모터, 임의로 오퍼레이터 서열 및 리보솜 결합 부위를 포함한다. 진핵세포는 프로모터, 폴리아데닐화 시그널 및 인핸서가 이에 포함된다. 플라스미드에서 유전자의 발현 양에 가장 영향을 미치는 인자는 프로모터이다. 고발현용의 프로모터로서 SR α 프로모터와 사이토메가로바이러스 (cytomegalovirus) 유래 프로모터 등이 바람직하게 사용된다.
- [0028] 본 발명의 DNA 서열을 발현시키기 위하여, 매우 다양한 발현 조절 서열중 어느 것이라도 벡터에 사용될 수 있다. 유용한 발현 조절서열의 예에는, 예를 들어, SV40 또는 아데노바이러스의 초기 및 후기 프로모터들, lac 시스템, trp 시스템, TAC 또는 TRC 시스템, T3 및 T7 프로모터들, 파지 램다의 주요 오퍼레이터 및 프로모터 영역, fd 코드 단백질의 조절 영역, 3-포스포글리세레이트 키나제 또는 다른 글리콜분해 효소에 대한 프로모터, 상기 포스포타제의 프로모터들, 예를 들어 Pho5, 효모 알파-교배 시스템의 프로모터 및 원핵세포 또는 진핵 세포 또는 이들의 바이러스의 유전자의 발현을 조절하는 것으로 알려진 구성과 유도의 기타 다른 서열 및 이들의 여러 조합이 포함된다. T7 RNA 폴리메라아제 프로모터  $\Phi 10$ 은 이. 콜라이에서 단백질 NSP를 발현시키는데 유용하게 사용될 수 있다.
- [0029] 핵산은 다른 핵산 서열과 기능적 관계로 배치될 때 “작동가능하게 연결 (operably linked)” 된다. 이것은 적절한 분자 (예를 들면, 전사 활성화 단백질)은 조절 서열(들)에 결합될 때 유전자 발현을 가능하게 하는 방식으로 연결된 유전자 및 조절 서열(들)일 수 있다. 예를 들면, 전서열(pre-sequence) 또는 분비 리더 (leader)에 대한 DNA는 폴리펩타이드의 분비에 참여하는 전단백질로서 발현되는 경우 폴리펩타이드에 대한 DNA에 작동가능하게 연결되고; 프로모터 또는 인핸서는 서열의 전사에 영향을 끼치는 경우 코딩서열에 작동가능하게 연결되거나; 또는 리보솜 결합 부위는 서열의 전사에 영향을 끼치는 경우 코딩 서열에 작동가능하게 연결되거나; 또는 리보솜 결합 부위는 번역을 용이하게 하도록 배치되는 경우 코딩 서열에 작동가능하게 연결된다. 일반적으로, “작동가능하게 연결된”은 연결된 DNA 서열이 접촉하고, 또한 분비 리더의 경우 접촉하고 리딩 프레임 내에 존재하는것을 의미한다. 그러나, 인핸서 (enhancer)는 접촉할 필요가 없다. 이들 서열의 연결은 편리한 제한 효소 부위에서 라이게이션(연결)에 의해 수행된다. 그러한 부위가 존재하지 않는 경우, 통상의 방법에 따른 합성 올리고뉴클레오티드 어댑터 (oligonucleotide adaptor) 또는 링커(linker)를 사용한다.
- [0030] 본원 명세서에 사용된 용어 “발현 벡터”는 통상 이중의 DNA의 단편이 삽입된 재조합 캐리어 (recombinant carrier)로서 일반적으로 이중 가닥의 DNA의 단편을 의미한다. 여기서, 이중 DNA는 숙주 세포에서 천연적으로 발견되지 않는 DNA인 이형 DNA를 의미한다. 발현 벡터는 일단 숙주 세포내에 있으면 숙주 염색체 DNA와 무관하게 복제할 수 있으며 벡터의 수 개의 카피 및 그의 삽입된 (이중) DNA가 생성될 수 있다.
- [0031] 당업계에 주지된 바와 같이, 숙주세포에서 형질감염 유전자의 발현 수준을 높이기 위해서는, 해당 유전자가, 선택된 발현 숙주 내에서 기능을 발휘하는 전사 및 해독 발현 조절 서열에 작동가능하도록 연결되어야만 한다. 바람직하게는 발현 조절서열 및 해당 유전자는 세균 선택 마커 및 복제 개시점 (replication origin)을 같이 포함하고 있는 하나의 발현 벡터 내에 포함되게 된다. 발현 숙주가 진핵세포인 경우에는, 발현 벡터는 진핵 발현 숙주 내에서 유용한 발현 마커를 더 포함하여야만 한다.
- [0032] 상술한 발현 벡터에 의해 형질전환 또는 형질감염된 숙주 세포는 본 발명의 또 다른 측면을 구성한다. 본원 명세서에 사용된 용어 “형질전환”은 DNA를 숙주로 도입하여 DNA가 염색체의 인자로서 또는 염색체 통합완성에 의해 복제가가능하게 되는 것을 의미한다. 본원 명세서에 사용된 용어 “형질감염”은 임의의 코딩 서열이 실제로 발현되든 아니든 발현 벡터가 숙주 세포에 의해 수용되는 것을 의미한다.
- [0033] 발명의 숙주 세포는 원핵 또는 진핵생물 세포일 수 있다. 또한, DNA의 도입효율이 높고, 도입된 DNA의 발현효율이 높은 숙주가 통상 사용된다. 이. 콜라이, 슈도모나스, 바실러스, 스트렙토마이세스, 진균, 효모와 같은 주지의 진핵 및 원핵 숙주들, 이 숙주 세포의 예이다.
- [0034] 물론 모든 벡터와 발현 조절 서열이 본 발명의 DNA 서열을 발현하는데 모두 동등하게 기능을 발휘하지는 않는다는 것을 이해하여야만 한다. 마찬가지로 모든 숙주가 동일한 발현 시스템에 대해 동일하게 기능을 발휘하지는 않는다. 그러나, 당업자라면 과도한 실험적 부담없이 본 발명의 범위를 벗어나지 않는 채로 여러 벡터, 발현 조절 서열 및 숙주 중에서 적절한 선택을 할 수 있다. 예를 들어, 벡터를 선택함에 있어서는 숙주를 고려하여야 하는데, 이는 벡터가 그 안에서 복제되어야만 하기 때문이다. 벡터의 복제 수, 복제 수를 조절할 수 있는

능력 및 당해 벡터에 의해 코딩되는 다른 단백질, 예를 들어 항생제 마커의 발현도 또한 고려되어야만 한다. 발현 조절 서열을 선정함에 있어서도, 여러 가지 인자들을 고려하여야만 한다. 예를 들어, 서열의 상대적 강도, 조절가능성 및 본 발명의 DNA 서열과의 상용성 등, 특히 가능성있는 이차 구조와 관련하여 고려하여야 한다. 단세포 숙주는 선정된 벡터, 본 발명의 DNA 서열에 의해 코딩되는 산물의 독성, 분비 특성, 단백질을 정확하게 폴딩시킬 수 있는 능력, 배양 및 발효 요건들, 본 발명 DNA 서열에 의해 코딩되는 산물을 숙주로부터 정제하는 것의 용이성 등의 인자를 고려하여 선정되어야만 한다.

[0035] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다. 이들 실시예는 오로지 본 발명을 예시하기 위한 것으로서, 본 발명의 범위가 이들 실시예에 의해 제한되는 것으로 해석되지는 않는 것은 당업계에서 통상의 지식을 가진 자에게 있어서 자명할 것이다.

[0036] **실시예 1: 코리네박테리움 글루타미쿰 형질전환용 과발현 재조합 벡터 제작**

[0037] 코리네박테리움 글루타미쿰 형질전환용 재조합 벡터인 pMT-s (도 2)에서 항시발현 p-out 촉진유전자와 발현 단백질의 세포외 분비를 촉진하는 신호 펩타이드 서열을 제거하기 위해 5'에 접착말단을 형성하는 제한효소 *Xho I* 와 3'에 제한효소 *Cla I*으로 절단하고, *lacUV5* 촉진유전자와 *trp* 촉진유전자로 결합된 과발현수준의 잠종촉진유전자인 *tac* 촉진유전자와 발현 조절억제자가 암호화된 *lacI* 유전자를 PCR 및 오버랩 중합효소연쇄반응(Overlap-PCR)을 수행하여 연결한 후, 상기 제한효소로 절단하고 pMT1-s에 삽입하여, 코리네박테리움 글루타미쿰 과발현 재조합 벡터 pMT-*tac*을 제조하였으며(도 3), pMT-*tac*의 염기서열을 서열번호 4에 나타내었다.

[0038] **실시예 2: 코리네박테리움 글루타미쿰 유래 L-시스테인 대사경로 관련 효소 유전자 확보**

[0039] 실시예 1에서 제작한 코리네박테리움 글루타미쿰 과발현 재조합 벡터 pMT-*tac*에 코리네박테리움 글루타미쿰 균주로부터 유래한 세린아세틸트랜스포레이즈(*cysE*), 아세틸세린 설파이드릴레이즈는 세린아세틸트랜스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*)의 유전자를 클로닝하기 위하여 각각 서열번호 1-3의 염기서열을 참고로 하여 정방향 프라이머(Forward primer)의 5'에는 제한효소 *ClaI*, 역방향 프라이머(Reverse primer)의 5'에는 제한효소 *NotI*의 서열이 각각 삽입된 프라이머를 합성하였다.

[0040] 세린아세틸트랜스포레이즈-리보솜 결합 부위(Ribosomal binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트랜스포레이즈(*cysEK*), 세린아세틸트랜스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysER*), 아세틸세린 설파이드릴레이즈 세린아세틸트랜스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysKR*), 세린아세틸트랜스포레이즈-리보솜 결합 부위(Ribosomal binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트랜스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysEKR*)을 연결하기 위하여 각각 연결부위에 리보솜 결합 부위(Ribosomal binding site)의 서열이 삽입된 프라이머(Forward 및 reverse primer)를 합성하였다. 이후, 상기 합성된 프라이머를 이용하여 PCR 및 오버랩 중합효소연쇄반응(Overlap-PCR)을 수행하였다. 사용된 프라이머의 서열을 표 1에 나타내었다.

표 1

tac F(서열번호 5)	CCATTCCATGGGTGCTTTGACAATTAATCATCGGCTCGTATAAATGTGT
tac R ( <i>Cla</i> I) (서열번호 6)	GGGATCGATATGATATCTCCTGTGTGAAATTGTTATCCG
lacI F( <i>Xho</i> I) (서열번호 7)GGG <u>CTCGAG</u> AGCCTGGGGTGCCATAATGAG	
lacI R (서열번호 8) <u>GACACC</u> ATGGAATGGTGCAAAACC	
cysE F( <i>Bam</i> HI) (서열번호 9)CGCGC <u>GGATCC</u> ATGCTCTCGACAATAAAAAATGATC	
cysE R ( <i>Kpn</i> I) (서열번호 10)	CA <u>GGTACCT</u> TAGTGGTGGTGGTGGTGGTGAATGTAATAGTCCGGATCGA
cysK F( <i>Bam</i> HI) (서열번호 11) GCGC <u>GGATCC</u> ATGGGCAATGTGTACAACAA	
cysK R ( <i>Kpn</i> I)(서열번호 12)	6)CAGA <u>GGTACCT</u> TAGTGGTGGTGGTGGTGGTGGTGAATGTAATAGTCCGGATCGA
cysR F ( <i>Bam</i> HI) (서열번호 13) GCGCGC <u>GGATCC</u> ATGATTGGCTATGGTTTACC	
cysR R ( <i>Kpn</i> I) (서열번호 14)	CGC <u>GGTACCT</u> TAATGATGATGATGATGATGATGGGGTACGAGAGTAAGTGG
OE cysEK MF (서열번호 15)	<u>CTAAAAGGAGATATAGATGGGCAATGTGTACAACAACATCACCGAAACC</u>
OE cysEK MR (서열번호 16)	<u>CATCTATATCTCCTTTTAGTGGTGGTGGTGGTGGTGAATGTAATAGTCC</u>
OE cysER MF (서열번호 17)	<u>AAAAGGAGATATAGATGATTGGCTATGGTTTACCTATGCCCCAATCAGGC</u>
OE cysER MR (서열번호 18)	<u>CATCTATATCTCCTTTTAGTGGTGGTGGTGGTGGTGAATGTAATAGTCC</u>
OE cysKR MF (서열번호 19)	<u>AAAAGGAGATATAGATGATTGGCTATGGTTTACCTATGCCCCAATCAGGC</u>
OE cysKR MR (서열번호 20)	<u>GCCAAATCATCTATATCTCCTTTTAGTGGTGGTGGTGGTGGTGAATGTAATAGTCC</u>
OE cysEKR MF (서열번호 21) <u>AATCCTCGAAGACACCGACGGCAAC</u>	
OE cysEKR MR (서열번호 22) <u>GTTGCCGTCGGTGTCTTCGAGGATT</u>	

\* 제관호소부위 (이빨릭), 5말단 뉴클레오타이드 확장부위 (밑줄), 리보솜결합부위 (음영)

[0041]

[0042]

실시예 3: *cysE*, *cysK*, *cysR*, *cysEK*, *cysER*, *cysKR*, *cysEKR* 유전자를 각각 삽입한 재조합 벡터의 제조 및 코리네박테리움 글루타미쿰 균주에 형질전환

[0043]

실시예 2에서 수득한 코리네박테리움 글루타미쿰 균주로부터 유래한 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드릴레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈(*cysEK*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysER*), 아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysKR*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysEKR*)의 유전자를 0.8% 아가로스 겔 상에서 전기영동하였고 아가로스 겔 상의 DNA 절편은 Gel extraction kit (GeneAll)를 사용하여 회수하였다.

[0044]

그 후, 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드릴레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal



binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈(*cysE*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysER*), 아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysKR*), 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-아세틸세린 설파이드릴레이즈 세린아세틸트렌스포레이즈-리보솜 결합 부위(Ribosomal binding site)-황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysEKR*)의 유전자는 제한효소 *NotI*과 제한효소 *ClaI*으로 절단한 후 제한효소 *NotI*과 제한효소 *ClaI*으로 절단된 코리네박테리움 글루타미쿰 과발현 재조합 벡터 pMT-*tac*에 라이게이션(ligation)시켜 에스세리시아 콜라이 DH5 $\alpha$ 에 형질전환을 하였다. 이어 형질전환체로부터 라이게이션(ligation)된 재조합 플라스미드 DNA를 분리하였다. 이어 분리된 재조합 플라스미드 DNA를 코리네박테리움 글루타미쿰 ATCC 13032에 형질전환을 하였다.

[0045] 상기 재조합 벡터를 pMT-*tac::cysE*, pMT-*tac::cysK*, pMT-*tac::cysR*, pMT-*tac::cysEK*, pMT-*tac::cysER*, pMT-*tac::cysKR*, pMT-*tac::cysEKR*이라 명명하였고 도 4 ~ 도 10에 도시하였다.

[0046] 상기 형질전환된 코리네박테리움 글루타미쿰 형질전환체를 코리네박테리움 글루타미쿰 CysE, 코리네박테리움 글루타미쿰 CysK, 코리네박테리움 글루타미쿰 CysR, 코리네박테리움 글루타미쿰 CysEK, 코리네박테리움 글루타미쿰 CysER, 코리네박테리움 글루타미쿰 CysKR, 코리네박테리움 글루타미쿰 CysEKR이라 명명하였다. 코리네박테리움 글루타미쿰 CysEKR은 2015년 12월 23일자로 한국생명공학연구원 수탁번호 : KCTC12970BP로 기탁하였다.

[0047] **실시예 4: *cysE*, *cysK*, *cysR* 형질전환의 시스테인 생산량 분석**

[0048] 실시예 3에서 확보한 코리네박테리움 글루타미쿰 CysE, 코리네박테리움 글루타미쿰 CysK, 코리네박테리움 글루타미쿰 CysR 형질전환체와 코리네박테리움 글루타미쿰 야생형 균주와 시스테인 생산량을 비교하였다.

[0049] 각각의 코리네박테리움 글루타미쿰 형질전환체 및 야생형 균주는 CGXII 액체 배지 플라스크에서 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드릴레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*) 단백질이 IPTG로 발현되도록 하는 조건을 조성하여 30 $^{\circ}$ C, 200 rpm에서 10시간 동안 발효를 진행하였다. 그 후, 배양액을 채취하고 원심분리하여 세포만을 회수하고 메탈올-퀵칭법으로 해당 세포를 용해시켰다. 용해된 세포시료를 건조한 후, Waters AccQ-tag 아미노산 분석 키트(Waters, 미국)를 이용하여 HPLC로 분석하였다.

[0050] 그 결과, 도 11에 나타난 바와 같이, 야생형에 비하여 코리네박테리움 글루타미쿰 CysE, 코리네박테리움 글루타미쿰 CysK, 코리네박테리움 글루타미쿰 CysR 형질전환체에서 비교적 L-시스테인을 더 많이 생성하는 것을 확인할 수 있었다.

[0051] **실시예 5: *cysEK*, *cysER*, *cysKR*, *cysEKR* 형질전환체의 시스테인 생산량 분석**

[0052] 실시예 3에서 수득한 코리네박테리움 글루타미쿰 CysEK, 코리네박테리움 글루타미쿰 CysER, 코리네박테리움 글루타미쿰 CysKR, 코리네박테리움 글루타미쿰 CysEKR 형질전환체와 코리네박테리움 글루타미쿰 야생형 균주와 시스테인 생산량을 비교하였다.

[0053] 각각의 코리네박테리움 글루타미쿰 형질전환체 및 야생형 균주는 CGXII 액체 배지 플라스크에서 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드릴레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*) 단백질이 IPTG로 발현되도록 하는 조건을 조성하여 30 $^{\circ}$ C, 200 rpm에서 10시간 동안 발효를 진행하였다. 그 후, 배양액을 채취하고 원심분리하여 세포만을 회수하고 메탈올-퀵칭법으로 해당 세포를 용해하였다. 용해된 세포시료를 건조한 후, Waters AccQ-tag 아미노산 분석 키트를 이용하여 HPLC로 분석하였다.

[0054] 그 결과, 도 12에 나타난 바와 같이, 코리네박테리움 글루타미쿰 CysEK, 코리네박테리움 글루타미쿰 CysER, 코리네박테리움 글루타미쿰 CysKR, 코리네박테리움 글루타미쿰 CysEKR 형질전환체는 모두 야생형 균주에 비하여 L-시스테인 생합성 활성이 증가하였으며, 특히, CysEKR 형질전환체의 경우는 CysEK, CysER 및 CysKR 형질전환체보다 월등히 높은 L-시스테인 생합성능을 나타내었다.

[0055]

[0056] **실시예 6: *cysR*, *cysEK*, *cysER*, *cysKR*, *cysEKR* 형질전환체와 야생형 균주의 세포내 황 축적능 분석**

[0057] 실시예 3에서 확보한 코리네박테리움 글루타미쿰 CysR, 코리네박테리움 글루타미쿰 CysEK, 코리네박테리움 글루

타믹쿰 CysER, 코리네박테리움 글루타믹쿰 CysKR, 코리네박테리움 글루타믹쿰 CysEKR 형질전환체와 코리네박테리움 글루타믹쿰 야생형 균주와 세포 내 설파이드 축적능을 비교하였다.

[0058] 각각의 코리네박테리움 글루타믹쿰 형질전환체 및 야생형 균주는 CGXII 액체 배지 플라스크에서 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드틸레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*) 단백질이 IPTG로 발현되도록 하는 조건을 조성하여 30℃, 200 rpm에서 10시간 동안 발효를 진행하였다. 그 후, 배양액을 채취하고 원심분리하여 세포만을 회수하고 6% (wt/vol) NaOH를 첨가하여 95℃에서 15분 동안 해당 세포를 용해시켰다. 용해된 세포시료는 메틸렌블루법을 이용하여 분광광도계로 흡광도를 분석하여, 세포내 황 축적능을 비교하였다.

[0059] 황을 정량하기 위해서 표준물질인 황화나트륨을 농도별로 흡광도를 측정하여 이를 표준곡선을 그려 일차방정식을 구하였다. 이후, 모든 황 농도는 이 일차방정식에 흡광도를 대입해 정량화하였다.

[0060] 그 결과, 도 13에 나타난 바와 같이, 세포내 황농도는 코리네박테리움 글루타믹쿰 CysR 균주에서 가장 높은 것으로 확인되었다.

[0061] 실시예 7: CysEKR 형질전환체의 시스테인 생산량 분석

[0062] 실시예 3에서 확보한 코리네박테리움 글루타믹쿰 CysEKR 형질전환체와 코리네박테리움 글루타믹쿰 야생형 균주의 시간에 따른 세포성장, 글루코스비량, 시스테인 생산량을 분석하였다.

[0063] 각각의 코리네박테리움 글루타믹쿰 형질전환체 및 야생형 균주는 CGXII 액체 배지 플라스크에서 세린아세틸트렌스포레이즈(*cysE*), 아세틸세린 설파이드틸레이즈는 세린아세틸트렌스포레이즈(*cysK*), 황 유입 유전자군 활성화 트랜스크립션레귤레이터(*cysR*) 단백질이 IPTG로 발현되도록 하는 조건을 조성하여 30℃, 200 rpm에서 15시간까지 발효를 진행하였다.

[0064] 그 후, 도 14에 나타난 바와 같이, 해당 시간에 배양액을 채취하고 원심분리하여 상등액은 글루코스량을 측정하기 위해 HPLC로 분석하였다. 원심분리 후, 회수된 세포는 메탈올-퀵칭법으로 해당 세포를 용해시켰다. 용해된 세포시료를 건조한 후, Wetters AccQ-tag 아미노산 분석 키트를 이용하여 HPLC로 분석하였다. 세포성장은 도 14와 같이 해당 시간에 배양액을 채취하고 배양액을 분광광도계로 흡광을 분석하였으며, 그 결과를 도 14에 나타내었다.

[0065] 이상으로 본 발명 내용의 특정한 부분을 상세히 기술하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적 기술은 단지 바람직한 실시양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의하여 정의된다고 할 것이다.

**수탁번호**

[0066]

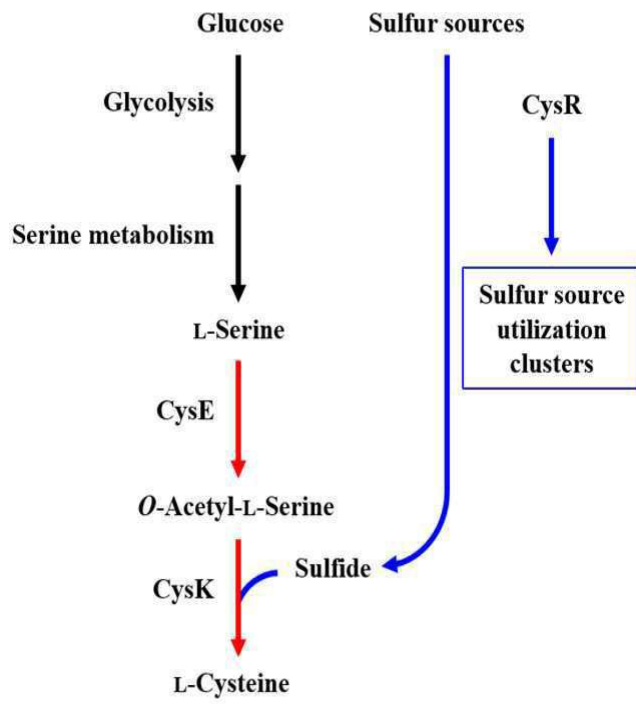
기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC12970BP

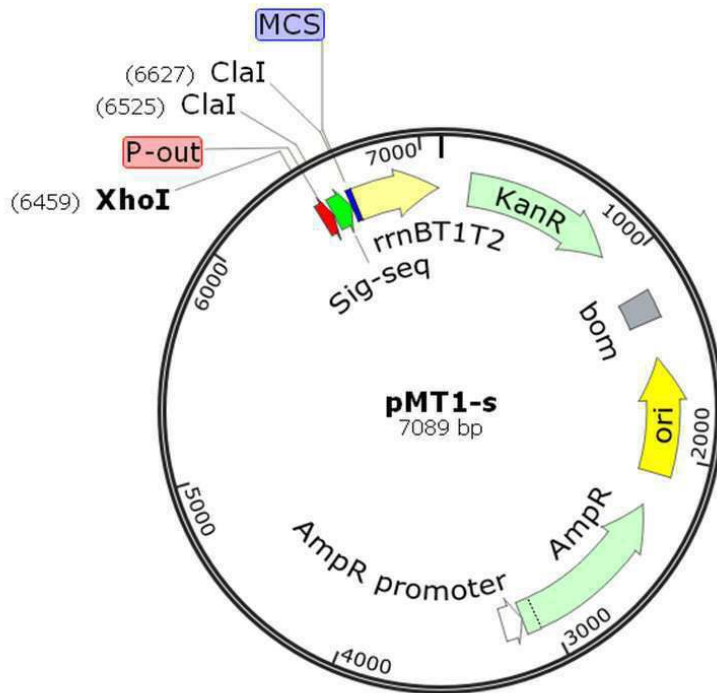
수탁일자 : 20151223

도면

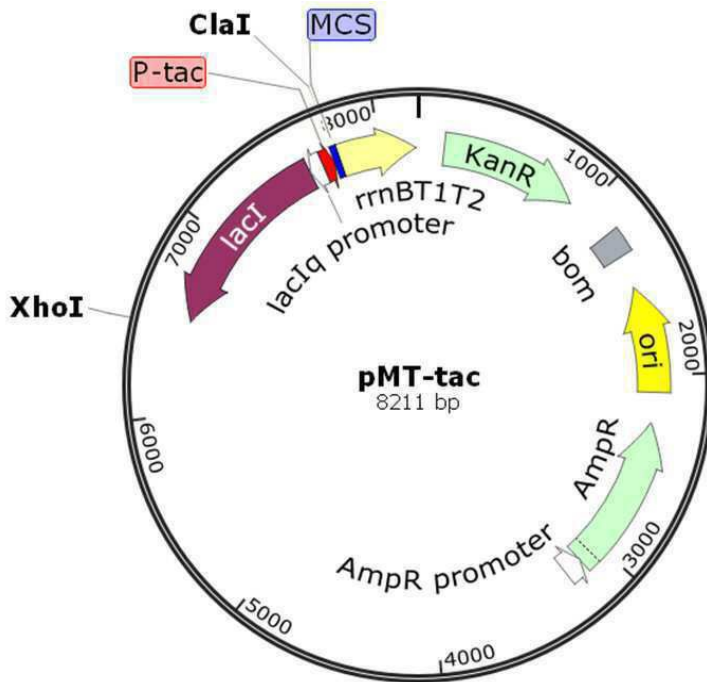
도면1



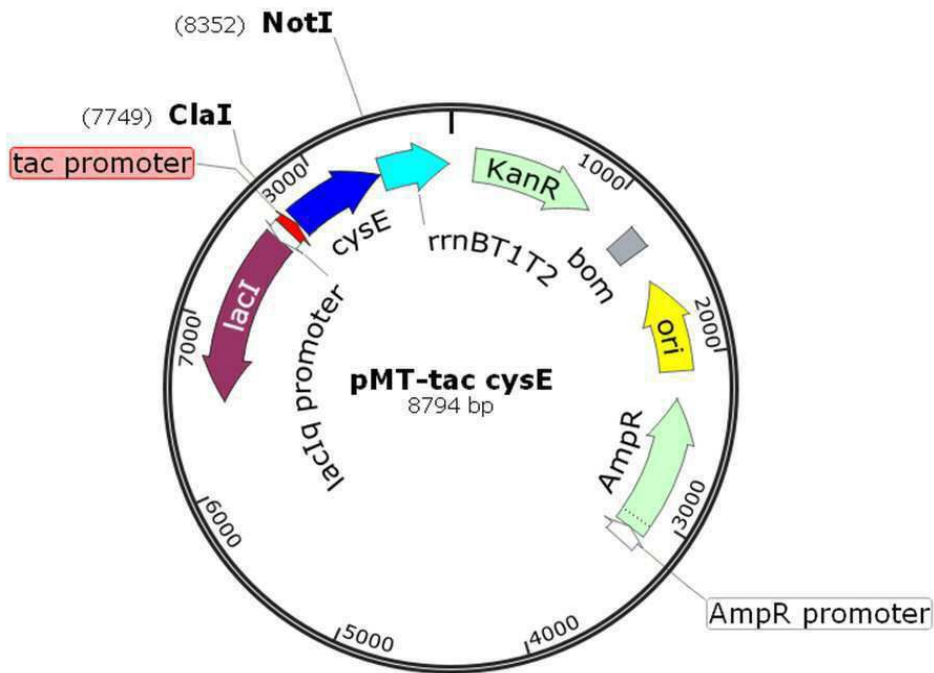
도면2



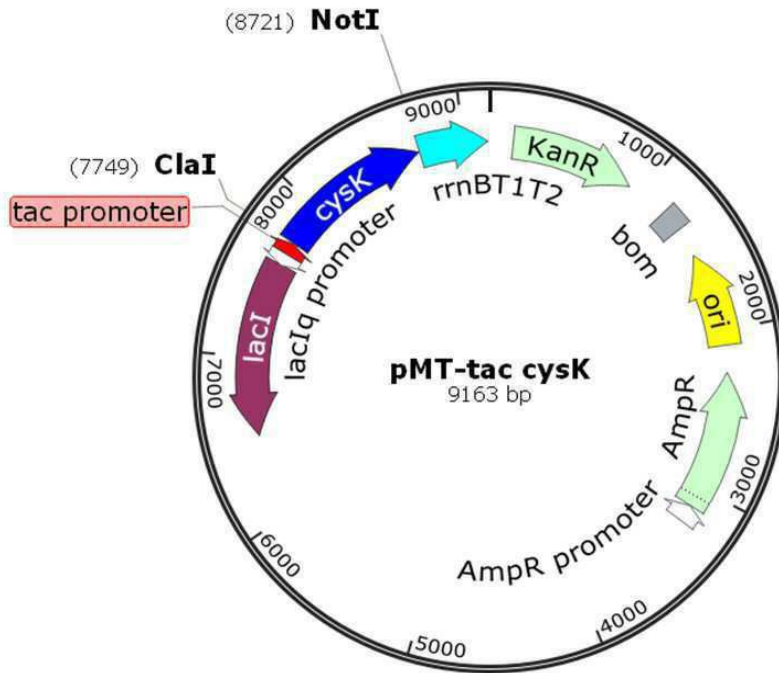
도면3



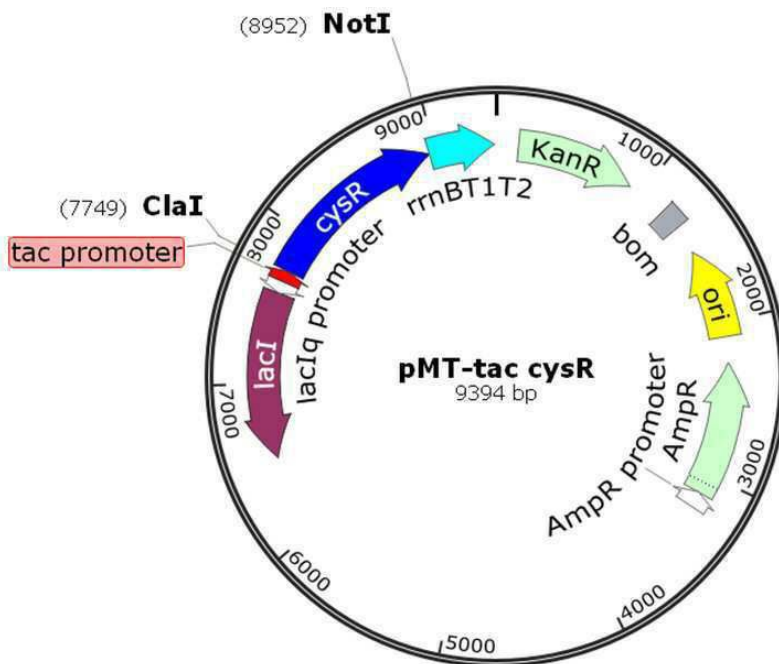
도면4



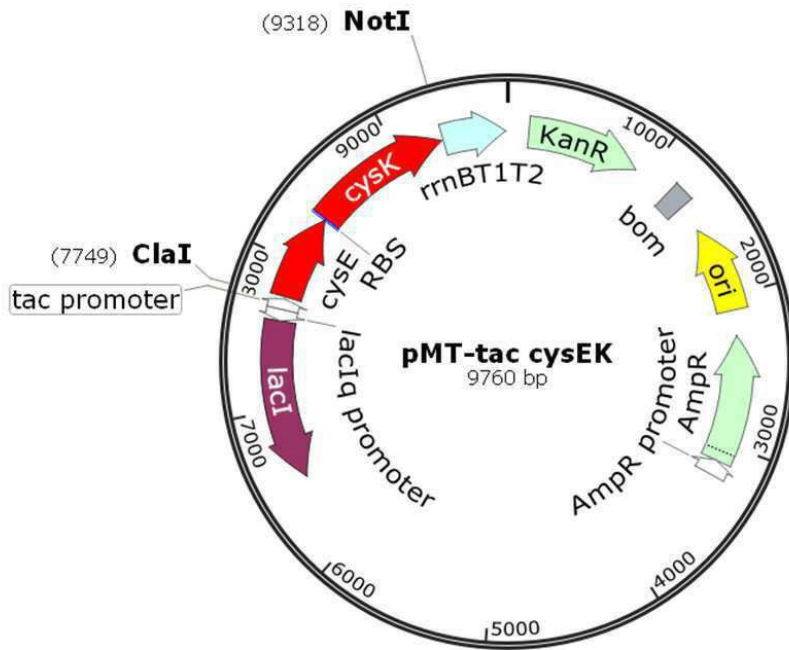
도면5



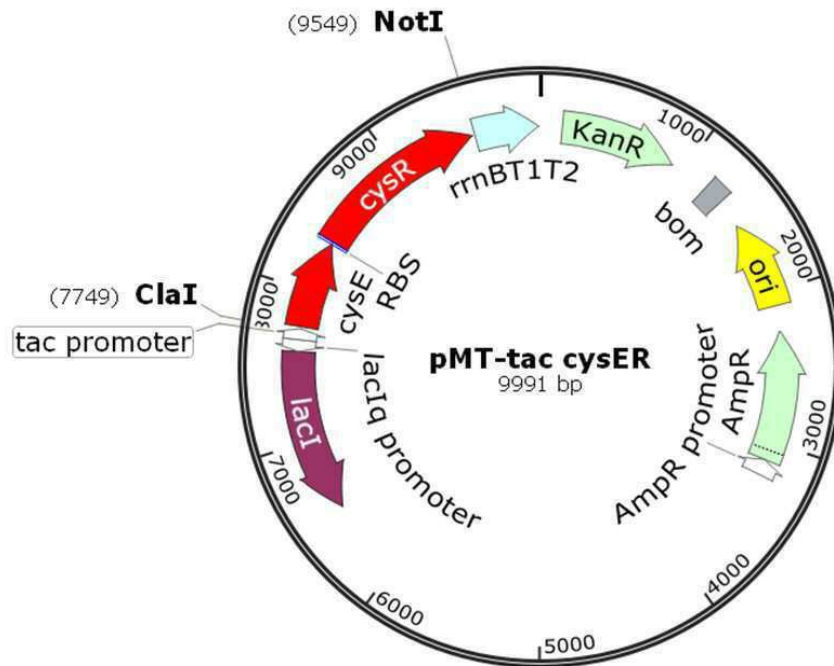
도면6



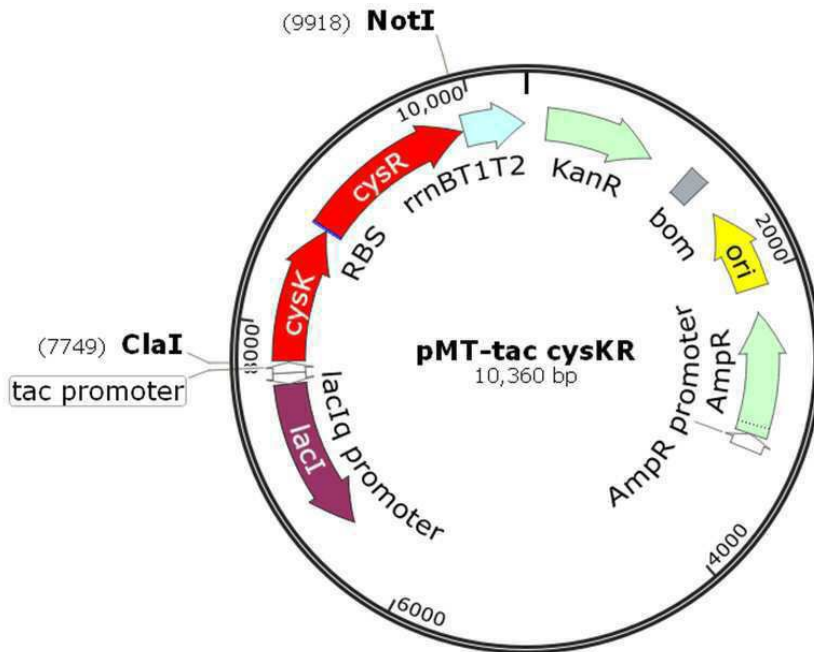
도면7



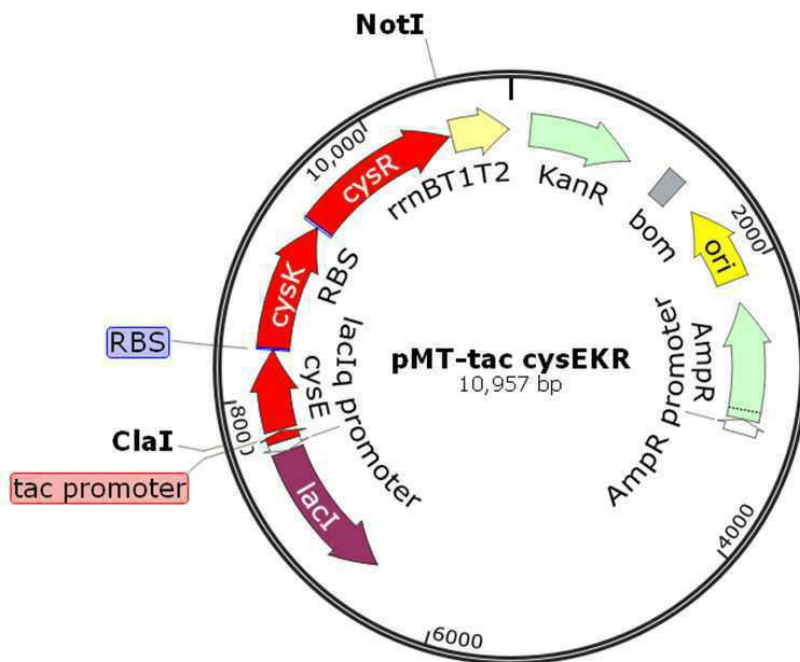
도면8



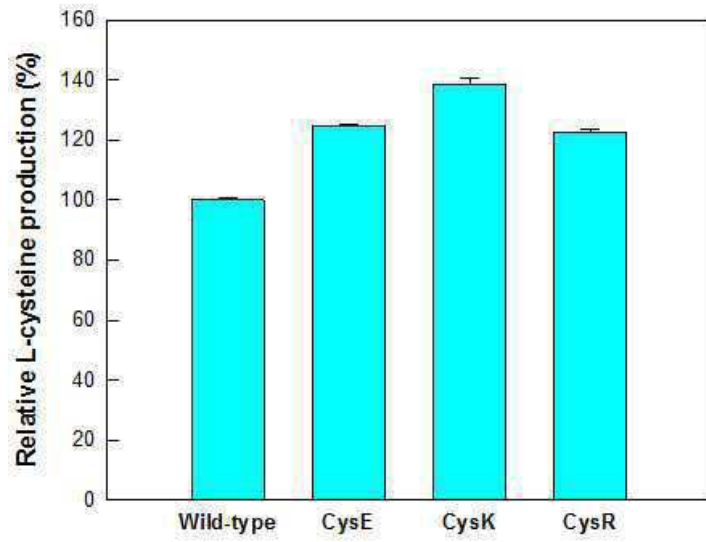
도면9



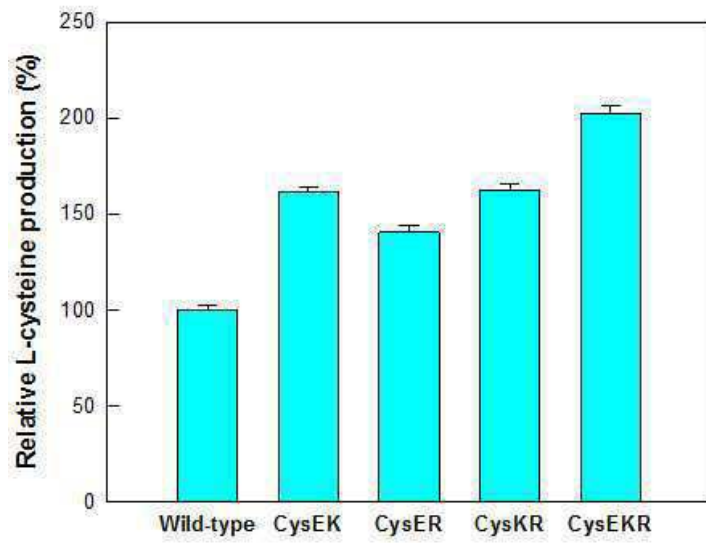
도면10



도면11

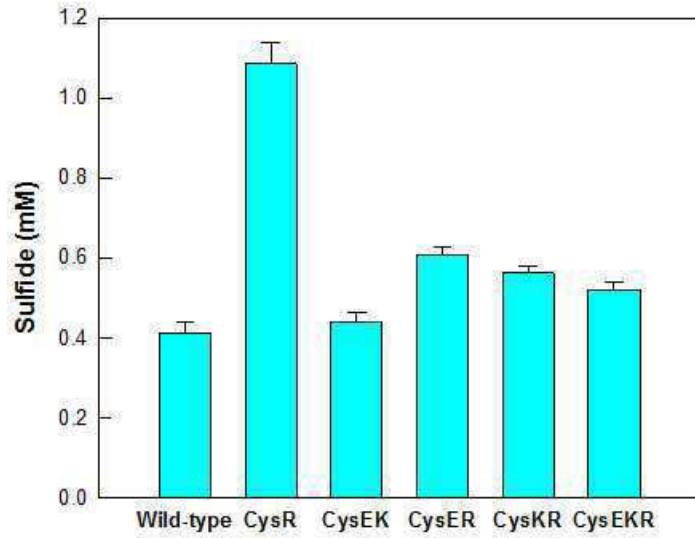


도면12





도면13



서열 목록

- <110> Korea University Research and Business Foundation
- <120> Variant Microorganism for Producing L-Cystein and Method for Preparing L-Cystein Using thereof
- <130> P15-B212
- <160> 22
- <170> KopatentIn 2.0
- <210> 1
- <211> 585
- <212> DNA
- <213> corynebacterium glutamicum
- <400> 1

```

atgctctcga caataaaaat gatccgtgaa gatctcgcaa acgctcgtga acacgatcca      60
gcagcccag gcgatttaga aaacgcagtg gtttactcgg gactccacgc catctgggca      120
catcgagttg ccaacagctg gtggaatcc ggtttccgcg gccccgccg cgtattagcc      180

caattcacc gattctcac cggcattgaa attcaccgcc gtgccacat tggctgtcgc      240
ttttttattg accacggaat gggaatcgtc atcggcgaaa ccgctgaaat cggcgaaggc      300
gcatgctct accacggcgt caccctcggc ggacaggttc tcacccaac caagcggcac      360
cccagctct gcgacaagt gacagtcggc gcggggcgcaa aatcttagg teccatcacc      420
atcggcgaag gctccgcaat tggcgccaat gcagttgtca ccaaagacgt gccggcagaa      480
cacatcgag tcggaattcc tgcgtagca cgcccacgtg gcaagacaga gaagatcaag      540
    
```

ctcgtcgatc cggactatta cattcaccac caccaccacc actaa 585

<210> 2

<211> 954

<212> DNA

<213> corynebacterium glutamicum

<400> 2

atgggcaatg tgtacaacaa catcaccgaa accatcggcc acaccccact ggtaaagctg 60

aacaagctca ccgaaggcct cgacgcaact gtctctgtca agcttgagtc attcaaccca 120

gcaaactccg tcaaggaccg tatcggctctg gccatcgttg aagatgcaga gaagtccggt 180

gcaactgaagc caggcggcac catcgttgaa gcaacctccg gcaacaccgg tatcgactg 240

gcaatggctg gcgctgcacg cggatacaac gttgttctca ccatgccgga gaccatgtcc 300

aacgagcgtc gcgttctcct ccgcgcttac ggtgcagaga tcgttcttac cccaggtgca 360

gcaggcatgc agggtgcaaa ggacaaggca gacgaaatcg tcgctgaacg cgaaaacgca 420

gtccttgctc gccagttcga gaacgaggca aaccacgcg tccaccgca caccaccgcg 480

aaggaaatcc tcgaagacac cgacggcaac gttgatatct tcgttgcaag cttcggcacc 540

ggcggaaaccg tcaccggcgt tggccaggtc ctgaaggaaa acaacgcaga cgtacaggtc 600

tacaccgtcg agccagaagc gtceccactt ctgaccgtg gcaaggctgg tccacacaag 660

atccagggca tcggcgcaaa cttcatccc gaggtcctgg accgcaaggt tctcgacgac 720

gtgctgaccg tctccaacga agacgaatc gcattctccc gcaagctcgc taccgaagag 780

ggcatcctcg gcggtatctc caccggcgca aacatcaagg cagctcttga ccttgacgca 840

aagccagaga acgctggcaa aaccatcgtc accgttgca ccgacttcgg cgagcgtac 900

gtctccaccg ttctttacga agacatccgc gaccatcacc atcaccacca ctaa 954

<210> 3

<211> 1185

<212> DNA

<213> corynebacterium glutamicum

<400> 3

atgattggct atggtttacc tatgeccaat caggcccact tctctgcgtc ctttgeccgc 60

ccctctacc cggctgcaaa gtgcatgcac catatccgc tcggccagca actcattaga 120

aatgagctgg tcgaggccac aggtctgtcc caaccgactg tcaccgcgc agtcaccgct 180

ttaatgcagg caggtttggc tcgtgaacgc cctgatctca cactctcacc gggccctggc 240

cgtccaata ttctctaga actcgtcca agtccatgga ttcattgcagg cgtggcaatc 300  
 ggcaccaagt ctctctactg cgctttgttt gataccaagg gtcgcaccct tcgtgatgcc 360  
 atgctggaaa tctcagcagc tgatttagat ccagacactt tcatcgaaca cctcattgct 420  
 ggtgtcaacc gcctcaccac tggctttgat ctaccactgg taggtattgg tgttgccacc 480  
 tcaggaaaag tcaccaacgc aggcgtttgc accgcaagca acttgggctg ggatggcgtt 540  
 gatatcgctg gccgcctgaa ctaccaattc agcgttcag caaccgtggc atcagcaatt 600  
  
 cctgccatcg cagcttctga actgcaggct tccccacttc cccaccctga gcagccaact 660  
 cccatcacct tgaccttcta cgccgatgac tctgtgggcg cggcctacag caatgatttg 720  
 ggagtacatg tcattggacc actggctaca actcgtggat caggtttgga tactttgggc 780  
 atggctgccg aagatgcgt gagcaccaa ggtttcttaa gcagggtttc tgatcagggt 840  
 atctttgcca acagccttgg tgagctagtc accattgcta aagacaatga aaccgcacgg 900  
 gaattcctca acgatcgcgc gaccctgctg gctcacactg ccgcagaagc tgccgaaaca 960  
 gttaagccat ccaccctggt tctctcggga tcggcgtttt ccgaagatcc acaagtcgg 1020  
  
 tcggtgttcg ctccaatt gaagaaggaa tacgacgag acattgagct ccggttgatc 1080  
 cccaccacc gggaaaactg ccgcgagca gctcgagcag tcgcacttga tcgactactc 1140  
 aacgagccac ttactctcgt accccatcat catcatcacc attag 1185  
  
 <210> 4  
 <211> 8211  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> pMT-tac  
 <400> 4  
  
 cagtttgttg aagattagat gctataattg ttattaaag gattgaagga tgcttaggaa 60  
 gacgagttat taatagctga ataagaacgg tgctctcaa atattctat ttagaaaagc 120  
 aaatctaaaa ttatctgaaa agggaatgag aatagtgaat ggaccaataa taatgactag 180  
  
 agaagaaaga atgaagattg ttcattgaaat taaggaacga atattggata aatatgggga 240  
 tgatgttaag gctattggtg tttatggctc tcttggctgt cagactgatg ggcctattc 300  
 ggatattgag atgatgtgtg tcatgtcaac agaggaagca gagttcagcc atgaatggac 360  
 aaccggtgag tggaaaggtg aagtgaattt tgatagcga gagattctac tagattatgc 420  
 atctcaggtg gaatcagatt ggccgcttac acatggtcaa ttttctcta ttttgccgat 480  
 ttatgattca ggtggatact tagagaaagt gtatcaaac gctaaatcgg tagaagccca 540

aacgttccac gatgcgattt gtgcccttat cgtagaagag ctgtttgaat atgcaggcaa 600  
  
 atggcgtaat attcgtgtgc aaggaccgac aacatttcta ccatccttga ctgtacaggt 660  
 agcaatggca ggtgccatgt tgattggctc gcatcatcgc atctgttata cgacgagcgc 720  
 ttcggtctta acigaagcag ttaagcaatc agatcttctc tcaggttatg accatctgtg 780  
 ccagttcgta atgtctggtc aactttccga ctctgagaaa cttctggaat cgctagagaa 840  
 tttctggaat gggattcagg agtggacaga acgacacgga tatatagtgg atgtgtcaaa 900  
 acgcatacca tttgaacga tgacctctaa taattgttaa tcatgttggg tacgtattta 960  
 ttaacttctc ctagtattag taattatcat ggctgtcatg gcgcattaac ggaataaagg 1020  
  
 gtgtgcttaa atcgggctgc ctcgcgctt tcggtgatga cggtgaaaac ctctgacaca 1080  
 tgcagctccc ggagacggtc acagcttctc tgtaagcggg tgccgggagc agacaagccc 1140  
 gtcagggcgc gtcagcgggt gttggcgggt gtcggggcgc agccatgacc cagtcacgta 1200  
 gcgatagcgg agtgtatact ggcttaacta tgcggcatca gagcagattg tactgagagt 1260  
 gcaccatatg cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg 1320  
 ctcttccgct tectcgtca ctgaectcgt gcgctcggtc gttcggctgc ggcgagcggg 1380  
 atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa 1440  
  
 gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgcctggc 1500  
 gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtacagag 1560  
 gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccttgaa gctccctcgt 1620  
 gcgctctcct gttccgacct tgccgcttac cggatactg tccgctttc tcccttcggg 1680  
 aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg aggtcgttcg 1740  
 ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg cttatccgg 1800  
 taactatcgt cttgagtcca acccgtaag acacgaetta tcgccactgg cagcagccac 1860  
  
 tggtaacagg attagcagag cgaggtatgt aggcgggtct acagagtctc tgaagtggg 1920  
 gcctaactac ggctacacta gaaggacagt atttggatc tcgctctgc tgaagccagt 1980  
 taccttcgga aaaagagttg gtagctcttg atccggcaaa caaacaccg ctggtagcgg 2040  
 tggtttttt gtttgaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 2100  
 tttgatctt tctaccgggt ctgacgctca gtggaacgaa aactcacgtt aagggatctt 2160  
 ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaattaaa aatgaagttt 2220  
 taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 2280

tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt 2340  
 cgtgtagata actacgatac gggagggtt accatctggc cccagtgtcg caatgatacc 2400  
 gcgagacca cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc 2460  
 cgagcgcaga agtggctctg caactttatc cgcctccatc cagtctatta attgttgccg 2520  
 ggaagctaga gtaagtagtt gccagttaa tagtttgcgc aacgttgttg ccattgtctg 2580  
 aggcatcgtg gtgtcacgct cgtcgttttg tatggcttca ttcagctcgg gttcccaacg 2640  
 atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct ctttcggtcc 2700  
  
 tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact 2760  
 gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc 2820  
 aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaac 2880  
 acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc 2940  
 ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaaccac 3000  
 tcgtgcacc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa 3060  
 aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact 3120  
  
 catactcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg 3180  
 atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg 3240  
 aaaagtgcc cctgatcgtc cattccgaca gcategccag tcaactatggc gtgctgctag 3300  
 cgctatatgc gttgatgcaa tttctatgcg cacccttct cggagcactg tccgaccgt 3360  
 ttggccgccg cccagtctg ctgcttctgc tacttggagc cactatcgac tacgcgatca 3420  
 tggcgaccac acccgtcctg tggatcgccg tagagcgatt gaagaccgtc aaccaaaggg 3480  
 gaagcctcca atcgacgcga cgcgcgctct acggcgatcc tgacgcagat ttttagctat 3540  
  
 ctgtcgcage gccctcaggg acaagccacc cgcacaacgt cgcgagggcg atcagcgacg 3600  
 ccgcagtact gatcaagtcg gtcaagccaa gcgcaaccag cggcaccgcc gcgagcaacg 3660  
 tcgcaagggc gatcagggga cgatttttgc gaagaatttc cacggtaaga atccaatctc 3720  
 tcgaatttag ggtgaaagaa gcttggcata ggggtgtgca cgaactcggg ggaggaaatt 3780  
 tccgcggggc aaggcttgc gaagcggagt cgcggcagtg gcttgaaga tctttgggag 3840  
 cagtcttgt gcgcttacga ggtgagccgg tggggaaccg ttatctgcct atggtgtgag 3900  
 cccccctaga gagcttcaag agcaatcagc ccgacctaga aaggaggcca agagagagac 3960  
  
 ccctacgggg ggaaccgttt tctgcctacg agatggcaca tttactggga agctttacgg 4020  
 cgtctctgtg gaagttcaat gcccgagac ttaagtctc tattcacggt ctgacgtgac 4080  
 acgctaaatt cagacatagc ttcatgatt gtcgccacg agccagtctc tcctcaaca 4140

gtcataaacc aacctgcaat ggtcaagcga tttcctttag ctttcctagc ttgtcgttga 4200  
 ctggacttag ctagtttttc tcgctgtgct cgggcgtact cactgttttg gtctttccag 4260  
 cgttctgcgg cctttttacc gccacgtctt cccatagtgg ccagagcttt tcgccctcgg 4320  
 ctgctctcgc tcctgtctg acgagcaggg acgactggct ggccttttagc gacgtagccg 4380  
  
 cgcacacgtc gcgccatcgt ctggcgggtca cgcacggcg gcagatcagg ctcacggccg 4440  
 tctgctccga ccgcctgagc gacggtgtag gcacgctcgt aggcgtcgat gatcttgggtg 4500  
 tcttttaggc gtcaccagc cgcttttaac tggtatccca cagtcaaagc gtggcgaaaa 4560  
 gccgtctcat cacgggcggc acgccctgga gcagtccaga ggacacggac gccgtcgatc 4620  
 agctctccag acgcttcagc ggcgctcggc aggcttgctt caagcgtggc aagtgctttt 4680  
 gcttccgcag tggcttttct tgccgcttcg atacgtgcc gtccgctaga aaactcctgc 4740  
 tcatagcgtt ttttaggttt ttctgtgcct gagatcatgc gagcaacctc cataagatca 4800  
  
 gctaggcgat ccacgcgatt gtgctgggca tgccagcggg acgcggtggg atcgtcggag 4860  
 acgtgcagtg gccaccggct cagcctatgt gaaaaagcct ggtcagcgcc gaaaacgcgg 4920  
 gtcatttctt cggctgttgc agccagcagg cgcatactcg ggctgctcat gectgctcgc 4980  
 gcatacaccg gatcaatgag ccagatgagc tggcatttcc cgctcagtgg attcacgccg 5040  
 atccaagctg gcgcttttcc caggcgtgcc cagcgtcca aaatcgcgta gacctcgggg 5100  
 tttacgtgct cgattttccc gccggcctgg tggctcggca catcaatgc caggacaagc 5160  
 acggctcgtg gctgcgcgtg cgtcagagca acatactggc accgggcaag cgattttgaa 5220  
  
 ccaactcggg ataacttcgg ctgtgtttct cccgtgtccg ggtctttgat ccaagcgtg 5280  
 gcgaagtgcg gggctttgct gccctggaaa ttttctctgc ccaggtgagc gaggaattcg 5340  
 cggcggctct cgtcgtcca gccacgtgat cgcagcgcga gctcgggatg ggtgtcgaac 5400  
 agatcagcgg aaaatttcca ggccggtgtg tcaatgtctc gtgaatccgc tagagtcatt 5460  
 tttgagcgtt ttctcccagg tttggactgg gggttagccg acgccctgtg agttaccgct 5520  
 cacggggcgt tcaacatttt tcaggtattc gtgcagctta tcgcttcttg ccgcctgtgc 5580  
 gctttttcga cgcgcgacgc tgctccgat tcggtgcagg tggcggcgc getgacacgt 5640  
  
 cctgggcggc cacggccaca cgaaacgcgg catttacgat gtttgcctg cctgcgggca 5700  
 ccgcgccacg atcgcggata attctcgtg ccgcttcag ctctgtgacg accatggcca 5760  
 aaatttcgct cgggggacgc acttcagcgc ccatttgcga cctagccgcc tccagctcct 5820  
 cggcgtggcg tttgttggcg cgctcgcggc tgctcgcggc acgacacgca tctgagcaat 5880  
 attttgcgcg ccgtcctcgc gggtcaggcc ggggaggaat caggccaccg cagtaggcgc 5940  
 aactgattcg atctccact actgtgcgtc ctctggcgc tgccgagcac gcagctcgtc 6000

agccagctcc tcaagatccg ccacgagagt ttctaggtcg ctcgcgccac tggcccagtc 6060

tcgtgatgct ggcgcgtccg tcgtatcgag agctcggaaa aatccgatca ccgtttttaa 6120

atcgacggca gcatcgagcg cgctcggactc cagcgcgaca tcagagagat ccatagctga 6180

tgattcgggc caatthtggc acttcgtcgt gaaggtcatg acaccattat aacgaacgtt 6240

cgttaaagt tttggcggaa aatcacgcgg cacgaaaatt ttcacgaagc gggactttgc 6300

gcagctcagg ggtgctaaaa atthtgtatc gcacttgatt ttcgaaaga cagatatctg 6360

caaacggtgt gtcgtattct ggcttggtht taaaaatctg gaatcgaat tgcggggcga 6420

ccgagaagtt tttacaaaag gcaaaaactt ttcgggatct cgagagcctg ggtgctctaa 6480

tgagtgagct aacttacatt aattcgttg cgctcactgc ccgctttcca gtcgggaaac 6540

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 6600

gggcgccagg gtggtthtcc tthtccaccag tgagacgggc aacagctgat tgcctttcac 6660

gcctggccc tgagagagtt gcagcaagcg gtccacgtg gthtgcctca gcaggcgaac 6720

atctgttht atggtggtta acggcgggat ataacatgag ctgtcttcgg tatcgtcgtg 6780

tcccactacc gagatatccg caccaacgcg cagcccggac tcggtaatgg cgcgcattgc 6840

gcccagcgc atctgatctg tggcaaccag catcgcagtg ggaacgatgc cctcattcag 6900

catttgcatg gthtgtgaa aaccggacat ggcactccag tcgctttccc gttccgctat 6960

cggctgaatt tgattgcgag tgagatattt atgccagcca gccagacgca gacgcgccga 7020

gacagaactt aatgggccc ctaacagcgc gatttgctgg tgaccaatg cgaccagatg 7080

ctccacgccc agtcgcgtac cgctttcatg ggagaaaata ataactgtga tgggtgtctg 7140

gtcagagaca tcaagaaata acgccggaac attagtgag gcagcttcca cagcaatggc 7200

atctgtgta tccagcgat agttaatgat cagcccactg acgcgttcg cgagaagatt 7260

gtgcaccgce gctttacagg ctccgacgce gcttcgttct accatcgaca ccaccacgt 7320

ggcaccagc tgatcggcgc gagatthaat cgccgcgaca atthtgcagc gcgcgtgcag 7380

ggccagactg gaggtggcaa cgcaatcag caacgactgt ttgcccgcca gthtgtgtgc 7440

cacgcggtg ggaatgtaat tcagctccgc catcgcctt tccacttht cccgcgtht 7500

cgcagaacg tggtggcct ggttcaccac gcgggaaacg gtctgataag agacaccgce 7560

atactctgce acatcgtata acgttactgg thtccattc accacctga attgactctc 7620

thtccggcgc tatcatgcca taccgcgaaa gthtthtgcac cattccatgg tgtcttgaca 7680

attaatcacc ggctcgtata atgtgtggaa thtgcagcgg ataacaattht cacacaggag 7740

atatcatatc gatggatccc gggtagccgc gccgcctgtt ttggcggatg agagaagatt 7800  
 ttcagcctga tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgcct 7860  
 ggccggcagta gcgcgggtgt cccacctgac cccatgccga actcagaagt gaaacgccgt 7920  
 agcgcctgat gtagtgtggg gtctcccat gcgagagtag ggaactgcca ggcattcaaat 7980  
 aaaacgaaag gctcagtcga aagactgggc ctttcgtttt atctgttgtt tgcggtgaa 8040  
 cgctctctg agtaggacaa atccgccggg agcggatttg aacgttgca agcaacggcc 8100  
 cggaggggtg cgggcaggac gcccgcata aactgccagg catcaaatta agcagaagc 8160

catcctgacg gatggccttt ttgcgtttct acaaactctt ctagagtga c 8211

<210> 5  
 <211> 47  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer

<400> 5  
 ccattccatg gtgtcttgac aattaatcat cggctcgtat aatgtg 47

<210> 6  
 <211> 39  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer

<400> 6  
 gggatcgata tgatatctcc tgtgtgaaat tgttatccg 39

<210> 7  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220><223> primer  
 <400> 7

gggctcgaga gcctgggggtg cctaagag 29

<210> 8  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence



<220><223> primer  
 <400> 8  
 gacacatgg aatggtgcaa aacc 24  
 <210> 9  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 9  
 gcgcgcgat ccattgctctc gacaataaaa atgatac 36  
 <210> 10  
 <211> 49  
  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 10  
 caggtacctt agtgggtgtg gtgggtgtga atgtaatagt ccggatcga 49  
 <210> 11  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 11  
 gcgcgatcc atgggcaatg tgtacaaca 30  
 <210> 12  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 12  
 cagaggtacc ttagtgggtg tgatgatgat ggtcgcggat gtcttcgta 49  
  
 <210> 13  
 <211> 32

<212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 13  
 gcgcgcgat ccatgattgg ctatggttta cc 32  
 <210> 14  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 14  
 cgcggtacce taatgatgat gatgatgat gggtagcaga gtaagtgg 48  
 <210> 15  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 15  
 ctaaaaggag atatagatgg gcaatgtgta caacaacatc accgaaacc 49  
  
 <210> 16  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 16  
 catctatate tccttttagt ggtggtggtg gtggtgaatg taatagtcc 49  
 <210> 17  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 17  
 aaaaggagat atagatgatt ggctatggtt tacctatgcc caatcaggc 49  
 <210> 18

<211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 18  
 catctatata tccttttagt ggtggtggtg gtggtgaatg taatagtcc 49

<210> 19  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 19  
 aaaaggagat atagatgatt ggctatggtt tacctatgcc caatcaggc 49

<210> 20  
 <211> 49  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 20  
 gccaatcacc tatatctcct tttagtgggtg gtgatgatga tggtcgagg 49

<210> 21  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 21  
 aatcctcgaa gacaccgacg gcaac 25

<210> 22  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> primer  
 <400> 22

gttgccgtcg gtgtcttcga ggatt

25