



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0048145
(43) 공개일자 2025년04월07일

(51) 국제특허분류(Int. Cl.)
C12N 15/81 (2006.01) C12R 1/84 (2006.01)
(52) CPC특허분류
C12N 15/815 (2013.01)
C12N 2830/002 (2013.01)
(21) 출원번호 10-2025-7009989(분할)
(22) 출원일자(국제) 2016년05월11일
심사청구일자 없음
(62) 원출원 특허 10-2023-7006299
원출원일자(국제) 2016년05월11일
심사청구일자 2023년03월24일
(85) 번역문제출일자 2025년03월26일
(86) 국제출원번호 PCT/US2016/031797
(87) 국제공개번호 WO 2016/183163
국제공개일자 2016년11월17일
(30) 우선권주장
62/159,899 2015년05월11일 미국(US)
(뒷면에 계속)

(71) 출원인
임파서블 푸즈 인크.
미국 94063 캘리포니아 레드우드 시티 새기노 드
라이브 400
(72) 발명자
샨카르, 스미타
미국 94030 캘리포니아주 밀브레 아파트먼트 115
에스. 엘 카미노 리얼 120
호이트, 마틴 앤드류
미국 94132 캘리포니아주 샌프란시스코 가필드 스
트리트 338
(74) 대리인
양영준, 김영

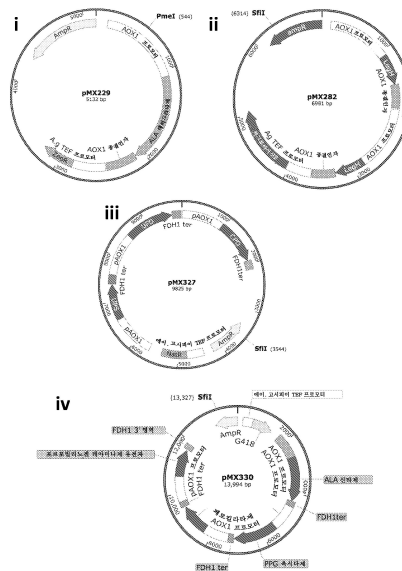
전체 청구항 수 : 총 1 항

(54) 발명의 명칭 메틸영양성 효모를 유전적으로 조작하는 발현 구축물 및 방법

(57) 요약

메틸영양성 효모를 유전적으로 조작하기 위한 방법 및 물질이 제공된다.

대표도 - 도2



(52) CPC특허분류
C12R 2001/84 (2021.05)

(30) 우선권주장
62/183,074 2015년06월22일 미국(US)
62/185,921 2015년06월29일 미국(US)
62/203,052 2015년08월10일 미국(US)
62/220,366 2015년09월18일 미국(US)
62/222,388 2015년09월23일 미국(US)
62/236,506 2015년10월02일 미국(US)
62/313,491 2016년03월25일 미국(US)

명세서

청구범위

청구항 1

효모 세포의 용도.

발명의 설명

기술 분야

- [0001] 일반적으로 본 개시내용은 메틸영양성 효모를 유전적으로 조작하기 위한 DNA 구축물 및 이러한 DNA 구축물을 사용하는 방법에 관한 것이다.

배경 기술

- [0002] 메틸영양성 효모 예컨대 피키아 파스토리스(*Pichia pastoris*)가 재조합 단백질의 발현에 통상적으로 사용된다. 메틸영양성 효모에서 하나 이상의 폴리펩티드를 효율적으로 발현시키는데 사용될 수 있는 구축물이 본원에서 제공된다.

발명의 내용

- [0003] 본 개시내용은 AOX1 프로모터로부터 전사 활성화제인 Mxr1을 과다발현하는 피. 파스토리스(*P. pastoris*) 균주를 사용하여 또한 AOX1 프로모터로부터 발현되는 트랜스진의 발현을 증가시키는 것을 개시하며, 이는 하나 이상의 단백질의 재조합 생산을 유의하게 개선한다. 추가적으로, AOX1 프로모터로부터의 Mxr1의 발현은 억제성 탄소원이 고갈될 때, 일반적으로 절대적인 유도제인 메탄올의 부재 하에 AOX1 프로모터로부터의 다른 트랜스진의 발현을 허용하는 양성 피드백 루프를 생성시킨다. Mxr1의 발현은 생산되는 단백질의 양의 유의한 증가를 초래한다.
- [0004] 한 측면에서, 재조합 핵산 분자를 포함하는 메틸영양성 효모 세포가 제공된다. 전형적으로 이러한 재조합 핵산 분자는 적어도 하나의 메탄올-유도성 프로모터 요소에 작동가능하게 연결된 전사 활성화제를 코딩하는 외인성 핵산을 포함한다. 대표적인 메틸영양성 효모는 칸디다(*Candida*), 한세놀라(*Hansenula*), 피키아(*Pichia*) 및 톨루프시스(*Torulopsis*) 속의 것일 수 있다. 대표적인 메틸영양성 효모는 피키아 파스토리스이다.
- [0005] 일부 실시양태에서, 재조합 핵산 분자는 메틸영양성 효모 세포의 게놈 내로 안정적으로 통합된다. 일부 실시양태에서, 재조합 핵산 분자는 복제-적격 플라스미드로부터 염색체외적으로 발현된다.
- [0006] 일부 실시양태에서, 전사 활성화제를 코딩하는 외인성 핵산은 피키아 파스토리스로부터의 Mxr1 서열, 한세놀라 폴리모르파(*Hansenula polymorpha*)로부터의 Adr1 서열, 칸디다 보이디니이(*Candida boidinii*)로부터의 Trm1 서열, 및 칸디다 보이디니이로부터의 Trm2 서열을 포함한다. 전사 활성화제를 코딩하는 대표적인 핵산이 DQ395124에서 제시된다. 대표적인 전사 활성화제는 ABD57365에 제시된 아미노산 서열을 갖는다.
- [0007] 일부 실시양태에서, 적어도 하나의 메탄올-유도성 프로모터 요소는 피키아 파스토리스로부터의 알콜 옥시다제 1 (AOX1) 프로모터 요소, 칸디다 보이디니이로부터의 AOD1 프로모터 요소, 한세놀라 폴리모르파로부터의 MOX 프로모터 요소, 피키아 메타놀리카(*Pichia methanolica*)로부터의 MOD1 프로모터 요소, 피키아 파스토리스로부터의 DHAS 프로모터 요소, 피키아 파스토리스로부터의 FLD1 프로모터 요소, 또는 피키아 파스토리스로부터의 PEX8 프로모터 요소이다.
- [0008] 일부 실시양태에서, 메틸영양성 효모 세포는 적어도 하나의 메탄올-유도성 프로모터 요소에 작동가능하게 연결된 폴리펩티드를 코딩하는 적어도 하나의 이중성 핵산을 포함하는 핵산 분자를 추가로 포함한다. 일부 실시양태에서, 적어도 하나의 이중성 핵산은 철 보조인자 예컨대 헴의 생합성에서 수반되는 적어도 하나의 폴리펩티드 (예를 들어, ALA 신타제, ALA 데히드라타제, 포르포빌리노겐 데아미나제, UPG III 신타제, UPG III 데카르복실라제, CPG 옥시다제, PPG 옥시다제, 및/또는 페로킬라타제)를 코딩한다. 일부 실시양태에서, 철 보조인자 생합성에서 수반되는 폴리펩티드 중 하나 이상이 적어도 하나의 메탄올-유도성 프로모터 요소에 연결된다.
- [0009] 또 다른 측면에서, 세포에서 이중성 폴리펩티드를 발현시키는 방법이 제공된다. 전형적으로 이러한 방법은 본

원에 기술된 바와 같은 메틸영양성 효모 세포를 제공하는 단계; 재조합 핵산 분자를 메틸영양성 효모 세포 내로 도입하며, 이러한 재조합 핵산 분자는 적어도 하나의 피키아 파스토리스 알콜 옥시다제 1 (AOX1) 프로모터 요소에 작동가능하게 연결된 폴리펩티드를 코딩하는 적어도 하나의 이중성 핵산을 포함하는 것인 단계; 및 세포를 재조합 핵산 분자의 발현에 적합한 조건 하에 배양하며, 이에 의해 이중성 폴리펩티드를 발현시키는 단계를 포함한다.

[0010] 일부 실시양태에서, 세포가 배양되는 조건은 철 또는 그의 제약상 또는 대사상 허용되는 염의 첨가를 포함한다. 일부 실시양태에서, 도입 단계는 형질도입, 전기천공, 바이오리틱 입자 전달 또는 화학적 형질전환과 같은 기술을 포함한다. 일부 실시양태에서, 배양 단계는 세포를 메탄올의 존재 하에 배양하는 것을 포함한다.

[0011] 또 다른 측면에서, 활성화시키는 프로모터에 작동가능하게 연결된 전사 활성화제를 포함하는 재조합 생물이 제공된다. 일부 실시양태에서, 이러한 프로모터에 작동가능하게 연결된 폴리펩티드를 발현하는 재조합 생물이 제공된다. 또 다른 측면에서, 본원에 기술된 바와 같이 유도제의 첨가 없이 유도성 프로모터로부터 폴리펩티드를 발현시키는 방법이 제공된다.

[0012] 달리 정의되지 않는 한, 본원에서 사용된 모든 기술 용어 및 과학 용어는 사안 방법 및 조성물이 속하는 분야의 통상의 기술자가 통상적으로 이해하는 바와 동일한 의미를 갖는다. 본원에 기술된 것과 유사하거나 등가인 방법 및 물질이 사안 방법 및 조성물의 실행 또는 테스트에서 사용될 수 있지만, 적합한 방법 및 물질이 하기에 기술된다. 추가적으로, 물질, 방법 및 예는 예시적일 뿐이고, 제한적인 것으로 의도되지 않는다. 본원에서 언급된 모든 발표물, 특허 출원, 특허 및 다른 참고문헌은 전문이 참조로 포함된다.

도면의 간단한 설명

[0013] 도 1은 헴 생합성 경로에서 수반되는 단계를 도시하는 개략도이다.

도 2는 생산 균주 MXY0183의 구축에서 사용된 플라스미드의 개략도이다.

도 3은 모 균주 Bg11로부터의 생산 균주 MXY0183 및 MXY0207의 생성을 나타내는 개략도이다.

도 4는 플라스미드 pGAB 및 pMx354를 나타내는 개략도이다.

도 5는 모 균주 Bg11로부터의 항생제 선택이 없는 생산 균주 MXY0291 및 MXY0338의 생성을 나타내는 개략도이다.

도 6은 생산 균주 MXY0291을 제조하기 위해 공동-형질전환에 의해 도입된 Mxr1 및 LegH 변이체 3을 함유하는 선형 DNA 조각의 개략도이다.

도 7은 피키아의 천연 비-pAOX1 구성적 프로모터의 제어 하에 LegH를 발현하는 선형 구축물을 나타내는 개략도이다.

도 8은 균주 MXY0183과 연관된 표현형 변화를 나타내는 사진이다. 유도 시작 (0 hr) 및 유도 72 hr 후의 진탕 플라스크가 제시된다. 1, MXY0051; 2, MXY0118; 3, MXY0183.

도 9는 변형된 피. 파스토리스 균주로부터의 LegH 생산을 나타낸다. 패널 A는 진탕 플라스크에서 성장된 피. 파스토리스 균주로부터의 용해물을 나타내는 SDS 겔이다: 51, MXY0051; 118, MXY0118; 183, MXY0183. 패널 B는 균주 MXY0118, MXY0183, 및 MXY0207로부터의 LegH 생산을 비교하는 표이다.

도 10은 균주 MXY0206으로의 실험으로부터의 데이터를 나타낸다. 패널 A는 억제성 탄소원에서의 48 hr의 성장 후의 균주 MXY0183 (왼쪽) 및 MXY0206 (오른쪽)의 진탕 플라스크 배양물의 사진이다. 패널 B는 BMY 배지에서의 48 hr의 성장 후의 균주 MXY0183 (왼쪽) 및 MXY0206 (오른쪽)의 진탕 플라스크 배양물로부터의 세포 펠렛의 사진이다. 패널 C는 (임의의 유도제의 부재 하에서의) 헴-로딩된 LegH의 상대적인 수율을 나타내는 그래프이다.

도 11은 2 L 발효기 탱크에서 메탄올과 글리세롤 또는 메탄올과 글루코스의 존재 하에서 성장되었을 때의 본원에 기술된 균주의 상대적인 수율을 나타내는 요약 표이다.

발명을 실시하기 위한 구체적인 내용

[0014] 폴리펩티드의 재조합 발현을 증가시키도록 세포를 유전적으로 조작하는 것을 허용하는 핵산 구축물이 본원에서 제공된다. 일부 실시양태에서, 유도 분자의 부재 하에 유도성 프로모터로부터의 폴리펩티드의 재조합 발현을 증가시키도록 세포를 유전적으로 조작하는 것을 허용하는 핵산 구축물이 본원에서 제공된다. 어떠한 특정 메커

니즘에도 제한되지 않으면서, 본원에 기술된 방법은 전사 활성화제의 낮은 수준의 천연 발현이 전사 활성화제에 작동가능하게 연결된 프로모터를 유도하는 양성 피드백 루프를 생성시킨다. 이는 전사 활성화제 뿐만 아니라 동일한 유도성 프로모터에 작동가능하게 연결된 하나 이상의 표적 폴리펩티드의 발현 증가에 이른다.

- [0015] 메틸영양성 효모 세포를 유전적으로 조작하는 것을 허용하는 핵산 구조물이 본원에서 제공된다. 본원에서 피키아 종 (즉, 피. 파스토리스)을 사용하여 방법이 예시되지만, 피키아 속의 다른 종, 또는 칸디다, 한세놀라, 피키아 및 토룰롭시스 속 중 임의의 것으로부터의 종을 사용할 수 있다.
- [0016] 메틸영양성 효모 세포를 유전적으로 조작하는 것은 세포 내로 재조합 핵산 분자를 도입하는 것을 전형적으로 포함한다. 본원에 기술된 바와 같이, 재조합 핵산 분자는 적어도 하나의 유도성 프로모터 요소에 작동가능하게 연결된 전사 활성화제를 코딩하는 외인성 핵산을 전형적으로 포함한다.
- [0017] 본원에 기술된 방법에서 사용되는 재조합 핵산 분자는 전형적으로 DNA지만, 적절한 환경 하에 RNA 분자를 사용할 수 있다. 본원에서 사용된 바와 같이, "외인성"은 외부 공급원으로부터 세포의 게놈 내로 도입된 임의의 핵산 서열을 지칭하고, 여기서 외부 공급원은 동일하거나 상이한 생물 또는 합성으로 생성된 핵산일 수 있다. 예를 들어, 외인성 핵산은 하나의 미생물 (예를 들어, 하나의 메틸영양성 효모 속 또는 종)으로부터의 핵산이 상이한 속 또는 종의 메틸영양성 효모 내로 도입된 핵산일 수 있지만, 외인성 핵산은 상응하는 천연 핵산 서열이 존재함에도 불구하고 추가적인 카피로서 메틸영양성 효모 내로 재조합으로 도입된 메틸영양성 효모로부터의 핵산일 수도 있다. 예를 들어, 피. 파스토리스는 Mxr1 전사 활성화제를 코딩하는 내인성 핵산을 함유한다; 추가적인 피. 파스토리스 Mxr1 핵산 (예를 들어, 피. 파스토리스 내로 재조합으로 도입됨) 또는 내인성 피. 파스토리스 Mxr1 핵산을 변형시키는 것이 외인성으로 간주된다.
- [0018] 전사 활성화제, 및 전사 활성화제를 코딩하는 핵산 (예를 들어, 전사 활성화제를 코딩하는 외인성 핵산)이 관련 기술분야에 공지되어 있다. 예를 들어, 피키아 파스토리스로부터의 전사 활성화제는 Mxr1 서열이지만, 적합한 전사 활성화제를 한세놀라 폴리모르파 (Adr1 서열; 예를 들어, 핵산 서열은 진뱅크(GenBank) 수탁 번호 AE0102000005의 염기 858873 내지 862352, 아미노산 서열은 진뱅크 수탁 번호 ESX01253 참조) 및 칸디다 보이디니이 (Trm1 서열; 예를 들어, 핵산 서열은 진뱅크 수탁 번호 AB365355, 아미노산 서열은 진뱅크 수탁 번호 BAF99700을 참조한다; Trm2 서열; 예를 들어, 핵산 서열은 진뱅크 수탁 번호 AB548760, 아미노산 서열은 진뱅크 수탁 번호 BAJ07608 참조)에서 또한 발견할 수 있다. 대표적인 피. 파스토리스 Mxr1 핵산 서열을, 예를 들어, 진뱅크 수탁 번호 DQ395124에서 확인할 수 있는 한편, 대표적인 피. 파스토리스 Mxr1 폴리펩티드 서열을, 예를 들어, 진뱅크 수탁 번호 ABD57365에서 확인할 수 있다.
- [0019] 전사 활성화제 예컨대 Mxr1은 정상적으로는 낮은 수준으로 발현될 수 있다. 따라서, 외인성 핵산 (즉, 전사 활성화제)를 유도성인 프로모터의 제어 하에 놓는 것이 바람직하다. 본원에서 사용된 바와 같이, "작동가능하게 연결된"은 프로모터 또는 다른 발현 요소(들)가 핵산 코딩 서열에 대해 이러한 핵산의 발현을 지시하거나 조절하는 방식으로 (예를 들어, 인-프레임(in-frame)으로) 위치하는 것을 의미한다.
- [0020] 메틸영양성 효모를 유전적으로 조작할 때 사용할 수 있는 다수의 유도성 프로모터가 있다. 예를 들어, 메탄올-유도성 프로모터, 또는 이로부터의 프로모터 요소를 사용할 수 있다. 관련 기술분야에 메탄올 유도성 프로모터가 공지되어 있다. 예를 들어, 피. 파스토리스로부터의 통상적으로 사용되는 메탄올-유도성 프로모터는 메탄올에 반응하여 강하게 전사되는 알콜 옥시다제 1 (AOX1) 유전자로부터의 프로모터 또는 그의 일부분이다. 그러나, 칸디다 보이디니이로부터의 알콜 옥시다제 (AOD1) 프로모터 (예를 들어, 진뱅크 수탁 번호 YSAAOD1A 참조), 한세놀라 폴리모르파로부터의 알콜 옥시다제 (MOX) 프로모터 (예를 들어, 진뱅크 수탁 번호 X02425 참조), 피키아 메타놀리카로부터의 MOD1 또는 MOD2 프로모터 (예를 들어, 문헌 [Raymond et al., 1998, Yeast, 14:11-23]; 및 [Nakagawa et al., 1999, Yeast, 15:1223-30] 참조), 피. 파스토리스로부터의 DHAS 프로모터 (예를 들어, 진뱅크 수탁 번호 FJ752551 참조) 또는 이로부터의 프로모터 요소, 피키아 파스토리스로부터의 포름알데히드 데히드로게나제 (FLD1) 프로모터 (예를 들어, 진뱅크 수탁 번호 AF066054 참조), 또는 피. 파스토리스로부터의 PEX8 프로모터 (예를 들어, 문헌 [Kranthi et al., 2010, Yeast, 27:705-11] 참조)를 비제한적으로 포함하여, 다른 메탄올-유도성 프로모터 또는 이로부터의 프로모터 요소를 사용할 수 있다. 일부 실시양태에서, 전사 활성화제는 피키아 파스토리스로부터의 Mit1 서열 (예를 들어, 진뱅크 수탁 번호 CAY70887 참조)이다. 이들 프로모터 모두가 메탄올에 의해 유도되는 것으로 공지되어 있다.
- [0021] 통상의 기술자는 본원에 기술된 재조합 핵산 분자가 메틸영양성 효모 세포의 게놈 내로 안정적으로 통합될 수 있거나, 또는 복제-적격 플라스미드로부터 염색체외적으로 발현될 수 있다는 것을 이해할 것이다. 둘 다를 달성하는 방법이 관련 기술분야에 널리 공지되어 있고, 일상적으로 사용된다.

- [0022] 본원에서 증명된 바와 같이, 피키아 내의 메탄올-조절 전사 활성화제가 AOX1 프로모터에 결합할 수 있고, Mxr1 과 협동으로 작용하여 AOX1 프로모터로부터의 전사를 활성화시킬 수 있다. 일부 실시양태에서, 2개의 메탄올-조절 전사 활성화제 (예를 들어, Mxr1 및 Mit1)가 메탄올 유도성 프로모터 요소에 작동가능하게 연결될 수 있다.
- [0023] 본원에 기술된 바와 같은 재조합 핵산 분자를 포함하는 균주를 메틸영양성 효모 세포 내에서 제2 재조합 핵산 분자를 조절하는데 (예를 들어, 과발현시키는데) 사용할 수 있다. 제2 재조합 핵산 분자는, 예를 들어, 하나 이상의 관심 폴리펩티드를 코딩하는 하나 이상의 이중성 핵산을 포함할 수 있다. 전사 활성화제를 코딩하는 외인성 핵산과 유사하게, 이중성 핵산은 계놈에 대해 천연이 아니거나 생물의 계놈 내에 있지 않은 임의의 핵산 서열을 지칭한다 (예를 들어, 이중성 핵산은 하나의 미생물 (예를 들어, 하나의 메틸영양성 효모 속 또는 종)으로부터의 핵산이 상이한 속 또는 종의 메틸영양성 효모 내로 도입된 핵산일 수 있음).
- [0024] 단순히 예로서, 하나 이상의 관심 폴리펩티드를 코딩하는 이중성 핵산은 헴-보조인자 생합성에서 수반되는 핵산 일 수 있다. 피키아 파스토리스 계놈의 서열로부터 결정되고 주해된 바와 같은 헴 생합성에서 수반되는 8개의 상이한 효소를 코딩하는 핵산이 본원에서 예시된다. 예를 들어, ALA 신타제, ALA 데히드라타제, 포르포빌리노겐 데아미나제, UPG III 신타제, UPG III 데카르복실라제, CPG 옥시다제, PPG 옥시다제, 및 페로킬라타제를 코딩하는 이중성 핵산이 본원에 기술된 메틸영양성 효모 균주에서 발현될 수 있다. 하나 초과와 이중성 핵산 (예를 들어, 트랜스진)을 함유하도록 메틸영양성 효모를 유전적으로 조작하기 위해, 메탄올-유도성 및 구성적 프로모터 또는 이로부터의 요소의 조합물을 조합하여 이러한 핵산의 발현을 추가로 증가시킬 수 있다.
- [0025] 사카로미세스 세레비시아에(*Saccharomyces cerevisiae*)에서의 기존의 연구에서, ALA 데히드라타제 및 포르포빌리노겐 데아미나제가 헴 생합성에서의 속도 제한 효소로서 확인되었다 (예를 들어, 문헌 [Hoffman et al., 2003, Biochem. Biophys. Res. Commun., 310(4):1247-53] 참조). 그러나, 글리세르알데히드-3-포스페이트 데히드로게나제 (GAP) 프로모터로부터의 피. 파스토리스에서의 개별적인 헴 효소의 이중성 발현은 헴 보조인자를 함유하는 재조합 단백질의 발현과 연관된 한계를 극복하지 못했다 (문헌 [Krainer et al., 2015, Microb. Cell Fact., 13:14:4] 참조). 본원에 기술된 바와 같이, 피. 파스토리스에서의 재조합 헴 함유 단백질의 고도로 효율적인 발현이 메탄올-유도성 프로모터로부터 전체 헴 생합성 경로를 공동-발현시킴으로써 달성되었지만, 헴 생합성 경로에서 수반되는 유전자 중 하나 이상이 하나 이상의 구성적 프로모터로부터 발현될 수 있다는 것이 이해될 것이다.
- [0026] 철-보조인자 생합성에서 수반되는 효소에 더하여, 식물 헤모글로빈을 포함하는 글로빈 단백질 패밀리에 (Pfam 데이터베이스 내의 PF00042)의 구성원을 코딩하는 핵산이 존재할 수 있다는 것이 이해될 것이다. 본원에서의 실시예에서, 대두 레그헤모글로빈 (LegH)을 코딩하는 핵산이 존재한다. LegH는 철 보조인자인 헴에 결합하는 단백질이고, 이는 415 nm에서의 특징적인 흡수 및 독특한 적색을 조래한다. LegH 단백질 (일명 LGB2)은 대두의 뿌리혹에서 자연적으로 발견되고 (예를 들어, 유니프록KB(UniprotKB) 수탁 번호 P02236 참조), 본원에서 사용된 핵산 서열은 피. 파스토리스에서의 발현에 대해 코돈이 최적화되었다. 예를 들어, WO 2014/110539 및 WO 2014/110532를 참조한다.
- [0027] 대안적으로, 관심 폴리펩티드를 코딩하는 이중성 핵산은, 예를 들어 비제한적으로, 데히드린, 피타제, 프로테아제, 카탈라제, 리파제, 피옥시다제, 아밀라제, 트랜스글루타미나제, 옥시도리덕타제, 트랜스퍼라제, 히드롤라제, 리아제, 이소머라제, 리가제, 또는 임의의 이러한 폴리펩티드에 대한 항체일 수 있다. 다른 실시양태에서, 이중성 핵산은 소형 분자, 예컨대 에탄올, 락트산, 부탄올, 아디프산 또는 숙신산의 생산을 위한 경로에서 수반되는 하나 이상의 효소를 코딩할 수 있다.
- [0028] 전사 활성화제를 코딩하는 외인성 핵산과 유사하게, 관심 폴리펩티드를 코딩하는 이중성 핵산이 유도성 프로모터 요소 (예를 들어, 메탄올-유도성 프로모터 요소)에 작동가능하게 연결될 수 있거나, 또는 관심 폴리펩티드를 코딩하는 이중성 핵산이 구성적 프로모터 또는 구성적 프로모터 요소에 작동가능하게 연결될 수 있다. 유도성 프로모터 및 이로부터의 요소는 상기에 논의되어 있다. 구성적 프로모터 및 구성적 프로모터 요소는 관련 기술 분야에 공지되어 있다. 예를 들어, 피. 파스토리스로부터의 통상적으로 사용되는 구성적 프로모터는 구성적 방식으로 강하게 전사되는 전사 신장 인자 EF-1 α 유전자 (TEF1)로부터의 프로모터 또는 그의 일부분이다. 그러나, 피. 파스토리스로부터의 글리세르알데히드-3-포스페이트 데히드로게나제 (GAPDH) 프로모터 (예를 들어, 진뱅크 수탁 번호 U62648.1 참조), 피. 파스토리스로부터의 잠재적 글리코실 포스포타일 이노시톨 (GPI)-앵커 단백질인 GCW14p (PAS_chr1-4_0586)로부터의 프로모터 (예를 들어, 진뱅크 수탁 번호 XM_002490678 참조), 또는 피. 파스토리스로부터의 3-포스포글리세레이트 키나제 유전자 (PGK1)로부터의 프로모터 (예를 들어, 진뱅크 수

탁 번호 AY288296 참조)를 비제한적으로 포함하여, 다른 구성적 프로모터 또는 이로부터의 프로모터 요소를 사용할 수 있다.

[0029] 본원에 기술된 재조합 핵산 분자와 유사하게, 제2 재조합 핵산 분자는 메틸영양성 효모 세포의 게놈 내로 안정적으로 통합될 수 있거나, 또는 복제-적격 플라스미드로부터 염색체외적으로 발현될 수 있다.

[0030] 통상의 기술자는 유도성 (예를 들어, 메탄올-유도성) 및 구성적 프로모터 (또는 이로부터의 프로모터 요소)의 조합물을 조합하여 이에 작동가능하게 연결된 임의의 핵산의 발현을 추가로 증가시킬 수 있다는 것을 이해할 것이다.

[0031] 통상의 기술자는 프로모터 요소에 작동가능하게 연결된 관심 폴리펩티드를 코딩하는 이중성 핵산이 본원에 기술된 재조합 핵산 분자로부터 분리될 수 있거나, 또는 본원에 기술된 재조합 핵산 분자 내에 함유된 프로모터 요소에 작동가능하게 연결된 전사 활성화제를 코딩하는 외인성 핵산에 인접할 수 있다는 것을 이해할 것이다. 통상의 기술자는 제2 핵산 분자가 본원에 기술된 재조합 핵산 분자에 인접하면, 단일 프로모터 또는 이로부터의 프로모터 요소가 양쪽 또는 모든 유전자 (예를 들어, 전사 활성화제를 코딩하는 외인성 핵산, 뿐만 아니라 관심 폴리펩티드(들)를 코딩하는 하나 이상의 이중성 핵산)의 전사를 구동시키는데 사용될 수 있다는 것을 또한 이해할 것이다.

[0032] 핵산을 메틸영양성 효모 세포 내로 도입하는 방법이 관련 기술분야에 공지되어 있고, 비제한적으로, 형질도입, 전기천공, 바이오폴리머 입자 전달, 및 화학적 형질전환을 포함한다.

[0033] 추가적으로, 메틸영양성 효모 세포를 배양하는 방법이 관련 기술분야에 공지되어 있다. 예를 들어, 문헌 [Pichia Protocols, *Methods In Molecular Biology*, 389, Cregg, Ed., 2007, 2nd Ed., Humana Press, Inc.]을 참조한다. 일부 환경 하에, 메탄올을 배양 배지에 도입 또는 첨가하는 것이 바람직할 수 있지만, 본원에서 증명된 바와 같이, 하나 이상의 관심 폴리펩티드의 높은 수준의 효율적인 발현을 수득하기 위해 메탄올이 요구되지 않는다. 일부 환경 하에 (예를 들어, 철-보조인자 생합성에서 수반되는 효소(들)를 코딩하는 하나 이상의 핵산이 발현되는 경우), 배양 배지에 철 또는 그의 제약상 또는 대사상 허용되는 (또는 GRAS) 염을 보충하는 것이 바람직할 수 있다.

[0034] 피키아 균주는 단독 탄소원으로서의 메탄올 상에서 성장할 수 있다. 알콜 옥시다제의 작용에 의해 메탄올을 포름알데히드로 전환하는 것에 의해 메탄올 이용이 시작된다. 메틸영양성 효모인 피키아 파스토리스는 알콜 옥시다제에 대한 2개의 유전자인 AOX1 및 AOX2를 함유한다. 알콜 옥시다제 활성이 감소된 균주 ("메탄올 이용 느림" 또는 MutS 균주)는 종종 알콜 옥시다제 활성이 감소되지 않은 균주보다 AOX1 프로모터로부터 발현된 재조합 단백질을 더 많이 생산한다. 양쪽 AOX 유전자에서 돌연변이되고 알콜 옥시다제 활성이 완전히 결여된 균주는 메탄올을 대사할 수 없지만, 여전히 메탄올에 의해 AOX1 프로모터로부터의 발현에 대해 유도될 수 있다. 이들 균주는 다른 탄소원을 성장에 이용하는 능력을 유지하지만, 여전히 메탄올 첨가 시 AOX1 프로모터로부터 이중성 단백질을 발현한다. 이들 균주는 메탄올을 대사하지 않기 때문에 ("메탄올 이용 마이너스" 또는 Mut- 균주), 훨씬 더 적은 메탄올이 단백질 발현 유도에 요구되고, 이러한 돌연변이를 보유하는 균주는 대규모 발효에서 메탄올 공급에 관련된 문제를 피한다. 예를 들어, 문헌 [Chiruvolu et al., 1997, *Enzyme Microb. Technol.*, 21:277-83]을 참조한다. Mut- 균주에서의 AOX1 프로모터로부터의 LegH 발현이 LegH 수율을 크게 개선한 것으로 본원에서 결정되었다. 따라서, AOX1 유전자 및 AOX2 유전자 양쪽 모두에서 돌연변이가 있는 메틸영양성 효모가 본원에 기술된 방법에서 사용될 수 있다.

[0035] 관심 단백질, 또는 하나 이상의 관심 단백질을 포함하는 복합체 (예를 들어, 헴-결합된 LegH, 데히드린, 피타제, 프로테아제, 카탈라제, 리파제, 퍼옥시다제, 아밀라제, 트랜스글루타미나제, 옥시도리덕타제, 트랜스퍼라제, 히드롤라제, 리아제, 이소머라제, 리가제, 또는 항체)를 이러한 효모 세포로부터 정제할 수 있다. 폴리펩티드를 정제하는 방법이 관련 기술분야에 공지되어 있다. 본원에서 사용된 바와 같이, "정제된" 폴리펩티드는 이에 천연적으로 동반되는 세포 성분으로부터 분리 또는 정제된 폴리펩티드이다. 전형적으로, 폴리펩티드는 자신이 천연적으로 회합되는 폴리펩티드 또는 천연 발생 분자가 건조 중량 기준으로 적어도 70% (예를 들어, 적어도 75%, 80%, 85%, 90%, 95%, 또는 99%) 없으면 "정제된" 것으로 간주된다. 화학적으로 합성된 폴리펩티드는 이에 천연적으로 동반되는 성분으로부터 선천적으로 분리되기 때문에, 합성 폴리펩티드는 "정제된" 것이다.

[0036] 본원에서 사용된 바와 같이, 핵산은 DNA 및 RNA를 포함할 수 있고, 하나 이상의 뉴클레오타이드 유사체 또는 백본 변형을 함유하는 핵산을 포함한다. 핵산은 단일 가닥 또는 이중 가닥일 수 있고, 이는 일반적으로 그의 의도되는 용도에 좌우된다. 소정의 서열과 상이한 핵산 및 폴리펩티드가 또한 제공된다. 핵산 및 폴리펩티드는 소정

의 핵산 또는 폴리펩티드 서열에 대한 적어도 50%의 서열 동일성 (예를 들어, 적어도 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%의 서열 동일성)을 가질 수 있다.

[0037] 퍼센트 서열 동일성의 계산에서, 2개의 서열을 정렬하고, 2개의 서열 사이의 동일하게 매칭되는 뉴클레오티드 또는 아미노산 잔기의 수를 결정한다. 동일한 매칭의 수를 정렬된 영역의 길이 (즉, 정렬된 뉴클레오티드 또는 아미노산 잔기의 수)로 나누고 100을 곱하여 퍼센트 서열 동일성 값에 도달한다. 정렬된 영역의 길이가 하나의 서열 또는 둘 다의 서열의 일부분에서 최단 서열의 전장 크기까지일 수 있다는 것이 이해될 것이다. 단일 서열이 하나 초과인 다른 서열과 정렬될 수 있고, 따라서 각각의 정렬된 영역에 걸쳐 상이한 퍼센트 서열 동일성 값을 가질 수 있다는 것이 또한 이해될 것이다.

[0038] 퍼센트 서열 동일성을 결정하기 위해 2종 이상의 서열을 정렬하는 것은 컴퓨터 프로그램 ClustalW 및 디폴트 파라미터를 사용하여 수행할 수 있고, 이는 핵산 또는 폴리펩티드 서열의 정렬이 이들의 전체 길이에 걸쳐 수행되는 것을 허용한다 (전체 정렬). 문헌 [Chenna et al., 2003, Nucleic Acids Res., 31(13):3497-500]. ClustalW는 질의 서열과 하나 이상의 대상 서열 사이의 최상의 매칭을 계산하고, 동일성, 유사성 및 차이가 결정될 수 있도록 이들을 정렬한다. 서열 정렬을 최대화하도록 하나 이상의 잔기의 갭이 질의 서열, 대상 서열, 또는 양쪽 내로 삽입될 수 있다. 핵산 서열을 신속하게 쌍으로 정렬하기 위해, 디폴트 파라미터를 사용할 수 있고 (즉, 워드 크기: 2; 윈도우 크기: 4; 채점 방법: 백분율; 상부 대각선의 수: 4; 및 갭 페널티: 5); 다중 핵산 서열의 정렬을 위해, 하기 파라미터를 사용할 수 있다: 갭 개방 페널티: 10.0; 갭 연장 페널티: 5.0; 및 가중치 변환: 예. 폴리펩티드 서열을 신속하게 쌍으로 정렬하기 위해, 하기 파라미터를 사용할 수 있다: 워드 크기: 1; 윈도우 크기: 5; 채점 방법: 백분율; 상부 대각선의 수: 5; 및 갭 페널티: 3. 폴리펩티드 서열의 다중 정렬을 위해, 하기 파라미터를 사용할 수 있다: 가중치 행렬: 블로썸; 갭 개방 페널티: 10.0; 갭 연장 페널티: 0.05; 친수성 갭: 커짐; 친수성 잔기: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg, 및 Lys; 및 잔기-특이적 갭 페널티: 커짐. 예를 들어, 월드와이드웹 상의 베이러 의과 대학의 서치 런처(Baylor College of Medicine Search Launcher) 웹사이트 또는 유럽 생물정보학 연구소(European Bioinformatics Institute) 웹사이트에서 ClustalW를 실행할 수 있다.

[0039] 변화를 핵산 분자 내로 도입함으로써, 코딩된 폴리펩티드의 아미노산 서열의 변화를 일으킬 수 있다. 예를 들어, 돌연변이유발 (예를 들어, 부위-지정 돌연변이유발, PCR-매개 돌연변이유발)을 사용하여 또는 변화가 있는 핵산 분자를 화학적으로 합성함으로써, 핵산 코딩 서열 내로 변화를 도입할 수 있다. 이러한 핵산 변화는 하나 이상의 아미노산 잔기에서의 보존적 및/또는 비-보존적 아미노산 치환에 이를 수 있다. "보존적 아미노산 치환"은 하나의 아미노산 잔기가 유사한 측쇄가 있는 상이한 아미노산 잔기로 교체되는 치환이고 (예를 들어, 문헌 [Dayhoff et al., 1978, Atlas of Protein Sequence and Structure, 5(Suppl. 3):345-352]을 참조하고, 이는 아미노산 치환에 대한 빈도 표를 제공함), 비-보존적 치환은 아미노산 잔기가 유사한 측쇄가 없는 아미노산 잔기로 교체되는 치환이다. 발현 (예를 들어, 전사 및/또는 번역) 증가, 더 엄격한 조절, 탈조절, 이화대사 산물 억제의 상실, 변형된 특이성, 분비, 열안정성, 용매 안정성, 산화 안정성, 프로테아제 저항성, 촉매 활성 및/또는 색상을 비제한적으로 포함하는 하나 이상의 성질을 개선하도록 핵산 및/또는 폴리펩티드 서열이 본원에 기술된 바와 같이 변형될 수 있다.

[0040] 본원에서 사용된 바와 같이, "단리된" 핵산 분자는 단리된 핵산 분자가 유래되는 생물의 게놈 (예를 들어, PCR 또는 제한 엔도뉴클레아제 소화에 의해 생산된 cDNA 또는 게놈 DNA 단편) 내에서 천연적으로 핵산의 한쪽 또는 양쪽 끝부분의 측면에 있는 서열이 없는 핵산 분자이다. 일반적으로 이러한 단리된 핵산 분자는 조작 편의를 위해 또는 하기에 더욱 상세하게 논의된 융합 핵산 분자를 생성시키기 위해 벡터 (예를 들어, 클로닝 벡터, 또는 발현 벡터) 내로 도입된다. 추가적으로, 단리된 핵산 분자는 조작된 핵산 분자 예컨대 재조합 또는 합성 핵산 분자를 포함할 수 있다.

[0041] 관련 기술분야에서 일상적인 기술을 사용하여 핵산을 단리할 수 있다. 예를 들어, 재조합 핵산 기술, 및/또는 폴리머라제 연쇄 반응 (PCR)을 비제한적으로 포함하는 임의의 방법을 사용하여 핵산을 단리할 수 있다. 일반적인 PCR 기술이, 예를 들어, 문헌 [PCR Primer: A Laboratory Manual, Dieffenbach & Dveksler, Eds., Cold Spring Harbor Laboratory Press, 1995]에 기술되어 있다. 재조합 핵산 기술은, 예를 들어, 제한 효소 소화 및 라이게이션을 포함하고, 이를 사용하여 핵산을 단리할 수 있다. 단리된 핵산은 단일 핵산 분자로서 또는 일련의 올리고뉴클레오티드로서 화학적으로 합성될 수도 있다.

[0042] 공지된 방법 예컨대 DEAE 이온 교환, 겔 여과, 및 히드록시아파타이트 크로마토그래피에 의해 천연 공급원 (예

를 들어, 생물학적 샘플)으로부터 폴리펩티드를 정제할 수 있다. 또한, 예를 들어, 발현 벡터 내의 핵산을 발현시킴으로써 폴리펩티드를 정제할 수 있다. 추가적으로, 정제된 폴리펩티드를 화학적 합성에 의해 수득할 수 있다. 임의의 적절한 방법, 예를 들어, 칼럼 크로마토그래피, 폴리아크릴아미드 겔 전기영동, 또는 HPLC 분석을 사용하여 폴리펩티드 순도의 정도를 측정할 수 있다.

[0043] 핵산 (예를 들어, 폴리펩티드를 코딩하는 핵산)을 함유하는 구축물 또는 벡터가 또한 제공된다. 발현 구축물 또는 벡터를 포함하는 구축물 또는 벡터는 시판되거나, 또는 관련 기술분야에서 일상적인 재조합 DNA 기술에 의해 생산될 수 있다. 핵산을 함유하는 구축물 또는 벡터는 이러한 핵산에 작동가능하게 연결된 발현 요소를 가질 수 있고, 선택성 마커 (예를 들어, 항생제 저항성 유전자)를 코딩하는 서열과 같은 서열을 추가로 포함할 수 있다. 핵산을 함유하는 구축물 또는 벡터는 키메라 또는 융합 폴리펩티드 (즉, 폴리펩티드의 N-말단 또는 C-말단 중 하나에 있을 수 있는 이중성 폴리펩티드에 작동적으로 연결된 폴리펩티드)를 코딩할 수 있다. 대표적인 이중성 폴리펩티드는 코딩된 폴리펩티드의 정제에서 사용될 수 있는 것 (예를 들어, 6xHis 태그, 글루타티온 S-트랜스퍼라제 (GST))이다.

[0044] 발현 요소는 핵산 코딩 서열의 발현을 지시 및 조절하는 핵산 서열을 포함한다. 발현 요소의 한 예는 프로모터 서열이다. 발현 요소는 인트론, 인핸서 서열, 반응 요소, 또는 핵산 발현을 조정하는 유도성 요소를 또한 포함할 수 있다. 발현 요소는 박테리아, 효모, 곤충, 포유동물 또는 바이러스 기원의 것일 수 있고, 벡터는 기원이 상이한 요소의 조합물을 함유할 수 있다.

[0045] 본원에 기술된 바와 같은 벡터는 숙주 세포 내로 도입될 수 있다. 본원에서 사용된 바와 같이, "숙주 세포"는 핵산이 도입된 특정 세포를 지칭하고, 벡터를 보유하는 이러한 세포의 자손을 또한 포함한다. 숙주 세포는 임의의 원핵생물 또는 진핵생물 세포일 수 있다. 예를 들어, 핵산이 박테리아 세포 예컨대 이. 콜라이(*E. coli*), 또는 곤충 세포, 효모 또는 포유동물 세포 (예컨대 차이나이즈 햄스터 난소 세포 (CHO) 또는 COS 세포)에서 발현될 수 있다. 다른 적합한 숙주 세포가 관련 기술분야의 통상의 기술자에게 공지되어 있다. 생체 내 및 시험관 내 둘 다에서 핵산을 숙주 세포 내로 도입하는 다수의 방법이 관련 기술분야의 통상의 기술자에게 널리 공지되어 있고, 전기천공, 인산칼슘 침전, 폴리에틸렌 글리콜 (PEG) 형질전환, 열 충격, 리포펙션(lipofection), 미세주입, 및 바이러스-매개 핵산 전달을 비제한적으로 포함한다.

[0046] 수많은 증폭 기술 (예를 들어, 문헌 [PCR Primer: A Laboratory Manual, 1995, Dieffenbach & Dveksler, Eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY]; 및 미국 특허 번호 4,683,195; 4,683,202; 4,800,159; 및 4,965,188 참조)을 적절한 올리고뉴클레오티드 (예를 들어, 프라이머) 쌍과 함께 사용하여 핵산을 검출할 수 있다. 원본 PCR에 대한 다수의 변형이 개발되었고, 핵산을 검출하는데 사용될 수 있다.

[0047] 혼성화를 사용하여 핵산을 검출할 수도 있다. 핵산들 사이의 혼성화가 문헌 [Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY]의 섹션 7.37-7.57, 9.47-9.57, 11.7-11.8, 및 11.45-11.57에 상세하게 논의되어 있다. 뉴클레오티드 약 100개 미만의 올리고뉴클레오티드 프로브에 대한 적합한 서던 블롯 조건이 문헌 [Sambrook et al.]에 개시되어 있다 (섹션 11.45-11.46). 길이가 뉴클레오티드 100개 미만인 서열과 제2 서열 사이의 Tm을 섹션 11.46에서 제공된 식을 사용하여 계산할 수 있다. 문헌 [Sambrook et al.]에는 뉴클레오티드 약 100개 초과인 올리고뉴클레오티드 프로브에 대한 서던 블롯 조건이 추가적으로 개시되어 있다 (섹션 9.47-9.54 참조). 길이가 뉴클레오티드 100개 초과인 서열과 제2 서열 사이의 Tm을 문헌 [Sambrook et al.]의 섹션 9.50-9.51에서 제공된 식을 사용하여 계산할 수 있다.

[0048] 핵산을 함유하는 막이 예비혼성화 및 혼성화되는 조건, 뿐만 아니라 핵산을 함유하는 막을 세정하여 과량의 비-특이적으로 결합된 프로브를 제거하는 조건이 혼성화의 엄격성에서 중요한 역할을 할 수 있다. 이러한 혼성화 및 세정을, 적절한 경우에, 중도 또는 고도의 엄격성 조건 하에 수행할 수 있다. 예를 들어, 세정 용액 내의 염 농도를 감소시킴으로써 및/또는 세정이 수행되는 온도를 증가시킴으로써 세정 조건을 더욱 엄격하게 만들 수 있다. 단순한 예로서, 고도의 엄격성 조건은 65°C에서 0.2X SSC에서 막을 세정하는 것을 전형적으로 포함한다.

[0049] 추가적으로, 혼성화의 양을 해석하는 것이, 예를 들어, 표지된 올리고뉴클레오티드 프로브의 특이적 활성화, 프로브가 혼성화된 주형 핵산 상의 프로브-결합 부위의 수, 및 자가방사선사진 또는 다른 검출 매질에 대한 노출량에 의해 영향을 받을 수 있다. 관련 기술분야의 통상의 기술자는 수많은 혼성화 및 세정 조건을 고정된 표적 핵산에 대한 프로브 핵산 분자의 혼성화를 검사하는데 사용할 수 있지만, 동일한 혼성화, 세정, 및 노출 조건 하에 표적 핵산에 대한 프로브의 혼성화를 검사하는 것이 더욱 중요하다는 것을 쉽게 이해할 것이다. 바람직한 예는, 표적 핵산이 동일한 막 상에 있다.

[0050] 한 핵산에 대한 혼성화가 또 다른 핵산에 대한 혼성화보다 적어도 5배 (예를 들어, 적어도 6배, 7배, 8배, 9배, 10배, 20배, 50배, 또는 100배) 더 크면 핵산 분자가 한 핵산에 혼성화하지만 또 다른 핵산에는 혼성화하지 않는 것으로 간주된다. 혼성화의 양을 직접적으로 막 상에서 또는 자가방사선사진으로부터, 예를 들어, 포스포이미저(PhosphorImager) 또는 덴서토미터(Densitometer) (몰레큘러 다이나믹스(Molecular Dynamics), 캘리포니아주 서니베일)를 사용하여 정량할 수 있다.

[0051] 항체를 사용하여 폴리펩티드를 검출할 수 있다. 항체를 사용하여 폴리펩티드를 검출하는 기술은 효소 결합 면역흡착 검정법 (ELISA), 웨스턴 블롯, 면역침전 및 면역형광을 포함한다. 항체는 폴리클로날 또는 모노클로날일 수 있다. 폴리펩티드에 대한 특이적 결합 친화성이 있는 항체를 관련 기술분야에 널리 공지된 방법을 사용하여 생성시킬 수 있다. 이러한 항체를 관련 기술분야에 공지된 방법을 사용하여 고체 지지체 예컨대 미량역가 플레이트에 부착시킬 수 있다. 폴리펩티드의 존재 하에, 항체-폴리펩티드 복합체가 형성된다.

[0052] 일반적으로 검출 (예를 들어, 증폭 생성물, 혼성화 복합체, 또는 폴리펩티드의 검출)은 검출가능한 표지를 사용하여 달성된다. 용어 "표지"는 직접적인 표지, 뿐만 아니라 간접적인 표지의 사용을 포괄하도록 의도된다. 검출가능한 표지는 효소, 치환기, 형광 물질, 발광 물질, 생물발광 물질 및 방사성 물질을 포함한다.

[0053] 선택을 위한 서열이 결합된 (즉, 선택성 마커가 결합된) 균주를 생성시키는데 사용될 수 있는 방법이 본원에서 기술된다. 이들 방법은 원형 플라스미드 DNA 벡터 및 선형 DNA 서열을 사용하는 것을 포함하고, 여기서 원형 플라스미드 DNA 벡터는 선택 마커 및 DNA 복제 기원 (일명 자가 복제 서열 (ARS))을 함유하고, 선형 DNA 서열은 동종 재조합에 의한 피키아 게놈 내로의 통합을 위한 서열을 함유한다. 선형 DNA 분자는 하나 이상의 관심 단백질, 예컨대 비제한적으로 렘-결합된 LegH, 데히드린, 피타제, 프로테아제, 카탈라제, 리파제, 퍼옥시다제, 아밀라제, 트랜스글루타미나제, 옥시도리덕타제, 트랜스퍼라제, 히드롤라제, 리아제, 이소머라제, 리가제, 소형 분자, 예컨대 에탄올, 락트산, 부탄올, 아디프산 또는 숙신산의 생산을 위한 경로에서 수반되는 하나 이상의 효소, 또는 임의의 이러한 단백질에 대한 항체를 코딩하는 핵산 서열을 추가적으로 포함할 수 있다.

[0054] 피키아 세포를 양쪽 DNA 분자로 형질전환시킬 수 있고, 원형 플라스미드 상의 선택성 마커의 존재에 의해 형질 전환체를 선택할 수 있다. 그 후, 예를 들어, PCR을 사용하여, 게놈 내로의 선형 DNA 분자의 통합에 대해 형질 전환체를 스크리닝할 수 있다. 마커가 없는 선형 DNA 분자가 올바르게 통합된 형질전환체가 확인되면, 세포를 원형 플라스미드에 대한 선택의 부재 하에 성장시킬 수 있다. 선택의 부재 하에 마커-보유 플라스미드가 안정적으로 유지되지 않기 때문에, 선택이 완료된 후에 이러한 플라스미드가 상실되고, 종종 매우 신속하게 상실된다. 생성된 균주는 통합된 선형 DNA를 선택을 위한 이중성 서열의 부재 하에 보유한다. 따라서, 이러한 접근법은 재조합 단백질 수율에 영향을 거의 미치지 않거나 미치지 않으면서 선택성 마커 (예를 들어, 이중성 선택 마커)가 결합된 피키아 균주를 구축하는데 사용될 수 있다.

[0055] 본 발명에 따라, 관련 기술분야의 기술 내의 통상적인 분자생물학, 미생물, 생화학, 및 재조합 DNA 기술을 사용할 수 있다. 이러한 기술은 문헌에 충분히 설명되어 있다. 본 발명이 하기 실시예에서 추가로 기술될 것이고, 이는 특허청구범위에 기술된 사안의 방법 및 조성물의 범주를 제한하지 않는다.

[0056] 실시예

[0057] 파트 A. 물질 및 방법

[0058] 실시예 1 - 폴리머라제 연쇄 반응

[0059] 퓨전 하이-피델리티(Phusion Hi-fidelity) DNA 폴리머라제 (뉴 잉글랜드 바이오랩스(New England Biolabs))를 사용하여 게놈 DNA 또는 플라스미드 DNA 주형으로부터 관심 유전자를 증폭시켰다. 간략하게, 1-2 U의 퓨전 DNA 폴리머라제의 존재 하에 각각 0.6 μ M의 정방향 및 역방향 프라이머를 10-50 ng의 주형 DNA 및 400 μ M의 뉴클레오티드 믹스와 함께 인큐베이션한다. 반응 조건은 하기와 같았다:

1 사이클	최초 변성	98°C	1분
25 사이클	변성	98°C	10초
	어닐링		20초
	연장	72°C	kb당 30초
1 사이클	최종 연장	72°C	5분
1 사이클	유지	4°C	계속

[0060]

- [0061] 실시예 2 - 라이게이션에 의한 플라스미드 구축
- [0062] 50-100 ng의 제한 효소로 소화된 플라스미드 및 3X 몰 과량의 PCR로 증폭된 삽입물을 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스)의 존재 하에 인큐베이션하였다. 16℃에서 2 hr 초과 동안 라이게이션을 수행하였다. 2 μ l의 라이게이션 반응물을 DH10B 전기천공적격 이. 콜라이 세포 내로 형질전환시켰다.
- [0063] 실시예 3 - 이. 콜라이 일렉트로맥스(ElectroMax) DH10B T1 파지-저항성 적격 세포 내로의 형질전환
- [0064] 1.5-2 μ l의 라이게이션 혼합물을 1 mm 갭 큐벳 (바이오래드(BioRad), 카탈로그 # 165-2089)을 사용하여 1.7 kV로 설정된 마이크로펄서(MicroPulser) (바이오래드)를 사용한 전기천공에 의해 20 μ l의 일렉트로맥스 DH10B T1 파지-저항성 적격 세포 (인비트로젠(Invitrogen), 카탈로그 # 12033-015) 내로 형질전환시켰다; 펄스 후, 1 ml SOC를 세포에 첨가하고, 200 rpm으로 진탕시키면서 세포를 37℃에서 1시간 동안 인큐베이션하였다. 10 μ l의 회수 혼합물을 100 μ g/ml의 농도로 앰피실린을 함유하는 LB 한천 플레이트 상에 플레이팅하였다. 플레이트를 37℃에서 인큐베이션하였다.
- [0065] 실시예 4 - 피. 파스토리스 내로의 형질전환을 위한 플라스미드 DNA의 선형화
- [0066] 플라스미드 DNA를 1-4시간 동안 37℃에서 1x 컷스마트 버퍼(CutSmart Buffer) 내의 PmeI 제한 엔도뉴클레아제 (뉴 잉글랜드 바이오랩스, 카탈로그 # R0560L)로 또는 1-4시간 동안 50℃에서 1x 컷스마트 버퍼 내의 SfiI 제한 엔도뉴클레아제 (뉴 잉글랜드 바이오랩스, 카탈로그 # R0123L)로 소화시켰다. 선형화된 플라스미드를 자이모클린 젤 DNA 회수 키트(Zymoclean Gel DNA Recovery Kit) (자이모 리서치(Zymo Research), 카탈로그 # D4002)를 사용하여 0.8% 아가로스 겔로부터 겔 정제하였다. DNA를 20 μ l H₂O에 용출시켰다.
- [0067] 실시예 5 - 피. 파스토리스 형질전환-적격 세포의 제조
- [0068] 선택된 피. 파스토리스 균주를 25 ml YPD 배지에서 지수 성장 중기 (~2 OD)까지 성장시켰다. 930 x g로 15분 동안 원심분리하여 세포를 수집하였다. 세포 펠릿을 80% YPD 및 200 mM HEPES, pH 6.8의 용액 2 ml에 재현탁시켰다. 75 μ l의 1 M DTT를 첨가하였다. 재현탁된 세포 펠릿을 100 rpm으로 30℃에서 25분 동안 혼합시켰다. 40 ml 부피의 빙냉 멸균수를 현탁액에 첨가하고, 1125 x g로 15분 동안 원심분리하여 세포를 수집하고, 얼음 상에 놓았다. 2회의 추가적인 세정 단계에 대해 이전과 같이 세포 펠릿을 40 ml 빙냉수에 재현탁시키고 수집하였다. 그 후, 세포 펠릿을 20 ml의 빙냉 1 M 소르비톨에 재현탁시키고, 이전과 같이 원심분리에 의해 수집하였다. 최종 세포 펠릿을 0.3 ml 빙냉, 멸균 소르비톨에 현탁시키고, 분취하고, -80℃에서 동결시켰다.
- [0069] 실시예 6 - 피. 파스토리스 내로의 형질전환
- [0070] 30-100 ng의 선형화된 플라스미드 DNA를 1 mm 갭 진펄서(GenePulser) 큐벳 (바이오래드)과 1.15 kV로 설정된 진펄서 (바이오래드)를 사용하여 30 μ l의 전기천공적격 피. 파스토리스 세포 내로 형질전환시켰다. 1 ml의 YPD/1 M 소르비톨을 1:1 비로 세포에 첨가하여 혼합시켰다. 100 rpm으로 진탕하면서 세포를 30℃에서 3시간 동안 인큐베이션하였다. 100 μ l의 회수 혼합물을 적절한 항생제를 함유하는 YPD 플레이트 상에 플레이팅하고, 나머지 세포를 적절한 항생제가 있는 YPD 플레이트 상에 플레이팅하였다. 플레이트를 30℃에서 48시간 동안 인큐베이션하였다. 1차 형질전환 플레이트를 적절한 항생제가 있는 YPD 플레이트 상에 스트리킹하고, 플레이트를 48시간 동안 30℃에서 인큐베이션하였다. 개별적인 클론을 항생제가 있는 YPD 플레이트 상에 패칭하고, 이러한 패치를 염색체 내로의 통합을 증명하기 위해 콜로니 PCR 또는 gDNA 프래그를 행하는데 또는 추가 분석을 위해 균주를 진탕 플라스크에서 성장시키는데 사용하였다.
- [0071] 실시예 7 - LegH 생산을 위한 진탕 플라스크에서의 배양물 성장
- [0072] 신선한 패치로부터의 균주를 성장 배지 BMGY (0.75% 글리세롤이 보충된 BMY) 내로 접종하고, 200 rpm으로 진탕시키면서 30℃에서 철야로 성장시켰다. 다음날, 철야 배양물을 0.1 mM 암모늄 Fe(III) 시트레이트가 보충된 BMMY 배지 (BMY+ 1% 메탄올)로 희석함으로써 메탄올로 LegH 발현을 유도하였다. 배양물을 0.5-0.7의 OD600으로 성장시켰다. 소포제를 0.01%의 최종 농도로 첨가하였다. 배양물을 총 72시간 동안 성장시켰다; 진탕 플라스크 부피의 1/10의 10x BMMY 배지 (BMY+10% 메탄올)를 첨가함으로써 배양물에 24시간 마다 메탄올을 보충하였다. 72시간의 성장 유도 후에 세포를 원심분리로 수확하였다.
- [0073] 실시예 8 - 진탕 플라스크 배지
- [0074] 10 g의 효모 추출물 및 20 g 소이톤을 790 ml 물에 용해시킴으로써 BMY 배지를 제조하였다. 혼합물을 오토클레이빙으로 멸균하고, 실온으로 냉각시켰다. 100 ml의 1 M 인산칼륨 완충제 (pH 6.0) 및 100 ml의 10X 아미노산

무함유 효모 질소 베이스 (100 mL당 13.4 g의 YNB 분말; 시그마-알드리치(Sigma-Aldrich))를 필터 멸균하고 (0.2 μ m 구멍 크기의 PES), 배지에 첨가하였다. pH 조정이 요구되지 않는다.

[0075] BMY 배지 성분

성분	양, 1L당
효모 추출물	10 g
대두 펩톤 (BD)	20 g
아미노산 무함유 효모 질소 베이스, 10X 용액	100 mL (BMY 내의 13.4 g/L를 생성함)
1M 인산칼륨 완충제, pH 6.0	100 mL

[0076] 최종 성분을 물에 용해시키고, 오토클레이빙하여 멸균하였다.

[0077] 진탕 플라스크용 저-오스몰농도 배지

성분	양, g/L
황산암모늄	15.7
일염기성 인산칼륨	9.4
황산칼슘 2수화물	0.43
황산마그네슘 7수화물	11.7
시트르산나트륨 2수화물	1.13

[0079] 실시예 9 - 발효 배지 및 공급원료

[0080] 하기 지시된 성분을 용해시키고, 물로 부피를 조정하였다. 성분은 FCC 식품 등급 또는 등가물이었다. 오토클레이빙에 의해, 정치 증기에 의해, 또는 등가물로 배지를 멸균하였다.

[0081] 95 g/L 글리세롤 함유 발효용 저-오스몰농도 배지

성분	양, g/L
황산암모늄	15.7
일염기성 인산칼륨	9.4
황산칼슘 2수화물	0.43
황산마그네슘 7수화물	11.7
시트르산나트륨 2수화물	1.13
글리세롤, USP 등급 99.7%	95

[0083] 멸균 후, 배지를 실온으로 냉각되게 하고, 하기를 첨가하였다:

성분	양, mL/L
미량 금속 PTM1 용액	2
비타민 용액	4
시그마 204 소포제 또는 등가물	1

[0084] 미량 금속 PTM1 용액은 분말 믹스로서 선라이즈 사이언스(Sunrise Science) (카탈로그 번호 4052-A-B-1L)로부터 입수가능하다. 파우치 A 및 파우치 B를 950 mL 물에서 혼합하고, 5 mL 황산을 첨가하였다. 혼합 시 약간의 침전이 예상된다; 혼합물을 필터 멸균하고 (0.2 μ m 구멍 크기의 PES), 4°C의 암실에서 보관하였다.

[0087] 비타민 용액 레시피

성분	양, g/L
비오틴	0.2
판토텐산칼슘	1
폴산	0.2
이노시톨	1
니아신	0.2
p-아미노벤조산	0.2
페리독신 히드로클로라이드	1
리보플라빈	0.5
티아민 히드로클로라이드	1
B12	0.1

[0088]

[0089] 대안적으로, 미량 금속 PTM1을 하기와 같이 제조할 수 있다:

성분	양
황산구리 - 5H ₂ O	6.0 g
아이오딘화나트륨	0.08 g
황산망간 - H ₂ O	3.0 g
몰리브덴산나트륨- 2H ₂ O	0.2 g
붕산	0.02 g
염화코발트	0.5 g
염화아연	0.5 g
황산철 - 7H ₂ O	65.0 g
비오틴	0.2 g
황산	5.0 ml
물	최종 부피 1 L가 되도록

[0090]

[0091] 성분을 함께 혼합하고, 필터 멸균하고, 실온에서 보관한다. 17.5 g의 앰버펄(AmberFerm) 4000을 320 mL 물 내로 혼합하고 교반하여 용해시킴으로써 글리세롤 공급원료 믹스를 제조하였다. 물-앰버펄 혼합물을 850 g의 글리세롤에 첨가하고, 격렬하게 교반함으로써 잘 혼합시켰다. 공급원료 믹스를 오토클레이빙으로 멸균하였다.

[0092]

글리세롤 공급원료 용액

성분	양, g/L
USP 등급 글리세롤	850
물	320
센시엔트 앰버펄4000 대두 가수분해물	17.5

[0093]

[0094] 12 mL/L의 PTM1 용액이 보충된 99-100% 메탄올을 사용하여 메탄올 공급원료를 제조하였다.

[0095] 실시예 10 - 실험실 규모의 고-산소 전달 발효를 위한 프로토콜

[0096] 시드 진탕 플라스크 프로토콜

[0097] 무균성 생물안전 후드에서, 저-오스몰농도 배지 및 BMV를 9:1 저-오스몰농도:BMV 비로 혼합하였다. 12.5 g/L 농도의 글리세롤을 배지에 첨가하였다. USP 식품 등급 글리세롤/글리세린 (50% v/v (63% w/w) 글리세롤/물 용액 내의 99.7% 순도)을 사용하였고, 오토클레이빙하여 멸균하였다. 시그마 204 또는 등가의 소포제를 0.25 mL/L의 농도로 배지에 첨가하였다. 글리세롤 시드 바이알을 회수하고, 70% IPA 또는 에탄올을 외부에 분무하고, 실온에서 약 5분 동안 생물안전 후드 내부에서 해동시켰다. 배플형 진탕 플라스크에 글리세롤 시드 바이알을 접종하였다; 1 L의 진탕 플라스크 배지마다 1 mL의 접종물 바이알을 사용하였다. 진탕 (1" 쓰로와 함께 200 RPM)하면서 배양물을 24시간 동안 30°C에서 성장시켰다. 1:10 내지 1:5의 실제 배지 부피 : 공칭 진탕 플라스크 부피 비를 사용하였다. 2.8 L 공칭 부피 플라스크와 250 내지 500 mL의 배지가 일상적으로 성공적으

로 사용되었다. 24시간의 성장 후 600 nm에서의 OD를 측정하였다; OD가 15 이상이면, 배양물을 발효기에 접종하는데 사용하였다. OD가 15 미만이면, 배양물을 1-2시간 더 성장시킨 후, OD를 다시 측정하였다. 15 내지 30 시간 후에 OD가 15에 도달하지 않았으면, 시드 플라스크를 실패한 것으로 간주하였다.

[0098] 발효 프로토콜

[0099] 발효 배지 및 공급원료를 본원에 기술된 바와 같이 제조하였다. 초기 부피는 최대 발효기 부피의 약 40%이어야 하고, 예를 들어, 발효기의 최대 작업 부피가 10 L이면, 4 L이어야 한다. 이는 발효가 끝날 때 공정이 최대 작업 부피에 접근할 것이기 때문이다. 10% 겔중물-발효기 비로 발효기에 진탕 플라스크 시드를 접종하고, 예를 들어, 4 L의 초기 배지가 발효기 내에 존재하면, 발효기에 약 0.4 L의 진탕 플라스크 시드를 접종한다. 이러한 시점의 발효기 내의 총부피를 T0 부피로 지칭하고, 예를 들어, 이러한 대표적인 예에서는 4.4 L이다. 공정 제어는 하기를 포함한다: 30°C 온도; 20% 포화 설정점을 유지하도록 교반-통기 케이스케이드에 의해 용존 산소를 제어함; 및 28% NH₄OH 첨가를 통해 pH를 제어하고, 설정점은 공정의 단계에 좌우됨.

[0100] 뱃치 단계 (접종에서 글리세롤 고갈까지이고, DO 급증에 의해 신호됨): 용존 산소에 대한 PID 제어의 반응성에 따라, 세포에서 배지 내에 존재하는 글리세롤이 고갈될 때 강한 DO 급증 또는 교반-통기 속도의 빠른 하락 또는 둘 다의 조합이 관찰될 수 있다. 이것이 발생할 때 페드-뱃치 단계가 시작된다. 뱃치 단계의 기간은 대략 20 시간이지만, 24시간까지 허용가능한 것으로 간주된다. pH 설정점은 5.0이다. 뱃치 단계 끝의 습식 세포 중량은 대략 220 g/L일 것이다.

[0101] 페드-뱃치 단계: T0 부피를 기초로 12-14 g/L/hr의 순수 글리세롤에 도달하도록 글리세롤 공급을 개시한다. 대략 350 g/L의 습식 세포 중량에 도달했을 때까지 공급속도를 유지하였고, 이는 약 7-10시간이 걸릴 것이다. pH 설정점은 5.0이다.

[0102] 전이 단계: 샘플을 취한 후 전이 단계를 시작한다. 발효 브로쓰에서 1-2 g/L 메탄올 농도에 도달했을 때까지, T0 부피를 기초로 1 g/L/hr의 순수 메탄올에 도달하도록 메탄올 공급을 개시하였다. 브로쓰 내의 0.25-1 g/L의 메탄올 농도를 유지하도록 나머지 발효 동안 메탄올 공급 속도를 조정하였다. 2시간 과정에 걸쳐 선형으로, T0 부피를 기초로 12-14 g/L/hr에서 8-9 g/L/hr의 순수 글리세롤로 글리세롤 공급 속도를 감소시켰다. 20분마다 공급 속도를 단계적으로 감소시키는 것이 또한 허용가능한 것이다. pH 설정점을 3.5로 변화시키고, 자연적으로 새로운 설정점으로 맞춰지도록 발효를 허용하였다 (즉, 산을 첨가하지 않음).

[0103] 생산 단계 (글리세롤 공급의 경사 감소 종료에서 발효 종료까지): pH 설정점은 3.5였다. 발효 브로쓰에서 0.25-1 g/L의 메탄올 농도를 유지하였다. 글리세롤의 공급 속도를 T0 부피를 기초로 8-9 g/L/hr의 순수 글리세롤로 유지하였다. 대략 12시간마다 샘플을 취했다. 4°C에서 4000 내지 7000 RCF로 샘플을 스피닝하고, 상청액을 경사분리하였다. 상청액을 별개의 튜브에 저장하였다. 각각의 시점의 5 mL의 상청액 샘플 3개 및 펠릿을 -80°C에서 동결시켰다. 용기에 대한 최대 통기 및 교반 속도에서도 생산 동안의 15-20% DO를 유지할 수 없으면, 글리세롤 공급 속도를 T0 부피를 기초로 5 g/L/hr의 순수 글리세롤까지 저하시킬 수 있다. 접종 60시간 후에 발효가 종료되었다. 1000 L 규모에서, 수확 공정은 공급원료 및 통기를 중단하고, 브로쓰를 8°C로 냉각하고, 샴플스 또는 디스크 스택 원심분리를 사용하여 페이스트를 농축하는 것으로 이루어졌다. 수확은 일반적으로 약 5-10시간이 걸리고, 생성물 품질의 검출가능한 손실을 초래하지 않는다. 실험실 규모의 경우, 3x5 mL 샘플에 더하여, 추가적인 50 mL 샘플을 종료 시에 수집하는 것이 충분하다. 습식 세포 중량은 >450 g/L이었고, 스피닝된 펠릿은 더 하얀색으로 보이는 유도 전 샘플로부터의 스피닝된 샘플과 대조적으로 분홍색으로 보였다. 약 6-12시간의 유도 후에 브로쓰의 색이 백색에서 더욱 뚜렷한 분홍색으로 변하기 시작하였다.

[0104] **파트 B. 생산 균주 구축**

[0105] **생산 균주 MX0183**

[0106] 실시에 11 - 헵 생합성 경로의 각각의 효소의 pGAN 또는 pGAZ 통합 벡터 내로의 클로닝

[0107] pGAN (nat 선택 마커가 있음) 및 pGAZ (제오신 선택 마커가 있음)를 바이오그래매틱스, 인크.(Biogrammatix, Inc.) (캘리포니아주 칼스배드)로부터 구입하였다. 각각의 유전자는 AOX1 프로모터의 제어 하에 놓였고, FDH 종결인자가 각각의 유전자의 정지 코돈 직후에 놓였다. 헵 생합성 경로 내의 유전자를 야생형 피. 파스토리스 균주로부터 PCR 증폭시키거나 또는 이전 구축물로부터 서브클로닝하였다.

[0108] 헵 생합성에서 수반되는 효소를 포함하는 헵 생합성 경로가 도 1에서 제시된다. 헵 생합성 동안 생산되는 중간체가 상자에서 제시되고, 각각의 단계를 촉매하는 효소가 오른쪽에 제시된다. 에스. 세레비시아에(S.

cerevisiae)에서 나타난 바와 같은 속도 제한 효소 단계가 밀줄이 그어져 제시된다.

[0109] ALA 신타제, ALA 데히드라타제, UPGIII 신타제, UPGIII 데카르복실라제, CPG 옥시다제 및 PPG 옥시다제 유전자를 제한 엔도뉴클레아제인 BsaI에 의한 인식을 위한 부위를 함유하는 프라이머로 PCR 증폭시켰다. 엘림바이오팜(ElimBiopharm)에 의해 올리고뉴클레오티드가 합성되었다.

프라이머 명칭	유전자	서열	SEQ ID NO:
MxO0187	ALAsynth_F	GAGGGTCTCGGATGGAGTTTGTGCGCCCGTC	19
MxO0188	ALAsynth_R	GAGGGTCTCGATTACAATCTGACTCCTGATGAGG	20
MxO0189	ALAdehyd_F	GAGGGTCTCGGATGGTGCATAAGGCTGAATACTTG	21
MxO0190	ALAdehyd_R	GAGGGTCTCGATTATTCAGATAACCACTCCAGG	22
MxO0191	UroporSynth_F	GAGGGTCTCGGATGCCAAAAGCCATTCTTCTGAAG	23
MxO0192	UroporSynth_R	GAGGGTCTCGATTAGTGCACCTTTTGTATAGAC	24
MxO0193	UroporDecarb_F	GAGGGTCTCGGATGAGTAGATTTCCAGAACTGAAG	25
MxO0194	UroporDecarb_R	GAGGGTCTCGATTATTGAGATCCAATGCG	26
MxO0195	CoproOx_F	GAGGGTCTCGGATGGCCATCGACTCTGATATC	27
MxO0196	CoproOx_R	GAGGGTCTCGATTATACCCATTCAATAGGAT	28
MxO0197	ProtoporOx_F	GAGGGTCTCGGATGCTGAAAAGTCTTGCACCAAA	29
MxO0198	ProtoporOx_R	GAGGGTCTCGATTAAATGCCACTGAGGGTAGC	30

[0110]

[0111] 퓨전 하이-피델리티 DNA 폴리머라제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0530L)를 사용하여 게놈 DNA로부터 유전자를 증폭시켰다. PCR 생성물을 수득하고, DNA 클린&컨센트레이터(DNA Clean&Concentrator)-5 (카탈로그 # D4004)를 사용하여 정제하고, DNA를 25 μ l의 H₂O에 용출시켰다. 벡터 DNA, pGAZ 및 pGAN, 및 PCR 생성물을 50 μ l 반응 부피로 37°C에서 BsaI (뉴 잉글랜드 바이오랩스, 카탈로그 # R0535S)로 소화시켰다.

[0112]

선형화된 벡터 및 소화된 PCR 생성물을 자이모클린 겔 DNA 회수 키트 (자이모 리서치 카탈로그 # D4002)를 사용하여 0.8% 아가로스 겔로부터 정제하였다. DNA를 20 μ l H₂O에 용출시켰다. T4 DNA 리가제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0202S)를 사용하여 철야로 16°C에서 10 μ l로 라이게이션 반응을 설정하였다.

[0113]

PBD 및 페로킬라타제 유전자를 이전에 구축된 플라스미드로부터 서브클로닝하였다: pJAZ_PBD를 37°C에서 5분 동안 1x 패스트 다이제스트(Fast Digest) 완충제 내의 BstBI(Bsp119I) (썬모사이언티픽(ThermoScientific), FD0124) 및 NotI (썬모사이언티픽, FD0596)로 소화시켰다. pJAZ_Ferroch를 37°C에서 5분 동안 1x 패스트 다이제스트 완충제 내의 MfeI (MunI, 썬모사이언티픽, FD0753) 및 NotI (썬모사이언티픽, FD0596)로 소화시켰다.

[0114]

소화된 생성물을 자이모클린 겔 DNA 회수 키트 (자이모 리서치 카탈로그 # D4002)를 사용하여 0.8% 아가로스 겔로부터 정제하였다. DNA를 20 μ l H₂O에 용출시켰다.

[0115]

T4 DNA 리가제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0202S)를 사용하여 철야로 16°C에서 10 μ l로 라이게이션 반응을 설정하였다.

[0116]

1.5 μ l의 라이게이션 혼합물을 1.7 kV로 설정된 마이크로펄스 (바이오래드)를 사용한 전기천공에 의해 20 μ l의 일렉트로맥스 DH10B T1 파지-저항성 적격 세포 (인비트로젠, 카탈로그 # 12033-015) 내로 형질전환시켰다; 200 rpm으로 진탕시키면서 세포를 37°C에서 1시간 동안 1 ml SOC 내에서 인큐베이션하였다. 10 μ l의 회수 혼합물을 100 μ g/ml의 농도로 앰피실린을 함유하는 LB 한천 플레이트 상에 플레이팅하였다. 플레이트를 철야로 37°C에서 인큐베이션하였다. 삽입물의 존재에 대해 콜로니 PCR에 의해 콜로니를 스크리닝하였다. 유전자의 서열을 확인하였다.

명칭	구조물	유전자
pMx0308	pGAN-ALAsynth	ALA 신타제
pMx0309	pGAN-ALAD	ALAD
pMx0310	pGAN-UPGIIIsyn	우로포르피리노겐 신타제
pMx0311	pGAN-UPGIIIdecarb	우로포르피리노겐 데카르복실라제
pMx0312	pGAN-CPGoxi	CPG 옥시다제
pMx0313	pGAN-PPGoxi	프로토포르피린 옥시다제
pMx0314	pGAZ-ALAsyn	ALA 신타제
pMx0315	pGAZ-ALAD	ALAD
pMx0316	pGAZ-UPGIIIsyn	우로포르피리노겐 신타제
pMx0317	pGAZ-UPGIII데카르복실라제	UPGIII 데카르복실라제
pMx0318	pGAZ-PPG옥시다제	PPG 옥시다제
pMx0319	pGAZ-CPG 옥시다제	CPG 옥시다제
pMx0320	pGAN-PBD	PBD
pMx0321	pGAZ-PGC	PBD
pMx0322	pGAZ-Fc	페로킬라타제
pMx0323	pGAN-Fc	페로킬라타제

[0117]

[0118] 실시예 12 - 피. 파스토리스 mutS 계놈 내로의 통합을 위한 플라스미드 상에서의 헵 생합성 유전자의 조립

[0119] 전체 카세트인 "프로모터-유전자-종결인자"를 피키아 계놈 내로의 통합을 위한 플라스미드를 조립하기 위해 제한 엔도뉴클레아제에 대한 부위를 함유하는 프라이머로 PCR 증폭시켰다.

[0120] pGAN 플라스미드 상에서의 Paox1_UPS_FDHterm-Paox1_UPD_FDHterm-Paox1_CPO_FDH term. 의 조립 (pMx327)

[0121] pGAN-CPG옥시다제 (pMx312)를 벡터로 사용하여 UPS 및 UPD 카세트를 클로닝하였다. UPG III 신타제 카세트를 상응하게 제한 엔도뉴클레아제에 대한 NheI 및 SphI 인식 부위를 함유하는 AOX1 프로모터 / FDH1 종결인자에 대한 프라이머로 pMx310으로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
Mx00399	NheI-pAOX1-F	CAA TCG CTA GCA TCC AAC ATC CAA AGA CGA AAG G	31
Mx00401	SphI-FDH1-R	GGA TAG CAT GCA CCT TAT CAA GAT AGC TAG AAA TAG AAA TGG	32

[0122]

[0123] UPG III 데카르복실라제 카세트를 상응하게 제한 엔도뉴클레아제에 대한 SphI 및 AgeI 인식 부위를 함유하는 AOX1 프로모터 / FDH1 종결인자에 대한 프라이머로 pMx311로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
Mx00402	SphI-pAOX1I-F	CAA TAG CAT GCA ACA TCC AAA GAC GAA AGG TTG AAT G	33
Mx00404	AgeI-FDH1-R	CAT GGT ACC GGT ACC TTA TCA AGA TAG CTA GAA ATA GAA ATGG	34

[0124]

[0125] 퓨전 하이-피델리티 DNA 폴리머라제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0530L)를 사용하여 플라스미드로부터 DNA를 증폭시켰다.

[0126] 수득된 PCR 생성물을 DNA 클린&컨센트레이터-5 (자이모 리서치, 카탈로그 # D4004)를 사용하여 정제하고, DNA를 25 μ l의 H₂O에 용출시켰다.

[0127] 벡터로 지정된 pGAN-CPG옥시다제 (pMx312)를 1x 컷스마트 버퍼에서 NheI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3131S) 및 AgeI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3552S)로 철야로 37°C에서 소화시켰다.

[0128] UPG III 신타제 카세트 PCR 생성물을 1x 컷스마트 버퍼에서 NheI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3131S) 및 SphI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3182S)로 철야로 37°C에서 소화시켰다.

[0129] UPG III 데카르복실라제 카세트 PCR 생성물을 1x 컷스마트 버퍼에서 SphI-HF (뉴 잉글랜드 바이오랩스, 카탈로그

그 # R3182S) 및 AgeI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3552S)로 철야로 37℃에서 소화시켰다.

[0130] 소화된 벡터 및 PCR 생성물을 자이모클린 겔 DNA 회수 키트 (자이모 리서치, 카탈로그 # D4002)를 사용하여 0.8% 아가로스 겔로부터 겔 정제하였다. DNA를 20 μ l H₂O에 용출시켰다.

[0131] NheI-SphI로 소화된 UPG III 신타제 카세트, SphI-AgeI로 소화된 UPG III 데카르복실라제 카세트 및 NheI-AgeI로 소화된 벡터 사이의 3-방식 라이게이션을 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0202S)를 사용하여 철야로 16℃에서 10 μ l로 설정하였다.

[0132] 1.5 μ l의 라이게이션 혼합물을 1.7 kV로 설정된 마이크로펄서 (바이오래드)를 사용한 전기천공에 의해 20 μ l의 일렉트로맥스 DH10B T1 파지-저항성 적격 세포 (인비트로젠, 카탈로그 # 12033-015) 내로 형질전환시켰다; 200 rpm으로 진탕시키면서 세포를 37℃에서 1시간 동안 1 ml SOC 내에서 인큐베이션하였다. 10 μ l의 회수 혼합물을 100 μ g/ml의 농도로 엠포실린을 함유하는 LB 한천 플레이트 상에 플레이팅하였다. 플레이트를 철야로 37℃에서 인큐베이션하였다. 삽입물의 존재에 대해 콜로니 PCR에 의해 콜로니를 스크리닝하였다. 벡터와 삽입물 사이의 접합부의 서열을 확인하였다.

[0133] *Paox1_ALA신타제_FDH1term.-Paox1_PPG옥시다제_FDH1term-Paox1_Fc_ FDH1term.-Paox1_PBD_ FDH1term 카세트 (pMx330)의 조립*

[0134] a. 유전자-카세트의 PCR 증폭:

[0135] ALA신타제 카세트를 상응하게 제한 엔도뉴클레아제에 대한 NheI 및 XhoI 인식 부위를 함유하는 AOX1 프로모터 / FDH1 종결인자에 대한 프라이머로 pMx310으로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
MxO0399	NheI-pAOX1-F	CAA TCG CTA GCA TCC AAC ATC CAA AGA CGA AAG G	35
MxO0400	XhoI-FDH1-R	GAT ATT GCT CGA GAC CTT ATC AAG ATA GCT AGA AAT AGA AAT G	36

[0136]

[0137] PPG옥시다제 카세트를 상응하게 제한 엔도뉴클레아제에 대한 XhoI 및 AfIII 인식 부위를 함유하는 AOX1 프로모터 / FDH1 종결인자에 대한 프라이머로 pMx313으로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
MxO0403	XhoI-pAOX1-F	CAA TCT CGA GAA CAT CCA AAG ACG AAA GGT TG	37
MxO0437	AfIII-FDH1-R	CAA CCA TTT CTA TTT CTA GCT ATC TTG ATA AGG TCT TAA GTC CA	38

[0138]

[0139] 페로킬라타제 카세트를 상응하게 제한 엔도뉴클레아제에 대한 AfIII 및 AgeI 인식 부위를 함유하는 AOX1 프로모터 / FDH1 종결인자에 대한 프라이머로 pMx323으로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
MxO0404	AgeI-FDH1-R	CAT GGT ACC GGT ACC TTA TCA AGA TAG CTA GAA ATA GAA ATG G	39
MxO0436	AfIII-pAOX1-F	TTA CTT AAG TCC AAC ATC CAA AGA CGA AAG GTT G	40

[0140]

[0141] G418 마커를 하기 프라이머를 사용하여 바이오그래메틱스로부터 구입한 pJAG 플라스미드로부터 PCR 증폭시켰다:

프라이머 명칭	인식 부위	서열	SEQ ID NO:
MxO0438	Mlu-G418-F	TCA CAG ACG CGT TGA ATT GTC C	41
MxO0439	BbvCI-G418-R	TTG CTC CTC AGC TTA GAA GAA CTC GTC CAA CAT CAA GTG	42

[0142]

[0143] 퓨전 하이-피델리티 DNA 폴리머라제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0530L)를 사용하여 플라스미드로부터 DNA를 증폭시켰다. PCR 생성물을 수득하고, DNA 클린&컨센트레이터-5 (자이모 리서치, 카탈로그 # D4004)를 사용하여 정제하고, DNA를 25 μ l의 H₂O에 용출시켰다.

- [0144] b. 벡터 제조
- [0145] 벡터로 지정된 pGAZ-PBD (pMx321)를 1x 컷스마트 버퍼에서 NheI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3131S) 및 XhoI (뉴 잉글랜드 바이오랩스, 카탈로그 # R0146S)로 철야로 37℃에서 소화시켰다.
- [0146] pGAZ-ALAsyn.-PBD (pMx328)를 1x NE버퍼3.1에서 MluI (뉴 잉글랜드 바이오랩스, 카탈로그 # R0198S) 및 BbvCI (뉴 잉글랜드 바이오랩스, 카탈로그 # R0601S)로 철야로 37℃에서 소화시켰다.
- [0147] pGAG-ALAsyn-PBD (pMx332)를 1x 컷스마트 버퍼에서 XhoI (뉴 잉글랜드 바이오랩스, 카탈로그 # R0146S) 및 AgeI-HF (뉴 잉글랜드 바이오랩스, 카탈로그 # R3552S)로 철야로 37℃에서 소화시켰다.
- [0148] c. 중간체 구축물 제조 및 최종 카세트 조립
- [0149] 소화된 벡터 및 PCR 생성물을 자이모클린 겔 DNA 회수 키트 (자이모 리서치, 카탈로그 # D4002)를 사용하여 0.8% 아가로스 겔로부터 겔 정제하였다. DNA를 20 μ l의 H₂O에 용출시켰다.
- [0150] NheI-XhoI 제한 엔도뉴클레아제로 소화된 pGAZ-PBD (pMx321) 벡터와 동일한 효소로 소화된 ALA신타제 카세트 PCR 생성물을 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0202S)를 사용하여 철야로 16℃에서 10 μ l 반응으로 라이게이션시켜, pGAZ-ALAsyn.-PBD 플라스미드 (pMx328)가 산출되었다.
- [0151] XhoI 및 AgeI-HF 제한 엔도뉴클레아제로 소화된 pGAG-ALAsyn-PBD (pMx332)와 XhoI, Af1II 및 Af1II, AgeI-HF 제한 엔도뉴클레아제로 상응하게 소화된 PPG옥시다제 카세트 및 페로킬라타제 카세트 PCR 생성물을 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스, 카탈로그 # M0202S)를 사용하여 3-방식 라이게이션 반응으로 라이게이션시켜, pGAG-ALA신타제_PPG옥시다제_Fc_PBD (pMx330)가 산출되었다.
- [0152] 1.5 μ l의 라이게이션 혼합물을 1.7 kV로 설정된 마이크로펄서 (바이오래드)를 사용한 전기천공에 의해 20 μ l의 일렉트로맥스 DH10B T1 파지-저항성 적적 세포 (인비트로젠, 카탈로그 # 12033-015) 내로 형질전환시켰다; 200 rpm으로 진탕시키면서 세포를 37℃에서 1시간 동안 1 ml SOC 내에서 인큐베이션하였다. 10 μ l의 회수 혼합물을 100 μ g/ml의 농도로 앰피실린을 함유하는 LB 한천 플레이트 상에 플레이팅하였다. 플레이트를 철야로 37℃에서 인큐베이션하였다. 삽입물의 존재에 대해 콜로니 PCR에 의해 콜로니를 스크리닝하였다. 벡터와 삽입물 사이의 접합부의 서열을 확인하였다.

명칭	구축물	유전자
pMx0327	pGAN-UPGsyn_UPGdecarb_CPGoxy	UPGsyn_UPGdecarb_CPGoxy
pMx0328	pGAZ-ALAsyn-PGC	ALAsyn-PBD
pMx0330	pGAG-ALAsyn_PBD_PPG_Fc	ALAsyn_PBD_PPG_Fc
pMx0332	pGAG-ALAsyn_PBD	G418 마커

- [0153]
- [0154] 실시예 13 - 유전자 카세트가 있는 선형화된 플라스미드의 피. 파스토리스 Bg11 게놈 내로의 통합
- [0155] 생산 균주인 MX0183을 생성시키는데 사용된 플라스미드가 도 2에서 제시된다. 생산 균주인 MX0183에 이른 변형을 만드는데 취해진 단계가 도 3에서 도시된다.
- [0156] 피. 파스토리스 Bg11 게놈 내로 도입될 제1 효소는 ALAD였다. pAOX1-구동 ALAD를 함유하는 플라스미드 (pMX229, 도 2i 및 도 3)를 PmeI 제한 효소 (뉴 잉글랜드 바이오랩스)를 사용하여 선형화시켰다. 선형화된 플라스미드를 기술된 바와 같이 0.8% 아가로스 겔로부터 정제하고, 천연 AOX1 유전자좌에서의 상동 재조합을 사용하여 피. 파스토리스 내로 형질전환시켜, 균주 MX099가 생성되었다 (도 3).
- [0157] pMX282 (도 2ii 및 도 3)로 지정된 pAOX1 프로모터 제어 하의 대두 LegH 유전자 (피. 파스토리스에 대해 최적화된 서열; 서열식별번호(SEQ ID NO): 3)의 2개의 카피를 함유하는 플라스미드를 SfiI 제한 효소를 사용하여 선형화시켰다. 선형화된 플라스미드를 기술된 바와 같이 0.8% 아가로스 겔로부터 정제하고, ALAD를 함유하는 피. 파스토리스 균주 내로 형질전환시켜, 균주 MX0118이 생성되었다 (도 3). qPCR을 사용하여, 균주 MX0118이 LegH 유전자의 여러 카피를 함유하였음이 결정되었으며, 이는 재조합 시점의 플라스미드 pMX282의 콘카테머화로 인해서일 것이다.
- [0158] AOX1 프로모터의 제어 하의 우로포르피리노겐 III 신타제 (UPS), 우로포르피리노겐 III 데카르복실라제 (UPD) 및 코프로포르피리노겐 III 옥시다제 (CPO) (각각 단계 4, 5 및 6을 촉매하는 효소)를 코딩하는 유전자를 함유하는 플라스미드 pMX327 (도 2iii 및 도 3)을 SfiI 제한 엔도뉴클레아제로 선형화시키고, MX0118 내로 도입하

여, 균주 MXY0170이 산출되었다 (도 3).

[0159] 피. 파스토리스 계놈으로부터의 ALA 신타제 (ALAS), 프로토포르피린 III 옥시다제 (PPO), 페로킬라타제 (FC) 및 포르포빌리노겐 데아미나제 (PBD) (각각 단계 1, 7, 8 및 3을 촉매하는 효소)를 코딩하는 유전자를 플라스미드 pMX330 (도 2iv 및 도 3) 상에서 조립시켰다. pMX330을 SfiI 제한 엔도뉴클레아제로 선형화하고, MXY0170 내로 형질전환시켜, 균주 MXY0183이 생성되었다 (도 3). MXY0183의 유전자형을 PCR 및 qPCR을 사용하여 확인하였다.

[0160] **생산 균주 MXY0207**

[0161] 실시예 14 - pGAB 발현 벡터의 구축

[0162] pGAZ 벡터 (바이오그래매틱스, 인크., 캘리포니아주 칼스배드) 내의 제오신 저항성 유전자의 오픈 리딩 프레임 을 아스페르길루스 테레우스(*Aspergillus terreus*)로부터의 블라스티시딘 S 데아미나제 (BSD) 유전자로부터의 오픈 리딩 프레임으로 교체함으로써 pGAB 발현 벡터 (도 4A)를 구축하였고, 이는 플라스미드를 보유하는 형질전환체를 항생제 블라스티시딘 S로 선택하는 것을 허용한다.

[0163] 본원에 기술된 바와 같이 하이-피델리티 폴리머라제 연쇄 반응을 사용하여 올리고뉴클레오타이드 프라이머 Mx00476 및 Mx00477을 사용하여 BSD 오픈 리딩 프레임을 상업적으로 합성된 DNA 분자로부터 증폭시켰다.

프라이머 명칭	설명	서열	SEQ ID NO:
MxO0477	BSD_정방향	TTA GTC TTG CTC CTC AGC TTA GCC	43
MxO0476	BSD_역방향	TCA CAG ACG CGT TGA ATT GTC C	44

[0164]

[0165] BSD PCR 생성물을 1x TBE 완충제 (89 mM 트리스(Tris), 89 mM 붕산, 2 mM EDTA, pH 8.3)에서 1% 아가로스 겔 상에서의 겔 전기영동으로 정제하고, SYBR 세이프(SYBR Safe) DNA 겔 염색 (라이프 테크놀로지스(Life Technologies), 캘리포니아주 칼스배드)를 사용하여 가시화하였다. 원하는 DNA 단편을 아가로스 겔로부터 절제하고, DNA를 자이모클린 겔 DNA 회수 키트 (자이모 리서치, 캘리포니아주 어바인)를 사용하여 회수하였다.

[0166] 정제된 BSD PCR 생성물 및 pGAZ 벡터를 1x NE버퍼 3.1 (100 mM NaCl, 50 mM 트리스-HCl, 10 mM MgCl₂, 100 μg/ml BSA, pH 7.9 @ 25°C)에서 각각 10 유닛의 MluI 및 BbvCI 제한 엔도뉴클레아제 (뉴 잉글랜드 바이오랩스, 매사추세츠주 업스위치)로 1시간 동안 37°C에서 소화시켰다. 소화된 DNA 생성물을 상기 기술된 바와 같이 겔 전기영동으로 회수하였다.

[0167] 정제된 MluI 및 BbvCI-소화 BSD 생성물 및 pGAZ 벡터를 20 μl 반응으로 1x T4 DNA 리가제 반응 완충제 (50 mM 트리스-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, pH 7.5 @ 25°C)에서 16°C에서 2시간 동안 20 μl 반응으로 400 유닛의 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스)와 함께 인큐베이션하였다. 전기천공적격 이. 콜라이 DH10B 세포를 2 μl의 라이게이션 반응물로 형질전환시키고, 항생제 저항성 형질전환체를 100 μg/μl 앰피실린 이 보충된 LSB 한천 플레이트 상에서 선택하였다.

[0168] 실시예 15 - Mxr1 발현 벡터의 구축

[0169] Mxr1 오픈 리딩 프레임을 pGAB 벡터 내로 도입함으로써 Mxr1 발현 벡터인 pMx354를 구축하였다 (도 4B). 번역 개시부가 피키아 파스토리스로부터의 메탄올-유도성 알콜 옥시다제 1 (AOX1) 프로모터의 바로 하류에 있고 번역 정지 신호에 피. 파스토리스 FDH1 유전자로부터의 전사 종결인자 서열이 바로 이어지면서 Mxr1 오픈 리딩 프레임이 pGAB 내로 삽입되었다.

[0170] Mxr1 단백질을 코딩하는 오픈 리딩 프레임을 바이오그래매틱스, 인크. (캘리포니아주 칼스배드)로부터 취득된 피키아 파스토리스 균주 Bgl1 MutS로부터 단리된 계놈 DNA로부터 증폭시켰다. Mxr1 오픈 리딩 프레임을 프라이머 Mx00495 (TTT TGC GGC CGC ATG AGC AAT CTA CCC CCA ACT TTT G (서열식별번호: 45)) 및 Mx00496 (AAA AGC GGC CGC CTA GAC ACC ACC ATC TAG TCG GTT (서열식별번호: 46))으로 피. 파스토리스 계놈 DNA로부터 증폭시켰고, 이들은 측면의 NotI 제한 엔도뉴클레아제 인식 부위를 부가하였다. 본원에 기술된 바와 같은 폴리머라제 연쇄 반응을 사용하여 증폭을 달성하였다.

[0171] 증폭된 Mxr1 PCR 생성물 및 pGAB 벡터를 1x NE버퍼 3.1 (100 mM NaCl, 50 mM 트리스-HCl, 10 mM MgCl₂, 100 μ

g/ml BSA, pH 7.9 @ 25°C)에서 10 유닛의 NotI 제한 엔도뉴클레아제 (뉴 잉글랜드 바이오랩스)로 1시간 동안 37°C에서 소화시켰다. 소화 후, NotI-소화 pMx352 벡터를 1x 엔탁틱(Antarctic) 포스파타제 완충제 (50 mM 비스-트리스-프로판-HCl, 1 mM MgCl₂, 0.1 mM ZnCl₂, pH 6 @ 25°C)에서 37°C에서 15분 동안 5 유닛의 엔탁틱 포스파타제 (뉴 잉글랜드 바이오랩스)로 처리하였다.

[0172] 정제된 NotI-소화 Mxr1 단편 및 pMx352 벡터를 1x TBE 완충제 (89 mM 트리스, 89 mM 붕산, 2 mM EDTA, pH 8.3)에서 1% 아가로스 겔 상에서의 전기영동으로 분리하고, SYBR 세이프 DNA 겔 염색 (라이프 테크놀로지스, 캘리포니아주 칼스배드)을 사용하여 가시화하였다. 원하는 DNA 단편을 아가로스 겔로부터 절제하고, DNA를 자이모클린 겔 DNA 회수 키트 (자이모 리서치, 캘리포니아주 어바인)를 사용하여 회수하였다.

[0173] Mxr1 오픈 리딩 프레임에 함유하는 NotI-소화 단편을 라이게이션에 의해 AOX1 프로모터의 바로 하류의 NotI 부위에서 pGAB 내로 도입하였다. 137 ng의 Mxr1 오픈 리딩 프레임을 코딩하는 NotI-소화 DNA 및 60 ng의 NotI-소화, 포스파타제-처리 pMx352를 함유하는 혼합물을 20 µl 반응으로 1x T4 DNA 리가제 반응 완충제 (50 mM 트리스-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, pH 7.5 @ 25°C)에서 16°C에서 2시간 동안 20 µl 반응으로 400 유닛의 T4 DNA 리가제 (뉴 잉글랜드 바이오랩스)와 함께 인큐베이션하였다. 전기천공격격 이. 콜라이 DH10B 세포를 2 µl의 라이게이션 반응물로 형질전환시키고, 항생제 저항성 형질전환체를 100 µg/µl 앰피실린이 보충된 LSB 한천 플레이트 상에서 선택하였다. 플레이트를 37°C에서 철야로 인큐베이션하였다. 프라이머 Mx00495 및 Mx00496을 사용하여 PCR에 의해 삽입물의 존재에 대해 콜로니를 스크리닝하였다. 최종 벡터의 서열을 DNA 시퀀싱으로 확인하였다.

[0174] 클로닝 동안, 6개의 추가적인 아미노산을 Mxr1의 N-말단에서 도입하였다. Mxr1 오픈 리딩 프레임이 "핵산 서열" 섹션에서 제시되고, 여기서 클로닝으로부터의 잔여 아미노산이 밑줄이 그어져 제시된다. N-말단에 추가적인 6-아미노산이 있는 Mxr1 서열을 함유하는 피키아 생산 균주 및 야생형 Mxr1 (즉, N-말단에 추가적인 6 아미노산이 없음)을 함유하는 피키아 균주가 발효 탱크에서 구별가능하지 않았다.

[0175] 실시예 16 - 천연 Mxr1 발현 벡터의 구축

[0176] pMX354로 지정된, pAOX1 프로모터의 제어 하의 Mxr1 전사 조절인자 유전자를 함유하는 플라스미드를 PCR 증폭용 주형으로서 사용하였다. AOX1 프로모터의 3' 끝부분, LegH 오픈 리딩 프레임, 및 AOX1 종결인자를 하기 제시된 프라이머 Mx00617 및 Mx00647을 사용하여 pMX354로부터 증폭시켰다. AOX1 종결인자, 링커 및 AOX1 프로모터의 5' 끝부분을 프라이머 Mx00618 및 Mx00646을 사용하여 pMX382로부터 증폭시켰다.

프라이머 명칭	서열	SEQ ID NO:
Mx00646	ACTAGATGGTGGTGTCTAGTCAAGAGGATGTCAGAATGCCATTTG	47
Mx00647	TCTGACATCCTCTTGACTAGACACCACCATCTAGTCGGTTTTCTAG	48

[0177]

[0178] PCR 생성물을 수득하고, DNA 클린&컨센트레이터-5를 사용하여 정제하고, DNA를 12 µl의 H₂O에 용출시켰다. 그 후, 정제된 PCR 생성물을 조합하고, 프라이머 Mx00617 및 Mx00618을 사용하는 후속 라운드의 PCR 증폭을 위한 주형으로서 사용하였다. 생성된 PCR 생성물은 AOX1 프로모터의 3' 끝부분에 이어지는 Mxr1 오픈 리딩 프레임, AOX1 종결인자, 짧은 링커 서열, 및 AOX1 프로모터의 5' 끝부분으로 구성되었다. PCR 생성물을 수득하고, 본원에 기술된 바와 같이 정제하였다. 정제된 PCR 생성물을 제로 블런트® TOPO® PCR 클로닝 키트(Zero Blunt® TOPO® PCR Cloning Kit) (인비트로젠, 카탈로그 # K2800-20)를 사용하여 pCR™-블런트(Blunt) II-TOPO® 벡터 내로 클로닝하여, pMX402 벡터가 생성되었다.

[0179] 실시예 17 - 피. 파스토리스 균주 MXY0206 및 MXY0207의 구축

[0180] pMx354 Mxr1 발현 벡터 (도 4B)를 DNA 형질전환에 의해 MXY0183 균주 내로 도입하였다 (도 3).

[0181] 1시간 동안 37°C에서 1x NE버퍼 4 (50 mM 아세트산칼륨, 20 mM 트리스-아세테이트, 10 mM 아세트산마그네슘, 1 mM DTT, pH 7.9@25°C)에서 20 유닛의 PmeI 제한 엔도뉴클레아제 (뉴 잉글랜드 바이오랩스)로 소화시킴으로써 pMx354 벡터 (1.5 µg)를 AOX1 프로모터 서열 내의 독특한 PmeI 부위에서 선형화시켰다.

[0182] PmeI-소화 pMX354 벡터를 상기 기술된 바와 같이 겔 전기영동으로 정제하고 자이모클린 겔 DNA 회수 키트를 사용하여 회수하였다. 선형화된 pMX354 벡터를 형질전환에 의해 균주 MXY0183 내로 도입하고, 블라스티시딘-함유 배지 상에서 선택하였다. 2개의 독립적인 클론이 형질전환으로부터 수득되었고, 이들을 MXY0206 및 MXY0207로

지정하였다. 이들 균주 내의 AOX1 프로모터의 제어 하의 Mxr1의 추가적인 카피의 존재를 PCR로 확인하였다.

[0183] **생산 균주 MXY0291**

[0184] 실시예 18 - 균주 MXY0213 및 MXY0260의 구축

[0185] 도 5는 헴 생합성 경로의 7개의 효소를 함유하는, 항생제 마커가 없는 균주 MXY0213을 구축하기 위해 취해진 단계를 나타낸다. 각각의 끝부분에서 pAOX1 프로모터에 대한 상동성이 있는 pAOX1 하의 변이체 Mxr1 (N 말단의 6개의 추가 아미노산)을 함유하는 선형 DNA 조각을 공동-형질전환을 사용하여 도입하였다 (도 5 및 도 6i). 동시에 이러한 선형 Mxr1 발현 카세트를 형질전환에 의해 pIL75 플라스미드와 함께 피키아 균주 MXY213 내로 도입하였다. pIL75 벡터는 형질전환된 세포의 게놈 내로의 통합 없이 플라스미드 벡터를 유지하는 것을 허용하는 panARS 자가 복제 서열 (Liachko & Dunham, 2014, FEMS Yeast Res., 14:364-7), 및 항생제 G418로 형질전환체를 선택하기 위한 kanMX 마커를 보유한다. 형질전환된 세포를 pIL75 플라스미드 상의 kanMX 마커의 존재에 대해 G418이 보충된 배지 상에서 선택하였다. pIL75 플라스미드를 받아들였고 또한 Mxr1 발현 카세트가 정확하게 통합된 형질전환체에 대해 콜로니 PCR에 의해 피키아 형질전환체를 스크리닝하였다.

[0186] 실시예 19 - LegH 발현 카세트를 피키아 내로 도입하기 위한 공동-형질전환

[0187] pMX399로 지정된, pAOX1 프로모터의 제어 하의 대두 LegH 유전자의 상이한 피키아 파스토리스-코돈 최적화 변이체 (변이체 3; 서열식별번호: 5)를 함유하는 플라스미드를 유전자의 PCR 증폭을 위한 주형의 공급원으로서 사용하였다. TOPO 클로닝 플라스미드 pMX401로부터의 백본을 PCR 증폭시켰다. 삽입물 및 벡터를 김슨(Gibson) 어셈블리 (NEB의 김슨 어셈블리 키트)를 사용하여 조립하여, 플라스미드 pMX422를 생성시켰다. 이러한 플라스미드를 하기 제시된 프라이머 Mx00617 및 Mx00618을 사용하는 후속 라운드의 PCR 증폭을 위한 주형으로서 사용하였다.

[0188]

프라이머 명칭	서열	SEQ ID NO:
MxO0617	AAACGCTGTCTTGGAACCTAATATGAC	49
MxO0618	AAACTGTCAGTTTTGGGCCATTTG	50

[0189] 생성된 PCR 생성물은, 5'에서 3' 방향으로, AOX1 프로모터의 3' 끝부분에 이어지는 LegH 변이체 3 오픈 리딩 프레임, AOX1 종결인자, 짧은 링커 서열, 및 AOX1 프로모터의 5' 끝부분으로 구성되었다 (도 6ii). PCR 생성물을 수득하고, 본원에 기술된 바와 같이 아가로스 겔 전기영동으로 정제하였다.

[0190] LegH 발현 카세트가 게놈 내로 통합된 형질전환체를 PCR로 스크리닝하고, qPCR을 사용하여 LegH 유전자 카피수에 대해 특성화하였다.

[0191] 실시예 20 - 형질전환체에서의 선택 마커를 보유하는 플라스미드 벡터의 교정

[0192] 대두 LegH 발현 카세트가 정확하게 통합된 것으로 콜로니 PCR에 의해 나타났고 카피수가 높은 것으로 qPCR에 의해 나타난 클론에서, G418 상에서의 선택에 요구되는 pIL75 플라스미드를 항생제에 대한 선택을 완화함으로써 제거하였다. 형질전환체를 G418 항생제가 결여된 배지 상에 단일 콜로니에 대해 스트리킹하였다. 선택의 부재 하에 panARS 플라스미드는 안정적으로 유지되지 않기 때문에, 이러한 조건 하에 pIL75가 형질전환된 세포로부터 신속하게 상실되었다. 생성된 피키아 균주 MXY0291은 MXY0207과 유사한 카피수로 LegH 발현을 위한 서열을 함유하지만, 선택을 위한 이중성 서열이 결여된다.

[0193] **생산 균주 MXY0330, MXY0333, 및 MXY0338**

[0194] 실시예 21 - 균주 MXY0306의 구축

[0195] 균주 MXY0291의 유전자형 PCR에서, 이러한 균주의 구축 동안 CPG옥시다제 코딩 서열의 일부분이 결실되었음이 밝혀졌다. 말단절단 카피의 교체에 의해 전장 CPG옥시다제 코딩 영역을 복원하였다. 간략하게, pAOX1 프로모터 및 전장 CPG옥시다제 코딩 영역을 함유하는 선형 DNA 단편을 하기 제시된 프라이머 Mx00866 및 Mx00867을 사용하여 플라스미드 pMX312로부터 PCR 증폭에 의해 생성시켰다.

프라이머 명칭	서열	SEQ ID NO:
MxO0866	ACGCTGTCTTGGAACCTAATATGAC	51
MxO0867	TACCCATTCAATAGGATTTTGTAGTACCTGC	52

[0196]

[0197]

선형 pAOX1-CPG옥시다제 DNA 단편을 pIL75 플라스미드와의 공동-형질전환에 의해 균주 MXY0291 내로 도입하였다. G418을 함유하는 배지 상에서 형질전환체를 선택한 후, PCR에 의해 전장 CPG옥시다제 코딩 영역의 존재에 대해 스크리닝하였다. 전장 CPG옥시다제를 함유하는 단리물을 확인하고, 이어서 상기 기술된 바와 같이 G418 상에서의 선택에 요구되는 플라스미드 벡터를 교정하였다. 이러한 균주를 MXY0306로 지정하였다 (도 5 참조).

[0198]

실시예 22 - 하이브리드 프로모터 균주를 위한 선형 구축물

[0199]

LegH 변이체 3을 본원에 지시된 3개의 피키아 파스토리스의 천연 구성적 프로모터 각각의 지시 하에 발현시켰다. 선형 구축물이 도 7에 제시되고, 이는 프로모터의 3' 절반, 이어지는 LegH var3, 이어지는 FDH1 전사 종결인자를 함유하였다. 여기에 아시비아 고시피이(*Ashbya gossypii*)로부터의 pTEF 프로모터, 아스페르길루스 니둘란스(*Aspergillus nidulans*)로부터의 아세트아미다제 유전자 (amdS) 및 아시비아 고시피이로부터의 TEF 종결인자를 함유하는 항생제 저항성 카세트가 바로 이어졌다. 마지막으로, 구축물은 프로모터의 5' 절반을 함유하였다. 이러한 선형 카세트를 하기 표에 열거된 올리고뉴클레오티드 프라이머를 사용하여 증폭시켜, 천연 피키아 게놈 내의 각각의 프로모터에 상동성인 5' 및 3' 끝부분 상의 수백개의 염기쌍을 함유하는 구축물을 생성시켰다.

프라이머 명칭	서열	SEQ ID NO:
MXO0718	GAGCTTCTTCTACGGCCCCC	53
MXO0723	TCCAGCAGAGTAAAATTTCTTAGGGAC	54
MXO0724	CTCTTTTAGGTTTAAAGTTGTGGGAACAGTAACA	55
MXO0729	GTGGGTGCTTCTTTGCGTGG	56
MXO0730	AGAATTGCCATCAAGAGACTCAGGACT	57
MXO0735	GATAGAGAGAAATCGCAAACCTTTGAGAGGAAG	58

[0200]

[0201]

선형 구축물을 증폭시키는데 사용된 프라이머

[0202]

적격 MXY0306 세포를 각각의 선형 카세트로 형질전환시키고, amdS 선택 카세트를 함유하는 형질전환체를 아세트아미드를 단독 질소원으로 함유하는 한천 플레이트 상에서 성장하는 능력을 기초로 선택하였다. 이들 균주를 정제하고, 단리하고, 구성적 프로모터의 제어 하의 LegH의 존재를 PCR로 증명하였다 (도 5).

[0203]

실시예 23 - 핵산 서열

[0204]

Mxr1 핵산 서열 (밑줄친 뉴클레오티드는 클로닝 동안 도입된 N-말단의 6개의 아미노산을 코딩함) (서열식별번호: 1)

ATGCGAGACCGCGGCCGCATGAGCAATCTACCCCCAACTTTTGGTTCCACTAGACAATCTCCAGA
AGACCAATCACCTCCCGTGCCCAAGGAGCTGTCATTCAATGGGACCACACCCCTCAGGAAAGCTAC
GCTTATTTGTCTGTCTCAGACATGTACTCGAGCATTTGCTCGTCAGGAACACTTGAAACGACACGAA
AGGTCTCACACCAAGGAGAAACCTTTTCAGCTGCGGCATTTGTTCTCGTAAATTCAGCCGTCGAGA
TCTGTTATTGAGACATGCCCAAAAACCTGCACAGCAACTGCTCTGATGCGGCCATAACAAGACTAA
GGCGCAAGGCAACTCGTCGGTCTTCTAATGCCGCGGGTTCCATATCTGGTTCTACTCCGGTGACA
ACGCCAAATACTATGGGTACGCCCAGATGGCGAGAAACGAAAAGTTTCAGAAACTGGCCGGCCG
CCGGGACTCAAATGAACAGAAACTGCAACTGCAACAACAACATCTACAGCAACAACCACAGTTGC
AATACCAACAATCTCTTAAGCAGCATGAAAATCAAGTCCAGCAGCCTGATCAAGATCCATTGATA
TCCCCGAGAATGCAATTATTCAATGATTCCAACCATCACGTAAACAATTTGTTTGATCTTGGACT
AAGAAGAGCTTCCCTTCTCCGCCGTTAGTGGAATAAATTATGCCCATTTATGTGAATAATTTTCAAC
AAGATGCCCTCTTCTACCAATCCAAATCAAGATTCAAATAATGCCGAATTTGAGAATATTGAATTT
TCTACCCACAAAATGATGCCCGTTGAAGATGCTGAACTTGGATGAACAACATGGGTCCAATTCC
GAAGTTCTCTCTCGATGTGAACAGGAACATTGGTGATAGCTTTACAGATATACAACACAAGAATT
CAGAGCTTATTATATCCGAACCGCCCAAGGACACCGCTCCAAACGACAAGAAGTTGAATGGCTAC
TCTTTTACGAAGCCCCCATCAAGCCATTAGAATCCCTATTTTCTGTCAGGAATACAAAGAGAAAA
CAAGTATAAAAACAAATGACGACTCTCCAGACACCGTGGATAATAACTCCGCACCGGCTGCTAATA
CCATTCAAGAACTTGAGTCTTCTTTGAATGCATCCAAGAATTTTGTGTTGCCAACTGGTTATTCC
TTCTATGGTAATTTGGACCAACAGACTTTCTCTAACACGTTATCATGCACTTCTTCTAATGCCAC
AATTTGCCCCATTCTACTCGATAACTCCATTAATAATAACTCCACTAGTGACGTGAGACCAGAAT
TTAGAACACAAAGTGTACCTCTGAAATGAGTCAAGCCCCCTCCCCCTCCTCAAAAAACAACCTCG
AAATATTCCACCGAAGTTCTTTTACCAGCAACATGCGGTCGTTTATTCACTACGCTCTTTCCAA
GTATCCTTTTATTGGTGTGCCCACTCCAACCTCTTCCGGAGAACGAAAGACTAAATGAATATGCTG
ATTCAATTCACCAACCGTTTCTTAAATCATTATCCTTTTCATACATGTCACGATTCTCAAAGAATAC
TCCCTTTTCAAGGCAATTTTAGATGAGAATGAGTCGACTAAGAACTGGGAAAATAATCAGTTTTA
CTTAGAGAACCACGAATATCAATTGTTTGTCTTCTCTTTTGGTGGCTACGATAGGTGCAGTAC
TATCAAAACAACAAAAAGGATGCTTCGAATTTATACGAAGCTTCAAGGCGTTGTATTTCATGTTTAC

[0205]

TTAGATTCCAGGAAAAAGATACCCACTTCCCTTGTCGCAATAACAATGACTCTCCACTTTGGCT
AATTCAATCCCTGACGTTATCTGTTATGTATGGGTATTTGCGGACAATGACATTAGTTTGAATG
TCGTGATCAGACAAGTTAACGCACTTAATTCTCTGGTCAAGACTTTCGGGCTGAATAGGACCTCA
ATTATAGATCTTTTCAACATCAACAAACCTTTGGATAATGAACCTCTGGAATCAATTCGTGAAAT
AGAGTCCACCGTAAGGACAATCCACACGATTTTCAAATCAGTTCCAACCTAAGCGCCTTGTACA
ATATTATTCCATCGTTGAAAATTGATGACCTAATGATTACTCTACCAGTTCCCACAACACTTTGG
CAAGCTGATTCTTTTGTGAAATTCAAAGTCTAAGTTACGGAAATCAGATCCCTTTTCAATATAC
AAGAGTACTACAGAAATTTGATTGATTACAATCAGCCATTGAGCGATGGAAAATTTTGTATGAAA
ACCATGTAAGTGAGTTTGGACTCATATGCCATACAGAATGGTCTACACCAATACAGCTATTTCCAA
AAATTGACTGCTGTCAATAACAGAGAAGATGCGCTATTCACAAAGGTTGTTAATTCACCTTCACAG
TTGGGATAGGATGATTTCGAATTCTGATTGTGTTCCAAAGAAGATATATCAGCAGAGTTGCTTGA
TTTTGGACTCAAAGTTGCTTAATAATTTCCCTGATTGTCAAGAGCTATTGAAAGTTTCGACCGGA
GACGTTAGTTCTTTGAATAAGTTAAAAGAAAACGTGTGGCTTAAAAACTGGAATCAAGTGTGTGC
TATCTATTATAACAGCTTCATGAACATTCTGCTCCAGTATTCAAAGAAGTACAATGACATAG
AGTTTGTGGATGACATGATTAAATTTGAGTCTAATCATCATCAAGATTATGAAACTCATTTTCTAT
AACAAATGTCAAAGACAAATTATGAGGATGAAAATGACTTCAAATTGCAAGAGTTAAATTTAACATT
TGACAATTTTGATGAGAAAATATCCTTGAATTTGACAATATTATTCGATATATTTTGTATGATCT
ACAAGATAATTACCAATTACGAAAAGTTTATGAAGATCAAACACAAGTTTAATTAATACTACAATTCT
AATTTCGAATATAAGCTTCTTGCATCATTTGCAACTCTCCTCGGTTATCAATAACACCCAAATGAA
CCAGAATGATTATATGAAAACAGATATTGATGAAAAGCTTGATCAGCTTTTCCACATCTATCAAA
CATTTTCCGGCTGTATCTGGATTTAGAAAAGTTTATGAAGTTCAAATTCAACTATCATGACTTT
GAGACAGAGTTTCAAGTCTCTCAATATCCAATATACTGAACACTCATGCTGCTTCTAACAATGA
CACAAATGCTGCTGATGCTATGAATGCCAAGGATGAAAAATATCTCCACAACCTTTGAATAGCG
TATTACTTGCTGATGAAGGAAATGAAAATTCGGTTCGTAATAACGATTCAGACCGCCTGTTTCATG
CTGAACGAGCTAATTAAATTTTGAAGTAGGTTTGAATTTCTCAAGATAGGTGAGTCATTTTGTGA
TTTCTTGATGAGAAATAACTACAAGTTTCATCCACTTCAAAAACCTTAAATGACGGAATGTTCCACA
TCAGGATATACCTAGAAAACCGACTAGATGGTGGTGTCTAG

[0206]

[0207] Mxr1 단백질 서열 (클로닝 동안 도입된 N-말단의 6개의 아미노산을 밑줄침) (서열식별번호: 2)

MRDRGRMSNLPPTFGSTRQSPEDQSPVPKELSFNGTTPSGKLRLFVCQTCTRAFARQEHKLRHE
RSHTKEKPFSCGICSRKFSRRDLLLRHAQKLHNSCSDAAITLRRKATRRESSNAAGSISGSTPVT
TPNTMGTPEDGEKRVQKLAGRRDSNEQKLQLOQQHLQQQPQLQYQQSLQHENVQVQPDQDPLI
SPRMQLFNDSNHVNNLFDLGLRRASFSAVSGNNYAHYVNNFQQDASSTNPNQDSNNAEFENIEF
STPQMMPVEDAETWMNNMGPIPNFSLDVNRNIGDSFTDIQHKNSEPIISEPPKDTAPNDKKLNGY
SFYEAPIKPLESLFSVRNTRKNKYKTNDSDPTVDNNSAPAANTIQELESSLNASKNFCLPTGYS
FYGNLDQQTFSTNLSTSSNATISPILLDNSINNNSTSDVRPEFRTQSVTSEMSQAPPPPQKNN
KYSTEVLFSTNMRSFIHYALSKYPFIGVPTPTLPENERLNEYADSFTNRFLNHYPFIHVITILKEY
SLFKAILDENESTKNWENNQFYLENQRISIVCLPLLVAITGAVLSNNKKDASNLYEASRRCIHVY
LDSRKKIPTSLANNNDSPWLIIQSLTSLVMYGLFADNDISLNVIRQVNALNSLVKTSGLNRTS
IIDLFNINKPLDNELWNQFVKIESTVRTIHTIFQISSNLSALYNIIPSLKIDDLMITLPVPTTLW
QADSFVKFKSLSYGNQIPFQYTRVLQNLIDYNQPLSDGKFLYENHVSEFGLICLQNLGHQYSYFQ
KLTAVERNREDALFTKVNSLHSDRMISNSDLFPKKIYQQSCLILDSKLLNNFLIVKSSSLKVSTG
DVSSLNKLKENVWLKNWNQVCAIYYNSFMNIPAPSIQKKYNDIEFVDDMINLSLIIKIMKLIFY
NNVKDNYEDENDFKLQELNLTDFNDFEKISLNLTLFDIFLMIYKIITNYEKFMKIKHKFNYYNS
NSNISFLHHFELSSVINNTQMNQNDYMKTDIDEKLDQLFHIYQTFFRLYLDLEKFMKFKFNHYHDF
ETEFSSLSISNINLTHAASNNDTNAADAMNAKDEKISPTTLNSVLLADEGNENSGRNNDSDRLFM
LNELINFEVGLKFLKIGESFFDFLYENNYKFIHFKNLNDGMFHIRIYLENRLDGGV*

[0208]

[0209] 피키아 파스토리스-코돈-최적화 LegH 핵산 서열 (서열식별번호: 3)

ATGGGTGCTTTACCGAGAAGCAGGAAGCACTTGTTCCTCTTCGTTTGAAGCTTTTAAGGCTAA
CATCCCTCAATACTCTGTTGTGTTTTACACGTCCATTCTAGAAAAAGCTCCTGCTGCCAAGGACC
TCTTCTCTTTTCTGTCCAACGGTGTAGATCCATCCAATCCCAAATTAACAGGTCACGCTGAGAAA
TTGTTCTGGTTTAGTCAGAGATAGCGCTGGACAATTGAAAGCAAATGGTACTGTGGTTGCTGATGC
TGCCCTTGGGCAGCATCCATGCACAGAAGGCAATTACAGACCCACAATTTGTTGTTGTGAAGGAAG
CTCTGCTTAAACCTATAAAGGAAGCCGTGCGAGACAAATGGAGTGACGAGTTGTGCATCAGCTTGG
GAGGTAGCTTATGATGAGTTGGCCGCAGCAATCAAAAAGGCATTCTAA

[0210]

[0211] 피키아 파스토리스-코돈-최적화 LegH 아미노산 서열 (서열식별번호: 4)

MGAFTKQEALVSSSF EAFKANIPQYSVVFYTSILEKAPAAKDLFSFLSNGVDPSNP KLTGHA EK
LFGLVRDSAGQLKANGTVVADAALGSIHAQKAITDPQFVVVKEALLKTIKEAVGDKWSELS SAW
EVAYDELA AAA
IKKAF

[0212]

[0213] 피키아 파스토리스-코돈-최적화 LegH 변이체 3 핵산 서열 (서열식별번호: 5)

ATGGGTGCATTTACAGAAAAACAAGAGGCTTTAGTATCCTCATCTTTTGAAGCTTTCAAAGCCAA
TATTCCTCAATACTCCGTTGTTTTCTATACGTCCATTTTGAAAAGGCTCCAGCAGCTAAGGACC
TTTTCTCTTTCTGTGCGAACGGCGTGGATCCCTCAAATCCTAAGCTGACTGGTCACGCCGAGAAG
CTTTTTGGTTTGGTCAGAGACAGCGCCGACAGCTGAAAGCTAACGGTACAGTTGTGGCAGATGC
TGCCCTTGGGATCTATACATGCACAAAAGGCTATCACCGACCCACAGTTTGTGGTTGTAAAAGAGG
CTCTACTCAAACTATCAAGGAAGCAGTTGGTGACAAATGGAGCGATGAATTGTCCAGTGCGATGG
GAGGTCGCTTACGATGAGTTAGCTGCTGCAATCAAAAAGGCTTTCTAA

[0214]

[0215] 피키아 파스토리스-코돈-최적화 LegH 변이체 3 아미노산 서열 (서열식별번호: 6)

MGAFTKQEALVSSSF EAFKANIPQYSVVFYTSILEKAPAAKDLFSFLSNGVDPSNP KLTGHA EK
LFGLVRDSAGQLKANGTVVADAALGSIHAQKAITDPQFVVVKEALLKTIKEAVGDKWSELS SAW
EVAYDELA AAAIKKAF

[0216]

[0217] 피키아 파스토리스 pA0X1 프로모터 (서열식별번호: 7)

GATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTCCAT
TCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAGCAGACCGTTGCAAACGCAG
GACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGCCCAGTTATTGGGC
TTGATTGGAGCTCGCTCATTCGAATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCT
GTCTATCCTGGCCCCCTGGCGAGGTTTCATGTTTGTATTTCGAATGCAACAAGCTCCGCATT
ACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCA
AATGGCCCCAAACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCAT
CCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAA
AAGTCGGCATACCGTTTGTCTTGTGTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAT
GCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGAA
ACACCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAA
TACTGCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATA
TATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATCATTATTAGCTTACTTTC
ATAATTGCGACTGGTTCGAATTGACAAGCTTTTGATTTTAAAGACTTTTAAACGACAACTTGAGAA
GATCAAAAACAACCTAATTATTCGAAACG

[0218]

[0219] 피키아 파스토리스 pGAP 프로모터 (서열식별번호: 8)

CGACTATTATCGATCAATGAAATCCATCAAGATTGAAATCTTAAATTTGCCCTTTCACTTGACA
GGATCCTTTTTTTGTAGAAATGTCTTGGTGTCTCGTCCAATCAGGTAGCCATCTCTGAAATATCT
GGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACCTAAATGT
GGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCTCTTCCACCGCCCGTTACCGTCCCTAG
GAAATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTGCAGCAATGCTCTTCCCAGCATTA
CGTTGCGGGTAAAACGGAGGTCTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCG
TCGCTGGCAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATAT
AAAAGGCGAACACCTTTCCAATTTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTT
GTCCCTATTTCAATCAATTGAACAACCTATCAAAACACG

[0220]

[0221] 피키아 파스토리스 pGCW14 프로모터 (서열식별번호: 9)

CAGGTGAACCCACCTAACTATTTTTAACTGGGATCCAGTGAGCTCGCTGGGTGAAAGCCAACCAT
CTTTTGTTCGGGGAACCGTGCTCGCCCCGTAAAGTTAATTTTTTTTTTCCCGCGCAGCTTTAATC
TTTCGGCAGAGAAGGCGTTTTTCATCGTAGCGTGGGAACAGAATAATCAGTTCATGTGCTATACAG
GCACATGGCAGCAGTCACTATTTTGTCTTTTAAAGTTCGTTTCATCAATCATTAACCTGACC
AATCAGATTTTTTTGCATTTGCCACTTATCTTAAATACTTTTTGTATCTCGCAGATACGTTCAAGT
GTTTCCAGGACAACACCCAAAAAAGGTATCAATGCCACTAGGCAGTCGGTTTTATTTTTTGGTCA
CCCACGCAAAAGAACCCACCTCTTTTAGGTTTTTAAAGTTGTGGGAACAGTAACACCGCCTAGAG
CTTCAGGAAAAACAGTACCTGTGACCGCAATTCACCATGATGCAGAATGTTAATTTAAACGAGT
GCCAAATCAAGATTTCAACAGACAAATCAATCGATCCATAGTTACCCATTCCAGCCTTTTCGTCTG
TCGAGCCTGCTTCATTCCTGCCTCAGGTGCATAACTTTGCATGAAAAGTCCAGATTAGGGCAGAT
TTTGAGTTTAAAAATAGGAAATATAAACAAATATACCGCGAAAAAGGTTTGTATTATAGCTTTTCGC
CTGGTGCCGTACGGTATAAATACATACTCTCTCCCCCCCCCTGGTTCTCTTTTTCTTTTGTACT
TACATTTTACCGTTCCGTCACCTCGCTTCACTCAACAACAAAA

[0222]

[0223] 피키아 파스토리스 pTEF1 프로모터 (서열식별번호: 10)

ATAACTGTGCCTCTTTTATCTGCCGCACTGCATGAGGTGTCCCCTTAGTGGGAAAGAGTACTGA
GCCAACCTTGAGGACAGCAAGGGAAAAATACCTACAACCTTGCTTCATAATGGTCGTAAAAACAA
TCCTTGTCGGATATAAGTGTGTAGACTGTCCCTTATCCTCTGCGATGTTCTTCTCTCAAAGTT
TGCGATTTCTCTCTATCAGAATTGCCATCAAGAGACTCAGGACTAATTTGCGAGTCCCACACGCA
CTCGTACATGATTGGCTGAAATTTCCCTAAAGAATTTCTTTTTTACGAAAATTTTTTTTTTACAC
AAGATTTTACGACAGATATAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTTTCTTTT
ATTCTCACTACATACATTTTAGTTATTTCGCCAAC

[0224]

[0225] 헴 생합성 효소 1- ALA 신타제 (서열식별번호: 11)

ATGGAGTTTGTGCGCCGTCAGTCCATGAATGCCTGTCCCTTTGTCAGGTCAACTTCTACCCACCA
TTTGAAGAAGTTGGCAGCAAACAGTTCTCTAGCTGCTACTGCTAGTCATTGTCCCGTGGTTGGCC
CTGCTCTCCAACAGCAGAGATACTACTCTCAACCTTCCAAGCCAGCCCAAGCCCAAACCTCCGAC
ATTGCTACTGGGATCAAGAAGGATGTTTCTCCGATCCGTATGGACTCTAATGAAACCGCCTTTGA
TTACAATGGAATGTATGAGTCTGATCTTGCGAATAAACGTAAAGATAACTCGTATCGTTATTTCA
ATAACATCAACCGTCTAGCCAAGGAGTTTCCAAGGCACATCGCCAGACCGAAGATGACAAGGTG
ACCGTCTGGTGCTCTAACGACTACTTAGGAATGGGTAGGCATCCTGAGATTATCAAAACCATGAA
GGCTACCATGGACAAGTACGGTTCCGGAGCAGGAGGAAC TAGGAACATTGCAGGTGATAACCACG
CCGCTATCAATTTGGAAGCGAGTTGGCTTGCTTGAACAAGAAGGAAGCGGCTCTGGTGTTTTCA
TCATGTTTTCATAGCTAACGATGCAATCATCTCGTTGTTGGGACAAAAAATCAAAAATTTGGTCAT
TTTCTCTGACCAGTCGAATCATGCTTCCATGATATTGGGTGTGCGTAAC TCCAAAGCGAAGAAGC
ACATCTTCAAGCACAACAATTTGAAGGATCTGGAGTCGCAGTTAGCTCAGTACCCCAAGTCGACT
CCTAAACTGATCGCCTTCGAGTCAGTTTACTCTATGTGTGGATCTGTGGCTCCCATTGAGAAGAT
TTGCGATTTGGCTAAAAGGTACGGTGCCCTCACCTTCTTGGATGAAGTTCATGCTGTTGGAATGT
ATGGTCCTCATGGACAGGGTGTAGCTGAGCATTTGGACTTTGATCTGCATTTACAGTCTGGAATC
GCCAGTCCTAGCGTGGTGGACAAACGCACCATATTGGATCGTGTGCGACATGATTACTGGTACTTG
CGGAAAGTCATTTGGTACTGTTGGAGGTTACGTTGCTGGTAGTGCCAACCTAATTGATTGGTTAA
GATCCTATGCGCCAGGTTTCAATTTTCACTACCACACTTCTCTGCTATCATGGCTGGTACAGCC
ACTTCTGTTTCGTATTGTTAGGGCCGACATTTGAGGCCCGTATCAAGCAACAGCTTAATACTCGCTA
CGTCAAAGACTCATTTGAAAACCTTGGTATTCCAGTCATTCCAAACCCAAGTCACATTGTTCTCTG
TTCTAGTTGGAAATGCTGCAGATGCCAAGAAGGCATCCGATATGTTAATGAACAAACACCGTATT
TATGTTCAAGCTATTAATACTACCTACTGTGCCTGTGCGGTGAAGAACGACTAAGGATTACTCTAC
TCCAGGTCATGGAAAGGAGATTTGTGACCAGCTGATCAGCGCTGTGACGATGTTTTTACTGAGC
TTAATTTACCAAGAATCAACAAATGGCAGTCCCAAGGTGGTCATTGCGGTGTTGGTGATGCTAAT
TACGTACCAGAACCCAACTCTGTGGACTCAGGACCAGCTCAGCTTGACAAACCAAGACTTGCACTC
CAATGTGCACAACCCAGTGATTGAGCAGATCGAAACCTCATCAGGAGTCAGATTGTAG

[0226]

[0227] 헴 생합성 효소 2- ALA 데히드라타제 (서열식별번호: 12)

ATGGTGCATAAGGCTGAATACTTGGACGACCACCCAACCTCAGATTTCCAGCATTCCTTTTCAGGAGG
TTACAACCACCCATTACTTTCGTGAATGGCAACATGAACGTCAACTCAACAAAAACATGTTTCATCT
TTCCCTGTGTTGTCACAGATCGACCAGACGAAGAAGAACTTATTCCTAGTCTACCTAATATCAAG
AGGTTTGGCGTTAACAAGTTGATTCCTTATGTAGGAGGTTTGGTTTTCCAAAGGATTGAGGGCGGT
GATCCTATTTGGTGTTCTCTGAAGCCCGGTGTGAAAGATGAAGAAGGAACGGCCGCTGATGATC
CAGAGGGACCTGTTATCCAAGCCATCAAACACTTGAGAAAGAACTTTCTTGACCTGTATATCATC
ACCGATGTCTGTCTATGTGAGTACACCAGCCATGGACATTGTGGAATACTATATGAGGATGGCAC
TATCAACAGAGAGCTCTCAGTCCGTCGTATTGCTGCTGTAGCTGTCAAATATGCTCAAGCTGGAG
CCAACCTCTGTGGCTCCTTCTGATATGACTGACGGCAGAATAAGAGATATTAAAGAAGGCTTACTA
AGTGCAGGACTGGCACATAAAACGTTTGTATGTCTTACGCTGCAAAATTCTCTGGTAATTTGTA
TGGCCCTTTTCAGAGATGCTGCAGGTTCTGTCCATCTCAAGGGGACAGAAAATGTTACCAGCTTC
CTTCTGGAGGAAAAGGTTGGCCCATCGTGCTCTGATTCGTGATATGAATGAAGGCATGATGGA
ATTATTGTCAAACCATCTACATTCTATTTGGACATTGTGCTGATGCTTATCAGCTTTGTAAAGA
CTATCCTATCTGTGTTACCAGGTTTCTGGAGAGTACGCCATGCTACATGCAGCGGCAGAGAAGA
ATATTGTTGATCTGAAATCAATCGCTTTTGAAGCTCATCAAGGATTCTTGCGGGCTGGAGCTCGT
TTAATCATTAGTTACTTTACCCCTGAATTCCTGGAGTGTTATCTGAATGA

[0228]

[0229] 헴 생합성 효소 3- 포르포빌리노젠 데아미나제 (서열식별번호:

ATGAACCAAATCGAACAGAGCGGACCCATTGATTGCAGTTCCTTGAAATTGGGGTCCCCGAAAAGTC
CGCTCTGGCTATAATCCAGGCAGAAATCGTCCGCCAATTGATATTGAAAGAATACCCTGAATTGG
AGACGAAGTTGGTCAGTGTGTCCACCCTGGGGGACCAAGTCCAGAATAAAGCACTTTTCACGTTT
GGAGGAAAATCTTTGTGGACCAAGAAGTGTGAGATGTTGTTGTTGGAGAGTGTGGGAGGATTTGA
CCAAATAGACATGATTGTACACTCGTTGAAAGACATGCCAACTCATTTACCAGACGAATTTGAGC
TGGGTTGCATTATTGAAAGAGAAGACCCTAGAGACGCTTTGGTTCGTGCAAGATGGTTTATCTTAC
AAGTCATTGGCCGACCTTCCAGAGGGAGCTGTGGTTCGGTACGTCTTCGGTTAGAAGATCGGCTCA
ACTACTGAAGAATTTCCCTCATCTGAAATTCAAATCTGTTAGAGGAAACCTTCAGACCAGACTAA
GAAAATTAGATGATCCAGATTCCGAGTACTGCTGTCTCCTCCTTGACAGCAGCCGGTTTTATCAGG
ACAGGCTTACAACACAGAATTTCAATGTATTTGAACGACGATGTGATGTACCACTCCGTCGGACA
AGGAGCATTAGGAGTAGAGATCAGAAAAGGTGACCAATTCATGAAAAATATCTGTGAAAAGATTG
GGCATAGAACCACCACCCTTCGTTGTCTTGCAGAGAGAGCACTGCTGAGATATCTAGAGGGAGGC
TGCTCGGTGCCAATTGGGGTCTCCACTATTTATAGCGAGGATACGAAGGAACCTTACCATGAACCT
CCTAGTCGTGAGTTGTAACGGTCGTGACTCGGTAACAGAATCAATGACTGAAGTCGTGACTACTG
AAGAGCAAGCTGAAGATTTCCGGTGAAGGCTGGCCCAGAAGCTCATAGATCAAGGTGCGAAACGC
ATTCTTGACGAGATCAACTTCAACAAGATCAAAGAGATTAAGGAAGAGGGTTTACATTAA

13)

[0230] 헴 생합성 효소 4- 우로포르피리노겐 III 신타제 (서열식별번호: 14)

ATGCCAAAAGCCATTCTTCTGAAGAATAAACTACACCGAAGGATCCTTATCTGGAGAAGTTTCGT
AAGTAGTGGCTACTCGACCGATTTTCGTACCCTTTTAGATCATATTCACATGGAGAAATCTGAGA
TCATCGCATTTCTCAAGACTGACTACTTTTTGCATAAACTTTGGCGTTTATTATTACGTCCCAA
AGAGCTGTAGAAATGCTGAATGAGTGTATGCAAATACTGAGACGTACTGATCCTGAAATTACACA
AATCATCTATAGTAAACCTGTCTATACAGTTGGCCCTGCCACCTACAGAATACTTGCGGATGCTG
GCTTCGTGGATCTACGAGGCGGAGATAAGGCAGGAAACGGATCCATTCTAGCCCAGATAATTTTG
AATGATGACATTTACACTGGAATTGAAGATTCTGACAAGCATATAACGTTTTTACGGGAGAAAC
AAGGAGAGACATAATTTCCAAATGTTTACTCTCTAACAACCTTTCACTTTACGAAAAGATTGTCT
ACAAGACTCTTCCTAGGGATGATATCGTGACTAGATTCAAGTCTGCCGTTGACAGCATCGACCAA
TCGCAAGAAGTTCCAGTTGGGTGGTCTTCTTTTCGCCCTCAAGGAACAGAGGACATTGTAACGTA
TCTTCAACACACCAAAGACCAATTTAATATTGCATCTATCGGGCCAACCACAGAAAAATACCTTC
TAAGCAAAAACCTGAAACCAAAGTTGTGGCACCTAAGCCAGAGCCTATCTCTTTACTATTGTCT
ATACAAAAAGTGCACTAA

[0231]

[0232] 헴 생합성 효소 5- 우로포르피리노겐 III 데카르복실라제 (서열식별번호: 15)

ATGAGTAGATTTCCAGAACTGAAGAATGACCTTATTTTAAGGGCAGCTCGTGGTGAAAAAGTTGA
ACGTCCCCCAATATGGATTATGAGACAGGCCGGAAGATATCTTCGGAGTACCATGAGGTCAAAG
GAGGTAGGGACTTCTTTGAAACTTGCAGGGATGCTGAGATTGCTTCTGAAATTACTATCCAGCCG
ATTACGCATTTTGACGGTCTGATCGATGCAGCTATTATCTTCAGTGATATCTTGGTGATTCTCA
AGCTATGGGCATGGAAGTTAAGATGGTGGACAAAGTTGGCCACAGTTCCCCAATCCGCTAAGAA
AACCGTCTGACTTGGATCATTTGAAAAAAGACGTTGACGTTTTGAAGGAACCTCGATTGGGCCTTC
AAAGCTATCTCATTGACCAGAAAAAACTCAATGGACGAGTGCCTTTGCTTGGATTTTGTGGTGC
TCCTTGGACTCTACTGGTTTATATGACTGAAGGAGGCGGTACCAAGATGTTTCGATTTGCAAAAG
AGTGGATCTACAAGTTTACCAAGGAATCTCATCAATTACTCCAACAGATCACTGACGTTGCAGTT
GAATTCTTAGCTCAGCAAGTTGTTGCAGGTGCCCAAATGTTACAAGTTTTTGAATCTTGGGGCGG
TGAATTGGGGCCTGATGAATTCGATGAGTTTTCTTTGCCTTATTTGAGACAGATTTCTCTAAAC
TTCCCCCTGAGGTTGAAGGAACCTGGAATCACAGAGAATGTTCCATAACTGTCTTTGCTAAAGGC
TCTTGGTACGCCTTGGAGCAATTGTGCGACAGTGGTTATGATGTTGTCTCGTTGGATTGGTTATT
CCGTCCAAGTGATGCTGTCCAGATTGCTAACGGAAGAATCGCATTGCAAGGTAATCTTGACCCTG
GAACCATGTACGGCTCCAAAGAAACCAATTTCCAAGAAAGTGGACAAAATGATCAAGGGTTTTGGT
GGAGGAAAAGCAAACTACATAATTAATTTTGGACACGGCACTCATCCATTCATGGATCCAGAACA
GATCAGATGGTTCTTACAAGAATGTCATCGCATTTGGATCTCAATAG

[0233]

[0234] 햄 생합성 효소 6- 코프로포르피리노겐 III 옥시다제 (서열식별번호: 16)

ATGGCCATCGACTCTGATATCAATCTAAGCTCTCCCAATGATTCCATCCGTCAAAGGATGTTCTGA
GCTTATCCAGCGGAAGCAACTCGAAATTGTCGCTGCATTGGAGGCAATTGAAGGAAACGATACCA
AATTTCTGTTCTGATTCTTGGGAAAGAGGAGCCGAAGGTGGAGGAGGAAGATCTATGCTTATTCAA
GATGGAAGAGTGTGTTGAAAAGGCTGGTGTAATATTTCCAAGGTTTCATGGCGTATTGCCTCCTCA
AGCTGTGAGCCAGATGAGAAATGACCACTCCAAGCTAGATCTGCCTGCGGGAACCTCTCTGAAGT
TCTTTGCCTGTGGGCTTTTCGTTGGTCATTCATCCCCATAATCCCCATGCTCCAACCTACCCATCTG
AATTATCGCTACTTCGAAACTTGGGATGAACTGGAAAGCCTCACACCTGGTGGTTTGGGGGCGG
TGCTGATTTAACGCCCTTCGTACCTGTATCCCGAGGATGCCAAGCAATTCATCAAGCCCATAAGG
ATGCCCTGGACAAACACGATGTTAGCTTGTACCCGAGATTCAAAAAGTGGTGTGATGAATACTTT
CTGATCAAAACATCGAAATGAAACTAGAGGTATTGGGGGTATTTTCTTTGATGATTTTGACGAGTT
TGATGCTGAGAGGTCCCTGAAGTTGGTTGAAGATTGTTTCAATGCTTTCTTGGAATCTTATCCCG
CTATCACTCGAAAAAGGATGGACACCCCTTCAACTGATGCTGAGAAGAACTGGCAACAAATTAGA
AGAGGAAGATATGTCGAATTCAACTTAGTATTGGATAGAGGTACTCAATTTGGTTTGAGAACGCC
TGGATCTCGTGTGAAAGTATTTTGATGTCGTTGCCAAGAACAGCTGGTTGGGTCTATGATCATC
ATCCAGAGCCTGGCTCCAGAGAAGAGGAGTTATTGCAGGTACTACAAAATCCTATTGAATGGGTA
TGA

[0235]

[0236] 햄 생합성 효소 7- 프로토포르피리노겐 옥시다제 (서열식별번호: 17)

ATGCTGAAAAGTCTTGCACCAAATTCCTCAATTGCCGTTTTAGGTTTCAGGGATATCTGGATTGAC
TTTCAGCTTTTTTTTTGAATCGGTTGCGTCCCGATGTTAAGATCCATATCTTTGAAAAATCCAAGC
AGGTTGGAGGATGGATCAGATCAGAAGAGCATGAACTTTTCATTTTGAAAAGGGACCCAGAAGT
TTGAGAGGCACAAATACGGGTACCTTGATGTTGTTGGATCTTCTTACCAAGATAGGAGCAAATGA
CAAGGTCCTGGGACTGCACAAAGATTCTCTTGCTAATAAAAAGTATCTGTTGTCCCCGTTCTCAG
ATGTTACGGAACAACGCAAAGCTTCTTCAAGTGCCACAGGATTTTCAGCTCTTTTGTAAGTTTC
ATGTTTGACCCGTTGTCTAAGGATCTCATCTCGGTCTTTTGAAAGAACCATGGCAACCAAATTT
AAAGTATTCAGATGAGTCGGTTGACCATTTTTTCAACAGAAGATTTGCTACCAAACATATCAGAGA
ATATCGTCAGCGCAATTGTGCATGGAATCTATGCGGGCGACGTGAAGAAGTTAAGTGTGAAAGCC
ATCTTCCCTAGGCTCCCTGAGATGGAACAGGAAAGTGGCTCTATTATAAGGTATATGATCGCCCCA
ATACAGGACAAAAAAGAAGCTCAAACAAAAAGTTGACCCTTTTTTGGCAGATTATGAAAAATTGA
TCGGTACATCTTTGAGTTTCAAAAATATTTCTTTGTTTCTGAAAAACTTTCCCATGCTGAGTTTT
CAGGGTGGACTACAGAACTTCCCATCTCATTTGAAGAACCATTTATCACAGATTGAAAACATCAA
GTTTCATTTTGACAGCAAAATCAAAAACATTGCTTTGGAGAGCGTAAGGTGGCATTGACTGACC
ATGATCAGGTTTATCTTGTGACCATGTGAGATCTACCATTAATACCAACGAATTGGCCAAAATC
ATTTACCCCGTTGTTCCAAGTTCTACTAAGAAAAAATCCGTTTTCAAATCCAAGCGAATTGGCCC
AGGGCTGGTCAAATGTTTGAGCTGGCTACACTATACAAATATACTAATGTGCAACATTTATATAC
CTAAGCACGTCTCAAAATCTATCACCGGATTTGGATACTTGGTTCCCTCGATCAATGTCTTCTCAG
GCATCCAAACTTCTCGGTGTCATATTTGACTCAGACATCGAGACTGCAATGACTCCTAATTTTAC
AGAGGCCAACATTACGGCGATAAACAGTAACCTCTGCATCTCCAAGCAACTCCAAAAGTTTTCTG
ACCAATTCGTCAATAATGATCTCCCTAAATACACCAAGTTGACGCTAATGCTTGAGGTCATTAT
CTCAAGTCGGAGGCAGACATGCCCGGTTCCGCAGAGAGTAAACATGCTGTCAAGGCGATTCTGTC
AAATCACCTGAATATTGATCTAGATGAGTTTGCATCTTTGCCAGACTTCAAGATGGAAATCACCA
AGATCCCCAACTGCATTCCCCAATATGAAGTTGGGTATCTTGATCTCAAGAGAAAGGTTTCAGAAT
GCAGCCTCCAAAGAGTTCAACGACCAATAAGTTTTGGAGGCATGGCATTGTTGGTGTGTTGGG
GATCCCTGACTGTGTCCAGAATGCATTCAAAGATTTCGGCTACCCTCAGTGGCATTTAA

[0237]

[0238] **헴 생합성 효소 8- 페로킬라타제 (서열식별번호: 18)**

ATGCTTAACCGTCGTTTCCAATCTACCGTGTCTCGAGTCTGAACAAGGGCACTGGAATAGTGTT
CATGAATATGGGTGGTCCCTCCACTGTCAAGGAAACCTATGACTTTTTATTTCGTCTTTTCTCGG
ACGGAGATTTAATCCCGTTTGGCAGATTTTCAAGACATCCTGGCCCGCTTCATTGCAAGTAGAAGA
ACACCCAAAATTGAATCCTACTACAAAGCTATCGGAGGTGGGTCTCCTATCCGAAAGTGGTCTGA
ATACCAGAGTTCTAAACTATGTGAAAAATTAGACATTATCAGTCCACAATCGGCTCCTCATAAGC
CTTATGTTGCCTTCAGATACGCTAATCCTCTCACTGAAGATACTTTACAAAAGATGAAAAATGAT
GGAATTACTAAGGCCATTGCCTTTTCTCAATATCCGCAATTTAGTTATTCAACCACCGGATCATC
GATTAACGAACTTTACAGGCAATCGAAAATTTTGGACCCTGATCAATCTATTAAATGGACAGTTA
TAGATCGCTGGCCTGACCACCCAGCCTTAGTTAAACTTTTCGCAGCTCATATCAAAGATACTCTA
AACAGATTCAAAACCTGAAAATGGACTGACTGACACAAAAGACGTCGTCCTCCAATTCAGTGCTCA
TTCTTTACCAATGGATATTGTCAATAAAGGAGATTTCGTATCCTGCAGAAGTCGCAGCGAGTGCTC
TTGCCATTATGAAAGAACTTAACCTTCTCAAATCCTTATAAATTAACCTGGCAATCACAGGTTGGC
CCAAAGCCTTGGCTGGGTGCTCAAACCTGAAAAAATTACCAAGCAGCTAGCATCCAGTGATGTTCC
TGGAGTCGTTTTGGTTCCCTATTGCCTTTACCTCTGATCATATTGAACTCTCCATGAATGGATA
TTGAACTGATTCAAGAAGTACCTAATCCTTCAAAGTAAAGCGAGTTGAATCGTTGAACGGAGAC
CAAACTTTCATTGACTCCTTGGCAGAACTAGTGAAGAGTCACATTGATTGGAAGGTTGTATTTTC
CAACCAGTTGCCATTGGATTCCATGCTGGGAGTAGTGTCAGATAATTCCCTCACAGATCCAAAAG
AGTTTTTTCAGAGCCCATTTGA

[0239]

[0240] **파트 C. 결과 및 논의**

[0241] 실시예 24 - 균주 MXY0183의 특성화

[0242] 균주 MXY0183에 대한 최적 성장 조건은 3.0 내지 6.0의 표적 pH 및 28-35℃의 온도를 포함한다. LegH 단백질을 생산하기 위해, 균주 MXY0183은 6일의 기간 동안 살아 있어야 하고 호기성으로 성장 중이어야 한다.

[0243] 균주 MXY0183과 연관된 유전자의 발현은 균주에 대한 표현형 변화를 초래하였다. 도 8은 유도 시작 (0 hr) 및 유도 72 hr 후의 진탕 플라스크의 사진을 나타낸다. #1로 지정된 플라스크는 숙주 균주인 MXY0051을 함유한다. #2 및 #3으로 지정된 플라스크는 각각 중간체 균주 (즉, MXY0118이고, LegH 유전자의 >10개의 카피 및 헴 생합성 경로로부터의 ALA 데히드라타제를 함유함) 및 생산 균주 (즉, MXY0183이고, LegH 유전자의 >10개의 카피 및 헴 생합성 경로로부터의 8개의 효소를 함유함) 중 하나를 함유한다. 72시간 후의 플라스크 #3에서의 특징적인 적색은 헴-결합된 LegH의 생산을 증명한다.

[0244] 진탕 플라스크에서의 성장 후, 상기 지시된 피. 파스토리스 균주 MXY0051, MXY0118, 및 MXY0183을 용해시키고, 단백질을 SDS 겔 상에 러닝시켰다 (도 9A). 화살표는 LegH 단백질의 위치를 나타낸다. 균주 MXY0183 및 균주 MXY0118에서의 LegH 생산의 비교가 도 9B에서 제시되고, 이는 MXY0183 균주에 의한 LegH 단백질의 헴 로딩의 효율을 증명한다.

[0245] 실시예 25 - 균주 MXY0207의 특성화

[0246] 그 후, 헴 생합성에서 수반되는 8개의 효소를 코딩하는 유전자의 존재 하에 전사 활성화제인 Mxr1을 과발현시키는 것의 이점을 결정하기 위해 실험을 수행하였다. LegH 서열의 >10개의 카피 및 헴 생합성에서 수반되는 8개의 효소를 코딩하는 유전자를 함유하는 균주 MXY0183, 및 LegH 서열의 >10개의 카피, 헴 생합성에서 수반되는 8개의 효소를 코딩하는 유전자, 및 Mxr1 전사 활성화제를 함유하는 자매 균주 MXY0206 및 MXY0207을 이들 균주에 대한 억제성 탄소원인 글리세롤의 존재 하에 진탕 플라스크 배양물에서 성장시켰다. 48 hr 후의 진탕 플라스크 배양물의 사진이 도 10A에서 제시되고, 추가적인 탄소원 없이 48시간 동안 BMY 배지에서 성장된 세포로부터의 펠렛의 사진이 도 10B에서 제시된다; 이들 실험은 AOX1 프로모터의 제어 하의 트랜스진 (예를 들어 헴 효소)의 유의한 발현이 Mxr1이 AOX1 프로모터로부터 또한 발현되는 균주의 성장 배지에서 억제성 탄소원이 소비될 때 유도성 탄소원의 부재 하에 발생한다는 것을 증명하였다. 진탕 플라스크 배양물이 유도제의 부재 하에 성장되었을 때의 헴-로딩된 LegH의 상대적인 수율이 도 10C에서 제시된다. 이들 실험은 재조합 헴-로딩된 단백질의 유의한 생산이 Mxr1 발현이 AOX1 프로모터에 의해 또한 구동되는 피키아 균주에서 메탄올 유도의 부재 하에 AOX1 프로모터-구동 트랜스진으로부터 달성된다는 것을 증명한다.

[0247] 선택된 균주를 2 L 발효기 탱크에서 성장시키고, LegH 및 헴-로딩된 LegH의 상대적인 수율을 결정하였다 (도 11). 균주 MXY0183에 비교하여, MXY0207 균주가 더욱더 많은 LegH를 생산하였고, 매우 효과적으로 LegH 단백질을 헴-로딩하기 위한 충분한 헴을 생산할 수 있었다.

[0248] 실시예 26 - 균주 MXY0291의 특성화

[0249] 상기 실시예 18-20에서 기술된 바와 같이, 균주 MXY0291은 항생제 저항성 유전자가 없으면서 MXY0207가 LegH를 생산하는 능력을 재현하도록 구축되었다. 균주 MXY0291이 LegH var3의 ~ 16개의 카피, Mxr1, 및 8개의 헵 생합성 효소 중 7개를 함유한 것으로 결정되었다. 2 L 발효기 탱크에서 성장시켰을 때, 이러한 균주는 MXY0207에 비교하여 개선된 LegH 수율을 나타냈다. 이러한 개선은 메탄올 / 글리세롤 및 메탄올 / 텍스트로스 (D-글루코스)를 함유하는 유도 배지 둘 다에서 나타났다 (도 11).

[0250] 실시예 27 - 하이브리드 프로모터 균주의 특성화

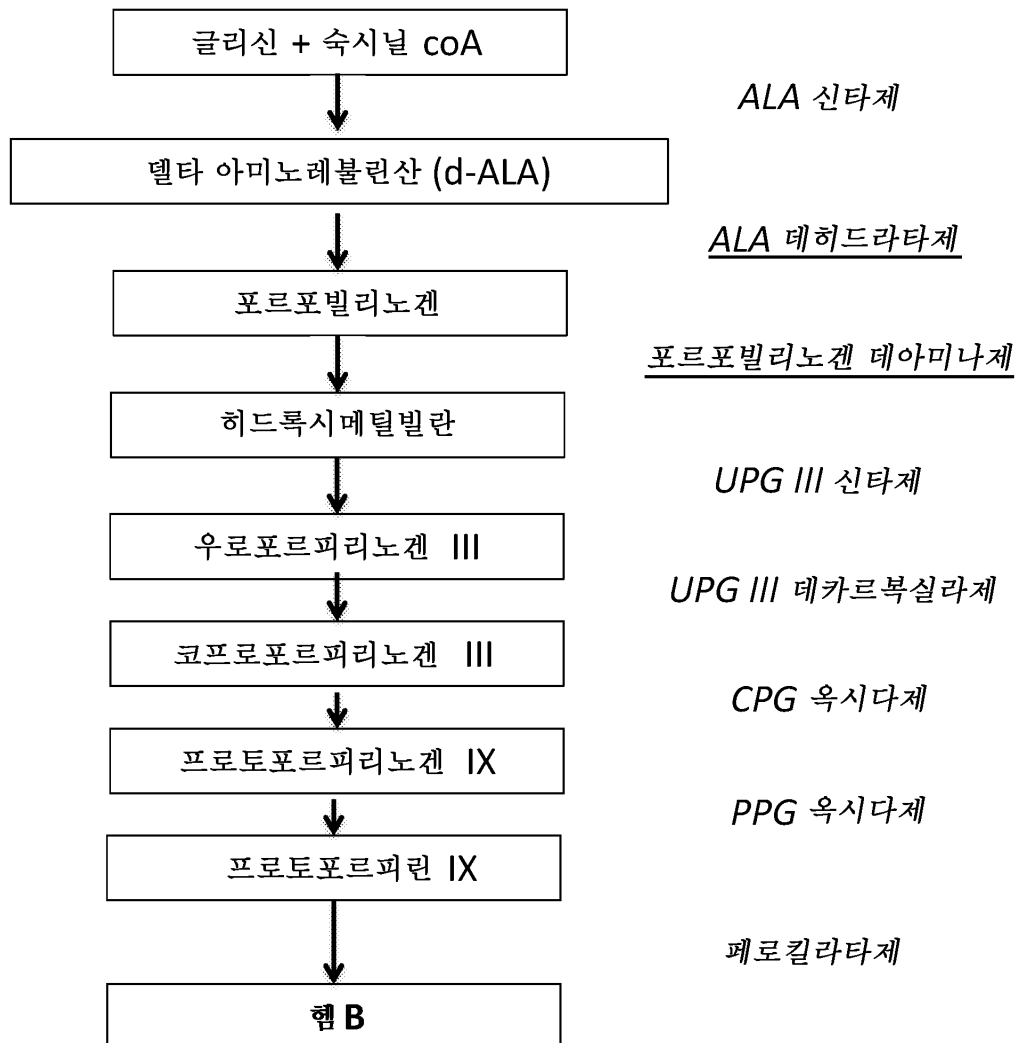
[0251] 프로모터 pAOX1의 제어 하의 LegH의 여러 카피, 모든 헵 생합성 효소, 및 전사 인자 Mxr1를 이미 함유하는 균주 (상기에서 MXY0291로 지칭됨)에서, 대두 레그헤모글로빈 (LegH)의 추가적인 카피를 3개의 상이한 구성적 프로모터인 pGAP, pGCW14 및 pTEF1 하에 발현시켰다. 텍스트로스 (즉, D-글루코스)의 존재 하에 메탄올로 유도되었을 때, 구성적 프로모터 및 pAOX1가 LegH의 발현을 구동시키는 반면, pAOX1 프로모터만 헵 효소의 발현을 구동시킨다. 이는 이전의 균주 MXY0291에 비교하여 LegH 수율에서의 추가적인 개선에 이른다 (도 11).

[0252] 사안 방법 및 조성물이 다수의 상이한 측면과 연계되어 본원에서 기술되었지만, 다양한 측면에 대한 상기 설명은 사안 방법 및 조성물을 예시하도록, 그리고 그의 범주를 제한하지 않도록 의도된다는 것을 이해하여야 한다. 다른 측면, 장점 및 변형이 하기 특허청구범위의 범주 내에 속한다.

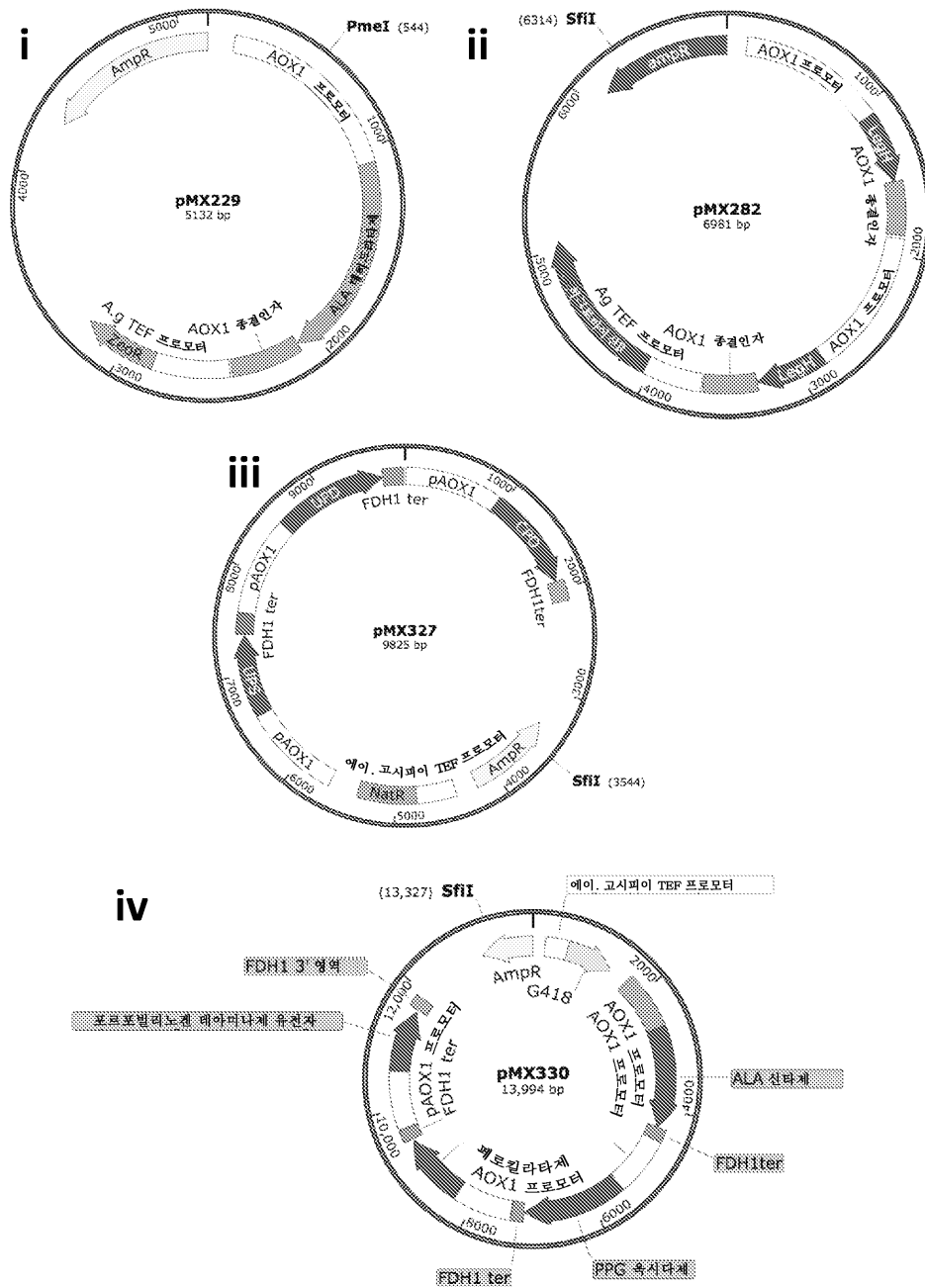
[0253] 개시된 방법 및 조성물에 사용될 수 있거나, 이와 연계되어 사용될 수 있거나, 이에 대비하여 사용될 수 있거나, 또는 그의 생성물인 방법 및 조성물이 개시된다. 이들 물질 및 다른 물질이 본원에서 개시되고, 이러한 방법 및 조성물의 조합, 부분집합, 상호작용, 군 등이 개시되는 것으로 이해된다. 즉, 이들 조성물 및 방법의 각각의 다양한 개별적 및 총괄적 조합 및 치환에 대한 구체적인 언급이 명시적으로 개시되지 않을 수 있지만, 각각이 본원에서 구체적으로 구상되고 기술된다. 예를 들어, 특정한 사안 조성물 또는 특정한 방법이 개시 및 논의되고, 다수의 조성물 또는 방법이 논의되면, 구체적으로 달리 대조적으로 지시되지 않는 한 조성물 및 방법의 각각의 모든 조합 및 치환이 구체적으로 구상된다. 유사하게, 이들의 임의의 부분집합 또는 조합이 또한 구체적으로 구상되고 개시된다.

도면

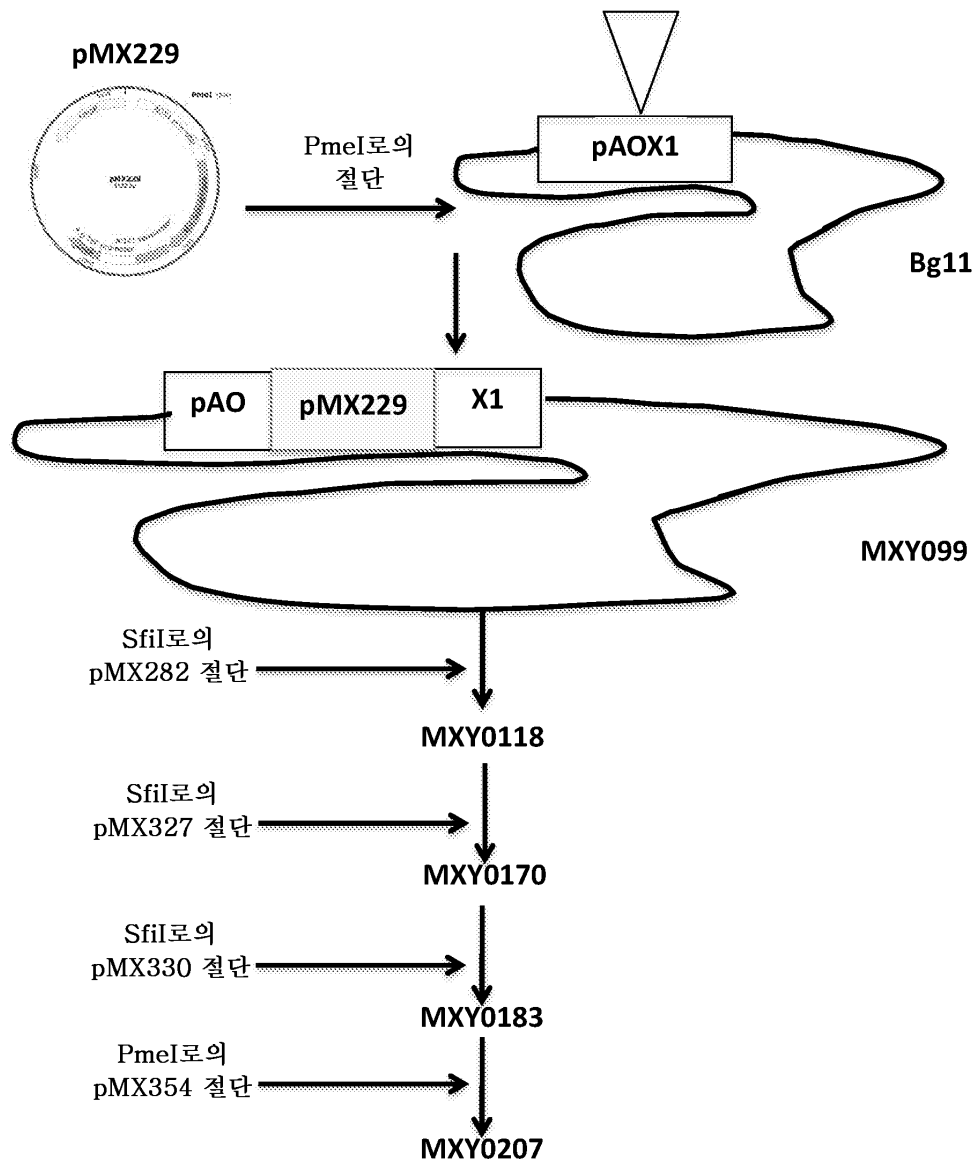
도면1



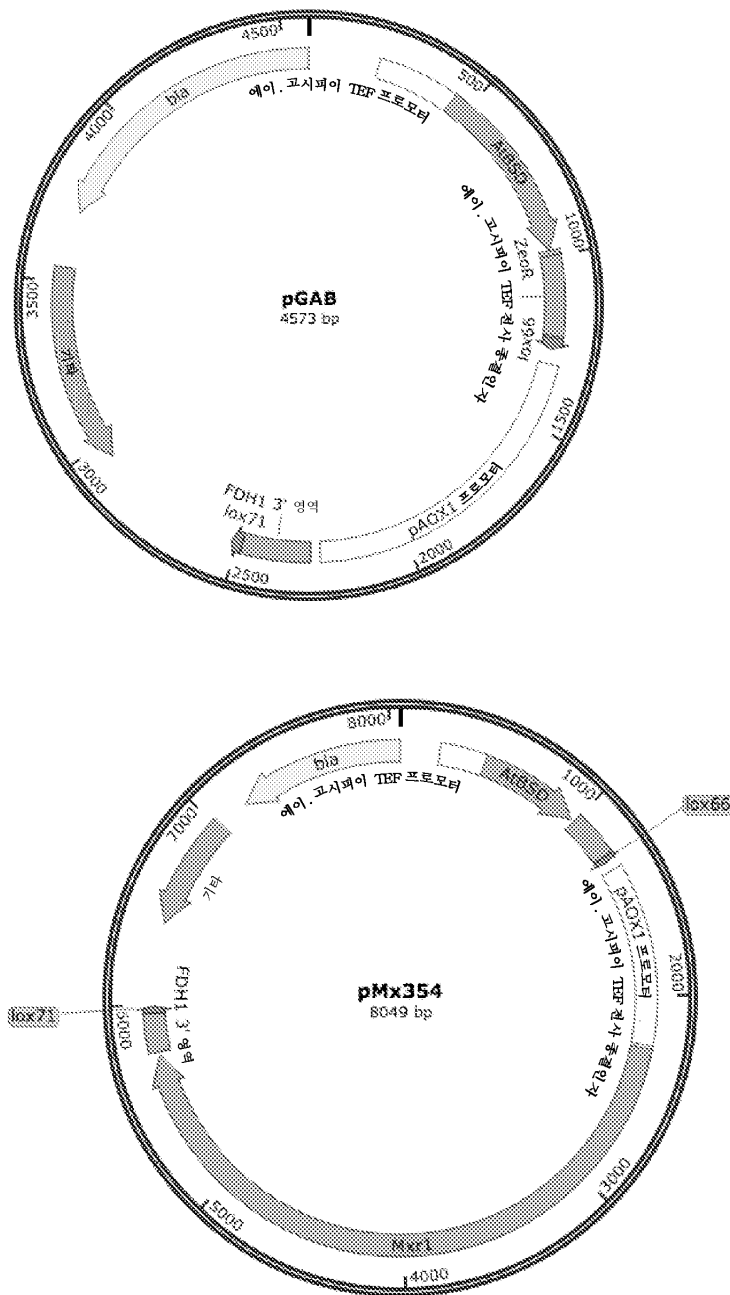
도면2



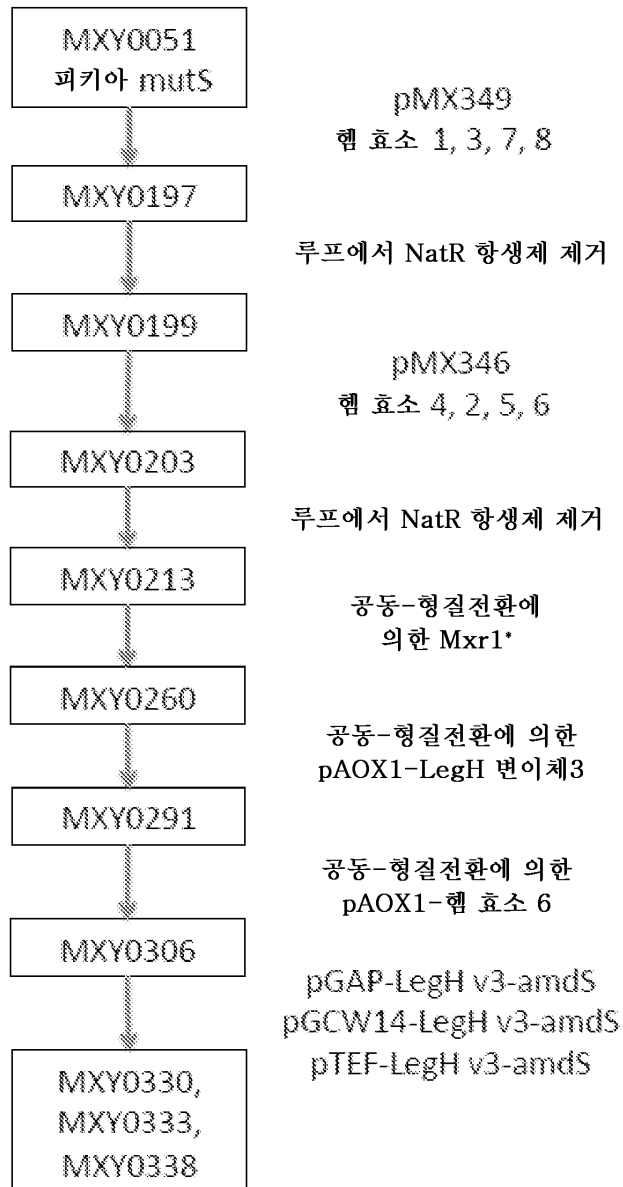
도면3



도면4

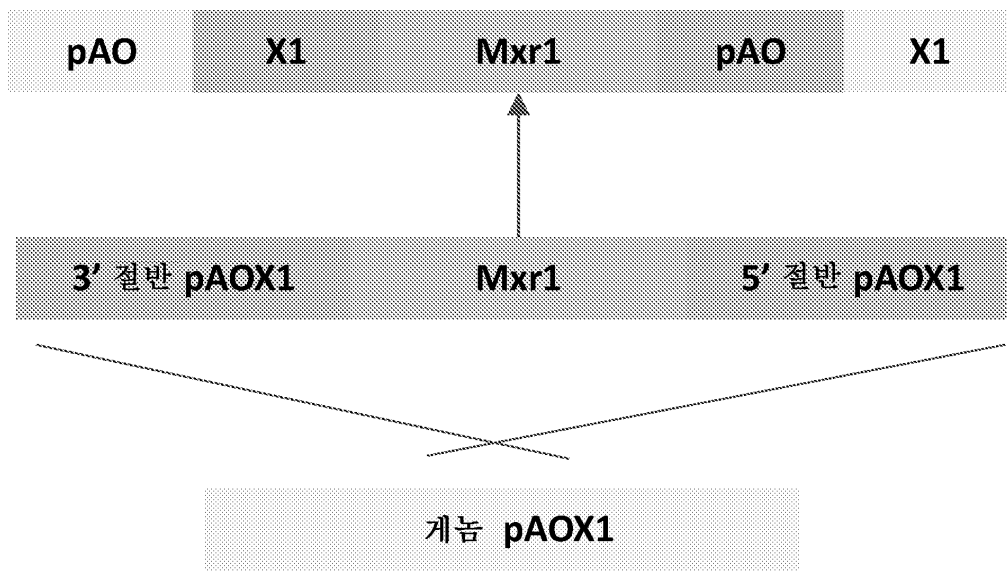


도면5

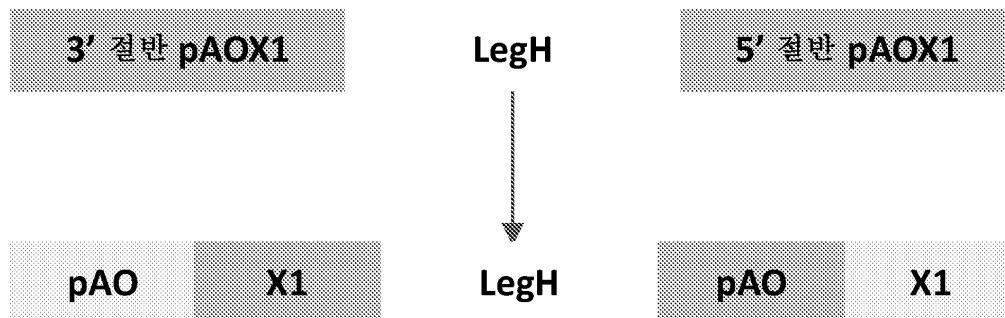


도면6

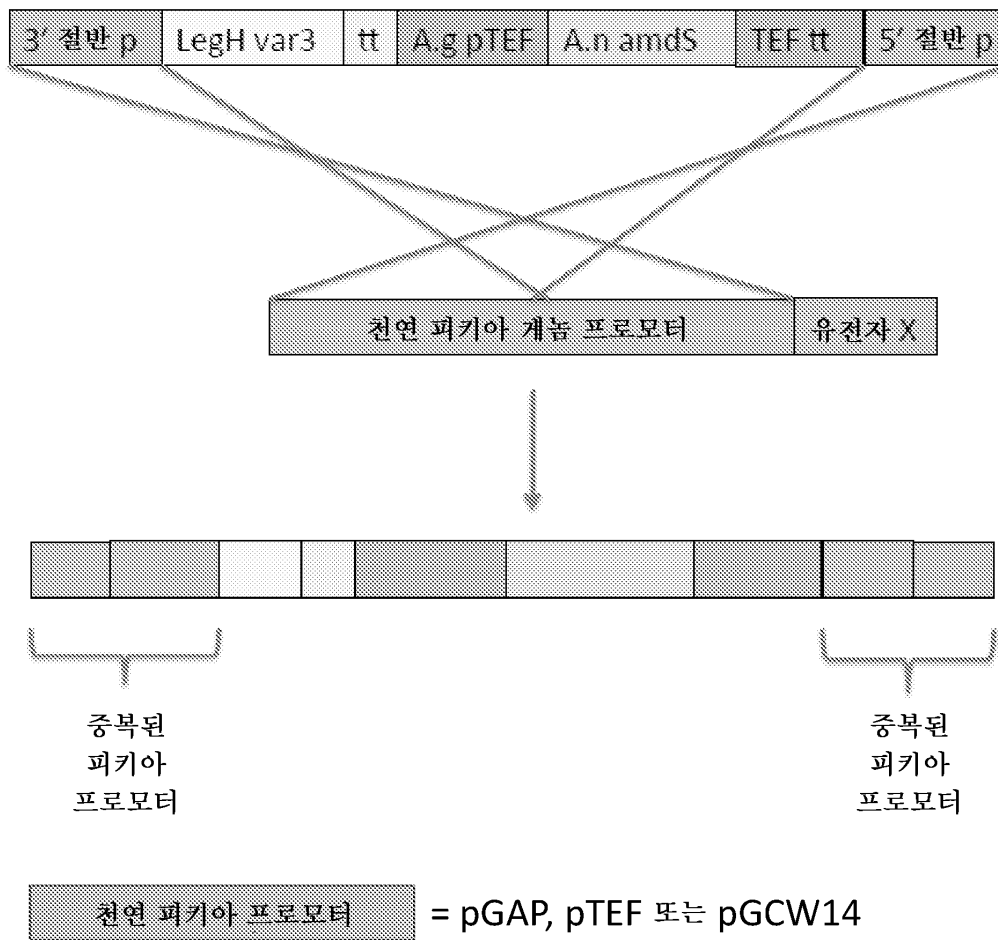
(i)



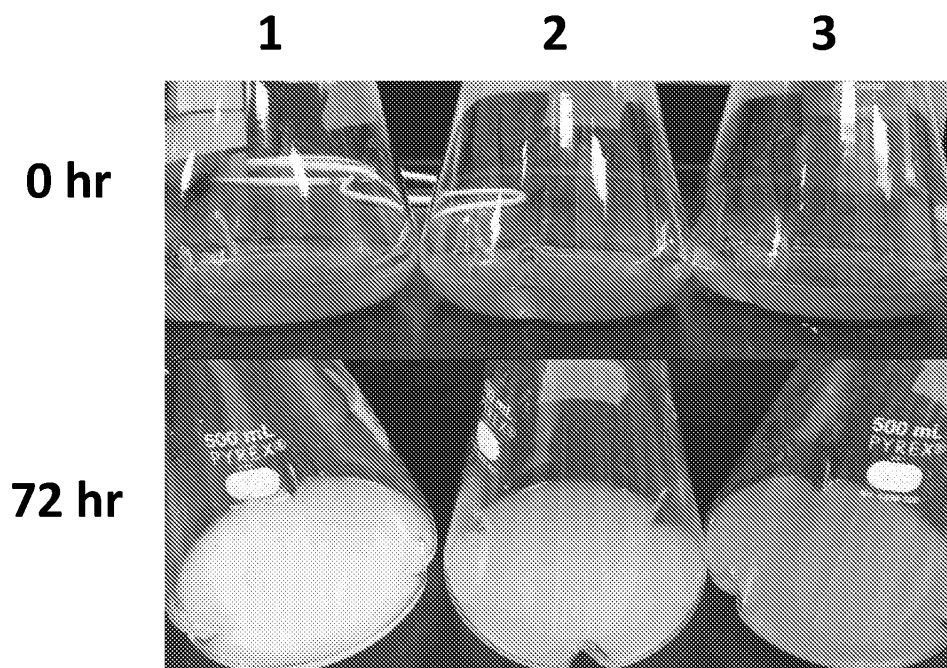
(ii)



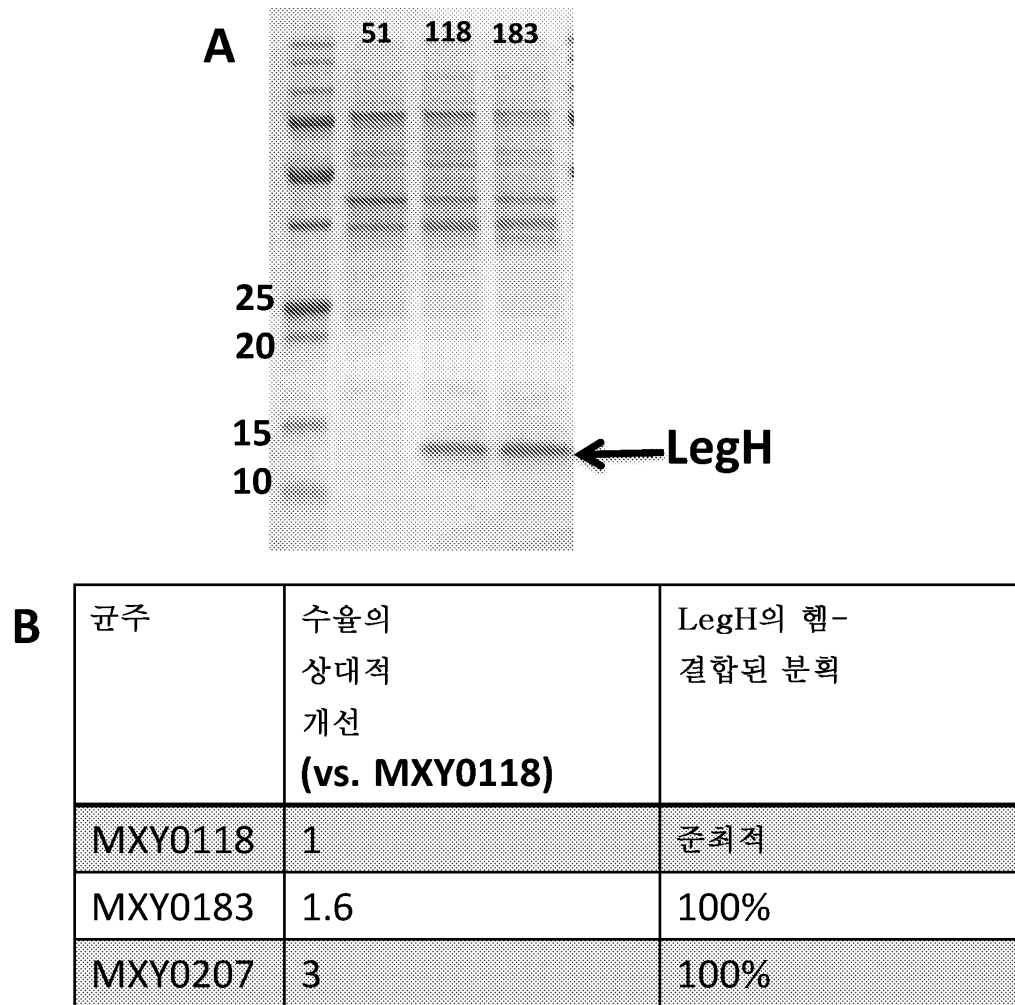
도면7



도면8

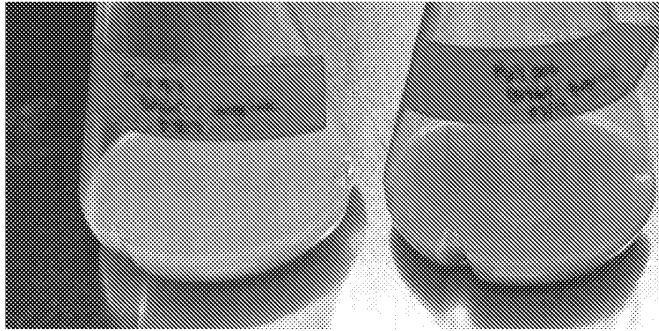


도면9

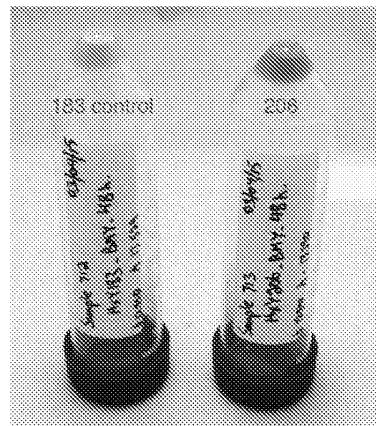


도면10

A



B



C

균주	혈 로딩된 LegH의 상대적 수율
183- BMGY	N/A
183- BMY	1
206- BMGY	3
206- BMY	5

도면11

MXV 균주	글리세롤 공동-공급 시의 수율	텍스트로스 공동-공급 시의 수율
183	1X	1X
207	4.4X	4.4X
291	5.8X	5.8X
306	5.8X	5.8X
330	6.9X	6.9X
333	5.8X	5.8X
338	6.9X	6.9X

서열 목록

SEQUENCE LISTING

<110> Impossible Foods Inc.

<120> EXPRESSION CONSTRUCTS AND METHODS OF GENETICALLY ENGINEERING
METHYLOTROPHIC YEAST

<130> 38767-0048W01

<140> PCT/US2016/031797

<141> 2016-05-11

<150> 62/313,491

<151> 2016-03-25

<150> 62/236,506

<151> 2015-10-02

<160> 58

<170> PatentIn version 3.5

<210> 1

<211> 3486

<212> DNA

<213> Komagataella phaffii

<400> 1

atgcgagacc gcggccgcat gagcaatcta cccccaactt ttggttcac tagacaatct

60

ccagaagacc aatcacctcc cgtgcccaag gagctgtcat tcaatgggac cacaccctca 120

ggaaagctac gcttatttgt ctgtcagaca tgtactcgag catttgctcg tcaggaacac 180

ttgaaacgac acgaaaggctc tcacaccaag gagaaacctt tcagctgcgg catttgttct 240

cgtaaattca gccgtcgaga tctgttattg agacatgccc aaaaactgca cagcaactgc 300

tctgatgcgg ccataacaag actaaggcgc aaggcaactc gtcggtcttc taatgccgcg 360

ggttccatat ctggtttctac tccggtgaca acgccaata ctatgggtac gcccgagat 420

ggcgagaaac gaaaagtcca gaaactggcc ggccgccggg actcaaatga acagaaactg 480

caactgcaac aacaacatct acagcaacaa ccacagtgc aataccaaca atctcttaag 540

cagcatgaaa atcaagtcca gcagcctgat caagatccat tgatatcccc gagaatgcaa 600

ttattcaatg attccaacca tcacgtaaac aatttgtttg atcttggact aagaagagct 660

tccttctccg ccgttagtgg aaataattat gccattatg tgaataattt tcaacaagat 720

gcctcttcta ccaatccaaa tcaagattca aataatgccg aatttgagaa tattgaattt 780

tctacccac aaatgatgcc cgttgaagat gctgaaactt ggatgaacaa catgggtcca 840

attccgaact tcctctcga tgtgaacagg aacattgggtg atagctttac agatatataa 900

cacaagaatt cagagcctat tataatccgaa ccgcccaagg acaccgtcc aaacgacaag 960

aagttgaatg gctactcttt ttacgaagcc cccatcaage cattagaatc cctattttct 1020

gtcaggaata caaagagaaa caagtataaa acaaatgacg actctccaga caccgtggat 1080

aataactccg caccggctgc taataccatt caagaacttg agtcttcttt gaatgcatcc 1140

aagaattttt gcttgccaac tggttattcc ttctatggta atttgacca acagactttc 1200

tctaacacgt tatcatgcac ttcttctaata gccacaattt cgccattct actcgataac 1260

tccattaata ataactccac tagtgacgtg agaccagaat ttagaacaca aagtgtcacc 1320

tctgaaatga gtcaagcccc tccccctcct caaaaaaaca actcgaaata ttccaccgaa 1380

gttcttttta ccagcaacat gcggtcgttt attcactacg ctctttccaa gtatcctttt 1440

attgggtgtc ccaactcaac tcttcgggag aacgaaagac taaatgaata tgctgattca 1500

ttaccaacc gtttcttaaa tcattatcct ttcatacatg tcacgattct caaagaatac 1560

tcccttttca aggcaatttt agatgagaat gagtcgacta agaactggga aaataatcag 1620

ttttacttag agaaccaacg aatatcaatt gtttgtcttc ctcttttggg ggctacgata 1680

ggtgcagtac tatcaacaa caaaaaggat gcttcgaatt tatacgaagc ttcaaggcgt 1740

tgtattcatg ttacttaga ttccaggaaa aagataccca cttccttgtc cgcaataaac 1800

aatgactctc cactttggct aattcaatcc ctgacgttat ctgttatgta tgggttattt 1860

gcggacaatg acattagttt gaatgtcgtg atcagacaag ttaacgcact taattctctg	1920
gtcaagactt cgggcctgaa taggacctca attatagatc ttttcaacat caacaaacct	1980
ttggataatg aactctggaa tcaattcgtg aaaatagagt ccaccgtaag gacaatccac	2040
acgatttttc aaatcagttc caacttaagc gccttgtaca atattattcc atcgttgaaa	2100
attgatgacc taatgattac tctaccagtt cccacaacac tttggcaagc tgattctttt	2160
gtgaaattca aaagtctaag ttacggaaat cagatccctt ttcaatatac aagagtacta	2220
cagaatttga ttgattacaa tcagccattg agcgaatggaa aatttttgta tgaaaacct	2280
gtaagtgagt ttggactcat atgcctacag aatggtctac accaatacag ctatttccaa	2340
aaattgactg ctgtcaataa cagagaagat gcgctattca caaaggttgt taattcactt	2400
cacagttggg ataggatgat ttcgaattct gatttgtttc caaagaagat atatcagcag	2460
agttgcttga ttttggactc aaagttgctt aataatttcc tgattgtcaa gagtcattg	2520
aaagtttcga cgggagacgt tagttctttg aataagttaa aagaaaacgt gtggcttaaa	2580
aactggaatc aagtgtgtgc tatctattat aacagcttca tgaacattcc tgctcccagt	2640
attcaaaaga agtacaatga catagagttt gtggatgaca tgattaattt gagtctaate	2700
atcatcaaga ttatgaaact cattttctat aacaatgtca aagacaatta tgaggatgaa	2760
aatgacttca aattgcaaga gttaaattta acatttgaca attttgatga gaaaatatcc	2820
ttgaatttga caatattatt cgatatattt ttgatgatct acaagataat taccaattac	2880
gaaaagttta tgaagatcaa acacaagttt aattactaca attctaattc gaataaagc	2940
ttcttgcatc atttcgaact ctctcgggtt atcaataaca cccaaatgaa ccagaatgat	3000
tatatgaaaa cagatatatga tgaaaagcct gatcagcttt tccacatcta tcaaacattt	3060
ttccggctgt atctggattt agaaaagttt atgaagtcca aattcaacta tcatgacttt	3120
gagacagagt tttcaagtct ctcaatatcc aatatactga aactcatgc tgcttctaac	3180
aatgacacaa atgctgctga tgctatgaat gccaaggatg aaaaaatgc tcccacaact	3240
ttgaatagcg tattacttgc tgatgaagga aatgaaaatt ccggtcgtaa taacgattca	3300
gaccgcctgt tcatgctgaa cgagctaatt aattttgaag taggtttgaa atttctcaag	3360
ataggtgagt catTTTTTga tttcttgtat gagaataact acaagttcat ccacttcaaa	3420
aacttaaatg acggaatgtt ccacatcagg atatacctag aaaaccgact agatggtggt	3480
gtctag	3486
<210> 2	
<211> 1161	
<212> PRT	

<213> Komagataella phaffii

<400> 2

Met Arg Asp Arg Gly Arg Met Ser Asn Leu Pro Pro Thr Phe Gly Ser

1 5 10 15

Thr Arg Gln Ser Pro Glu Asp Gln Ser Pro Pro Val Pro Lys Glu Leu

20 25 30

Ser Phe Asn Gly Thr Thr Pro Ser Gly Lys Leu Arg Leu Phe Val Cys

35 40 45

Gln Thr Cys Thr Arg Ala Phe Ala Arg Gln Glu His Leu Lys Arg His

50 55 60

Glu Arg Ser His Thr Lys Glu Lys Pro Phe Ser Cys Gly Ile Cys Ser

65 70 75 80

Arg Lys Phe Ser Arg Arg Asp Leu Leu Leu Arg His Ala Gln Lys Leu

85 90 95

His Ser Asn Cys Ser Asp Ala Ala Ile Thr Arg Leu Arg Arg Lys Ala

100 105 110

Thr Arg Arg Ser Ser Asn Ala Ala Gly Ser Ile Ser Gly Ser Thr Pro

115 120 125

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

130 135 140

Lys Val Gln Lys Leu Ala Gly Arg Arg Asp Ser Asn Glu Gln Lys Leu

145 150 155 160

Gln Leu Gln Gln Gln His Leu Gln Gln Gln Pro Gln Leu Gln Tyr Gln

165 170 175

Gln Ser Leu Lys Gln His Glu Asn Gln Val Gln Gln Pro Asp Gln Asp

180 185 190

Pro Leu Ile Ser Pro Arg Met Gln Leu Phe Asn Asp Ser Asn His His

195 200 205

Val Asn Asn Leu Phe Asp Leu Gly Leu Arg Arg Ala Ser Phe Ser Ala

210 215 220

Val Ser Gly Asn Asn Tyr Ala His Tyr Val Asn Asn Phe Gln Gln Asp

225 230 235 240
 Ala Ser Ser Thr Asn Pro Asn Gln Asp Ser Asn Asn Ala Glu Phe Glu
 245 250 255
 Asn Ile Glu Phe Ser Thr Pro Gln Met Met Pro Val Glu Asp Ala Glu
 260 265 270
 Thr Trp Met Asn Asn Met Gly Pro Ile Pro Asn Phe Ser Leu Asp Val
 275 280 285
 Asn Arg Asn Ile Gly Asp Ser Phe Thr Asp Ile Gln His Lys Asn Ser
 290 295 300

 Glu Pro Ile Ile Ser Glu Pro Pro Lys Asp Thr Ala Pro Asn Asp Lys
 305 310 315 320
 Lys Leu Asn Gly Tyr Ser Phe Tyr Glu Ala Pro Ile Lys Pro Leu Glu
 325 330 335
 Ser Leu Phe Ser Val Arg Asn Thr Lys Arg Asn Lys Tyr Lys Thr Asn
 340 345 350
 Asp Asp Ser Pro Asp Thr Val Asp Asn Asn Ser Ala Pro Ala Ala Asn
 355 360 365

 Thr Ile Gln Glu Leu Glu Ser Ser Leu Asn Ala Ser Lys Asn Phe Cys
 370 375 380
 Leu Pro Thr Gly Tyr Ser Phe Tyr Gly Asn Leu Asp Gln Gln Thr Phe
 385 390 395 400
 Ser Asn Thr Leu Ser Cys Thr Ser Ser Asn Ala Thr Ile Ser Pro Ile
 405 410 415
 Leu Leu Asp Asn Ser Ile Asn Asn Asn Ser Thr Ser Asp Val Arg Pro
 420 425 430

 Glu Phe Arg Thr Gln Ser Val Thr Ser Glu Met Ser Gln Ala Pro Pro
 435 440 445
 Pro Pro Gln Lys Asn Asn Ser Lys Tyr Ser Thr Glu Val Leu Phe Thr
 450 455 460
 Ser Asn Met Arg Ser Phe Ile His Tyr Ala Leu Ser Lys Tyr Pro Phe
 465 470 475 480

Ile Gly Val Pro Thr Pro Thr Leu Pro Glu Asn Glu Arg Leu Asn Glu
485 490 495

Tyr Ala Asp Ser Phe Thr Asn Arg Phe Leu Asn His Tyr Pro Phe Ile
500 505 510

His Val Thr Ile Leu Lys Glu Tyr Ser Leu Phe Lys Ala Ile Leu Asp
515 520 525

Glu Asn Glu Ser Thr Lys Asn Trp Glu Asn Asn Gln Phe Tyr Leu Glu
530 535 540

Asn Gln Arg Ile Ser Ile Val Cys Leu Pro Leu Leu Val Ala Thr Ile
545 550 555 560

Gly Ala Val Leu Ser Asn Asn Lys Lys Asp Ala Ser Asn Leu Tyr Glu
565 570 575

Ala Ser Arg Arg Cys Ile His Val Tyr Leu Asp Ser Arg Lys Lys Ile
580 585 590

Pro Thr Ser Leu Ser Ala Asn Asn Asn Asp Ser Pro Leu Trp Leu Ile
595 600 605

Gln Ser Leu Thr Leu Ser Val Met Tyr Gly Leu Phe Ala Asp Asn Asp
610 615 620

Ile Ser Leu Asn Val Val Ile Arg Gln Val Asn Ala Leu Asn Ser Leu
625 630 635 640

Val Lys Thr Ser Gly Leu Asn Arg Thr Ser Ile Ile Asp Leu Phe Asn
645 650 655

Ile Asn Lys Pro Leu Asp Asn Glu Leu Trp Asn Gln Phe Val Lys Ile
660 665 670

Glu Ser Thr Val Arg Thr Ile His Thr Ile Phe Gln Ile Ser Ser Asn
675 680 685

Leu Ser Ala Leu Tyr Asn Ile Ile Pro Ser Leu Lys Ile Asp Asp Leu
690 695 700

Met Ile Thr Leu Pro Val Pro Thr Thr Leu Trp Gln Ala Asp Ser Phe
705 710 715 720

Val Lys Phe Lys Ser Leu Ser Tyr Gly Asn Gln Ile Pro Phe Gln Tyr

725 730 735
 Thr Arg Val Leu Gln Asn Leu Ile Asp Tyr Asn Gln Pro Leu Ser Asp
 740 745 750

 Gly Lys Phe Leu Tyr Glu Asn His Val Ser Glu Phe Gly Leu Ile Cys
 755 760 765
 Leu Gln Asn Gly Leu His Gln Tyr Ser Tyr Phe Gln Lys Leu Thr Ala
 770 775 780
 Val Asn Asn Arg Glu Asp Ala Leu Phe Thr Lys Val Val Asn Ser Leu
 785 790 795 800
 His Ser Trp Asp Arg Met Ile Ser Asn Ser Asp Leu Phe Pro Lys Lys
 805 810 815

 Ile Tyr Gln Gln Ser Cys Leu Ile Leu Asp Ser Lys Leu Leu Asn Asn
 820 825 830
 Phe Leu Ile Val Lys Ser Ser Leu Lys Val Ser Thr Gly Asp Val Ser
 835 840 845
 Ser Leu Asn Lys Leu Lys Glu Asn Val Trp Leu Lys Asn Trp Asn Gln
 850 855 860
 Val Cys Ala Ile Tyr Tyr Asn Ser Phe Met Asn Ile Pro Ala Pro Ser
 865 870 875 880

 Ile Gln Lys Lys Tyr Asn Asp Ile Glu Phe Val Asp Asp Met Ile Asn
 885 890 895
 Leu Ser Leu Ile Ile Ile Lys Ile Met Lys Leu Ile Phe Tyr Asn Asn
 900 905 910
 Val Lys Asp Asn Tyr Glu Asp Glu Asn Asp Phe Lys Leu Gln Glu Leu
 915 920 925
 Asn Leu Thr Phe Asp Asn Phe Asp Glu Lys Ile Ser Leu Asn Leu Thr
 930 935 940

 Ile Leu Phe Asp Ile Phe Leu Met Ile Tyr Lys Ile Ile Thr Asn Tyr
 945 950 955 960
 Glu Lys Phe Met Lys Ile Lys His Lys Phe Asn Tyr Tyr Asn Ser Asn
 965 970 975

Ser Asn Ile Ser Phe Leu His His Phe Glu Leu Ser Ser Val Ile Asn
 980 985 990
 Asn Thr Gln Met Asn Gln Asn Asp Tyr Met Lys Thr Asp Ile Asp Glu
 995 1000 1005

Lys Leu Asp Gln Leu Phe His Ile Tyr Gln Thr Phe Phe Arg Leu
 1010 1015 1020
 Tyr Leu Asp Leu Glu Lys Phe Met Lys Phe Lys Phe Asn Tyr His
 1025 1030 1035
 Asp Phe Glu Thr Glu Phe Ser Ser Leu Ser Ile Ser Asn Ile Leu
 1040 1045 1050
 Asn Thr His Ala Ala Ser Asn Asn Asp Thr Asn Ala Ala Asp Ala
 1055 1060 1065

Met Asn Ala Lys Asp Glu Lys Ile Ser Pro Thr Thr Leu Asn Ser
 1070 1075 1080
 Val Leu Leu Ala Asp Glu Gly Asn Glu Asn Ser Gly Arg Asn Asn
 1085 1090 1095
 Asp Ser Asp Arg Leu Phe Met Leu Asn Glu Leu Ile Asn Phe Glu
 1100 1105 1110
 Val Gly Leu Lys Phe Leu Lys Ile Gly Glu Ser Phe Phe Asp Phe
 1115 1120 1125

Leu Tyr Glu Asn Asn Tyr Lys Phe Ile His Phe Lys Asn Leu Asn
 1130 1135 1140
 Asp Gly Met Phe His Ile Arg Ile Tyr Leu Glu Asn Arg Leu Asp
 1145 1150 1155

Gly Gly Val
 1160

<210> 3

<211> 438

<212> DNA

<213> Artificial Sequence

<220><223> Glycine max leghemoglobin codon-optimized for Pichia pastoris

<400> 3

atgggtgctt tcaccgagaa gcaggaagca cttgtttcct cttcgttcga agcttttaag 60

gctaacatcc ctcaatactc tgttgtgttt tacacgtcca ttctagaaaa agctcctgct 120

gccaaaggacc tcttctcttt tctgtccaac ggtgtagatc catccaatcc caaattaaca 180

ggtcacgctg agaaattgtt cggtttagtc agagatagcg ctggacaatt gaaagcaaat 240

ggtactgtgg ttgctgatgc tgccttgggc agcatccatg cacagaaggc aattacagac 300

ccacaatttg ttgttgtgaa ggaagctctg cttaaaacta taaaggaagc cgtcggagac 360

aaatggagtg acgagttgtc atcagcttgg gaggtagctt atgatgagtt ggccgcagca 420

atcaaaaagg cattctaa 438

<210> 4

<211> 145

<212> PRT

<213> Artificial Sequence

<220><223> Glycine max leghemoglobin codon-optimized for Pichia pastoris

<400> 4

Met Gly Ala Phe Thr Glu Lys Gln Glu Ala Leu Val Ser Ser Ser Phe

1 5 10 15

Glu Ala Phe Lys Ala Asn Ile Pro Gln Tyr Ser Val Val Phe Tyr Thr

20 25 30

Ser Ile Leu Glu Lys Ala Pro Ala Ala Lys Asp Leu Phe Ser Phe Leu

35 40 45

Ser Asn Gly Val Asp Pro Ser Asn Pro Lys Leu Thr Gly His Ala Glu

50 55 60

Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Gly Gln Leu Lys Ala Asn

65 70 75 80

Gly Thr Val Val Ala Asp Ala Ala Leu Gly Ser Ile His Ala Gln Lys

85 90 95

Ala Ile Thr Asp Pro Gln Phe Val Val Val Lys Glu Ala Leu Leu Lys

100 105 110

Thr Ile Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Ser Ser

115 120 125

Ala Trp Glu Val Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala

130 135 140

Phe

145

<210> 5

<211> 438

<212> DNA

<213> Artificial Sequence

<220><223> Glycine max leghemoglobin codon-optimized for Pichia pastoris

<400> 5

atgggtgcat ttacagaaaa acaagaggct ttagtatacct catcttttga agctttcaaa 60

gccaatattc ctcaatactc cgttgttttc tatacgtcca ttttggaaaa ggctccagca 120

gctaaggacc ttttctcttt cttgtcgaac ggcgtggatc cctcaaatcc taagctgact 180

ggtcacgccg agaagctttt tggtttggtc agagacagcg ccggacagct gaaagctaac 240

ggtacagttg tggcagatgc tgccttggga tctatacatg cacaaaaggc tatcaccgac 300

ccacagtttg tggttgtaaa agaggctcta ctcaaaacta tcaaggaagc agttggtgac 360

aaatggagcg atgaattgtc cagtgcattg gaggtcgctt acgatgagtt agctgctgca 420

atcaaaaagg ctttctaa 438

<210> 6

<211> 145

<212> PRT

<213> Artificial Sequence

<220><223> Glycine max leghemoglobin codon-optimized for Pichia pastoris

<400> 6

Met Gly Ala Phe Thr Glu Lys Gln Glu Ala Leu Val Ser Ser Ser Phe

1 5 10 15

Glu Ala Phe Lys Ala Asn Ile Pro Gln Tyr Ser Val Val Phe Tyr Thr

20 25 30

Ser Ile Leu Glu Lys Ala Pro Ala Ala Lys Asp Leu Phe Ser Phe Leu

35 40 45

Ser Asn Gly Val Asp Pro Ser Asn Pro Lys Leu Thr Gly His Ala Glu

50 55 60

Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Gly Gln Leu Lys Ala Asn

65 70 75 80
 Gly Thr Val Val Ala Asp Ala Ala Leu Gly Ser Ile His Ala Gln Lys
 85 90 95
 Ala Ile Thr Asp Pro Gln Phe Val Val Val Lys Glu Ala Leu Leu Lys
 100 105 110
 Thr Ile Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Ser Ser

 115 120 125
 Ala Trp Glu Val Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala
 130 135 140

Phe

145

<210> 7

<211> 939

<212> DNA

<213> *Pichia pastoris*

<400> 7

gatctaacaat ccaaagacga aaggttgaat gaaacctttt tgccatccga catccacagg	60
tccatttctca cacataagtg ccaaacgcaa caggagggga tacactagca gcagaccgtt	120
gcaaacgcag gacctccact cctcttctcc tcaacacca cttttgcat cgaaaaacca	180
gcccagttat tgggcttgat tggagctcgc tcattccaat tccttctatt aggctactaa	240
caccatgact ttattagcct gtctatcctg gccccctgg cgaggttcat gtttgtttat	300
ttccgaatgc aacaagctcc gcattacacc cgaacatcac tccagatgag ggctttctga	360
gtgtggggtc aaatagtttc atgttcccca aatggcccaa aactgacagt ttaaagctg	420
tcttgaacc taatagaca aaagcgtgat ctcatccaag atgaactaag tttggttcgt	480
tgaaatgcta acggccagtt ggtcaaaaag aaacttccaa aagtcggcat accgtttgtc	540
ttgtttggta ttgattgacg aatgctcaaa aataatctca ttaatgtta gcgcagtctc	600
tctatcgctt ctgaaccccg gtgcacctgt gccgaaacgc aaatggggaa acaccgctt	660
tttggatgat taigcattgt ctccacattg tatgcttcca agattctggt gggaaactg	720
ctgatagcct aacgttcatg atcaaaattt aactgttcta acccctactt gacagcaata	780
tataaacaga aggaagctgc cctgtcttaa accttttttt ttatcatcat tattagctta	840
ctttcataat tgcgactggg tccaattgac aagcttttga ttttaacgac ttttaacgac	900
aacttgagaa gatcaaaaaa caactaatta ttcgaaacg	939

<210> 8

<211> 558

<212> DNA

<213> *Pichia pastoris*

<400> 8

cgactattat cgatcaatga aatccatcaa gattgaaatc ttaaaattgc ccctttcact	60
tgacaggatc cttttttgta gaaatgtctt ggtgtcctcg tccaatcagg tagccatctc	120
tgaaatatct ggctccgttg caactccgaa cgacctgctg gcaacgtaaa attctccggg	180
gtaaaactta aatgtggagt aatggaacca gaaacgtctc ttcccttctc tctccttcca	240
ccgcccgta ccgtccctag gaaattttac tctgctggag agcttcttct acggccccct	300
tgcagcaatg ctcttccag cattacgttg cgggtaaaac ggaggtcgtg taccgacct	360
agcagcccag ggatggaaaa gtcccggccg tcgctggcaa taatagcggg cggacgcatg	420
tcatgagatt attggaacc accagaatcg aatataaaag gcgaacacct ttcccaattt	480
tggtttctcc tgacccaaag actttaaat taatttatit gtcctatit caatcaattg	540
aacaactatc aaaacacg	558

<210> 9

<211> 822

<212> DNA

<213> *Pichia pastoris*

<400> 9

caggtgaacc cacctaacta tttttaactg ggatccagt agctcgtgg gtgaaagcca	60
accatctttt gtttcgggga accgtgctcg ccccgtaaag ttaatttttt ttcccgcg	120
agctttaatc tticggcaga gaaggcgttt tcatcgtagc gtgggaacag aataatcagt	180
tcatgtgcta tacaggcaca tggcagcagt cactatitg ctttttaacc ttaaagtcgt	240
tcatcaatca ttaactgacc aatcagattt ttgcatitg ccacttatct aaaaatactt	300
ttgtatctcg cagatacgtt cagtggtttc caggacaaca ccaaaaaaa ggtatcaatg	360
ccactaggca gtcggtttta tttttgtca cccacgcaa gaagcaccca cctcttttag	420
gttttaagtt gtgggaacag taacaccgcc tagagcttca ggaaaaacca gtacctgtga	480
ccgcaattca ccatgatgca gaatgttaat ttaaagcagt gccaaatcaa gatttcaaca	540
gacaaatcaa tcgatccata gttacccatt ccagcctttt cgtcgtcgag cctgttcat	600
tcctgcctca ggtgcataac ttgcatgaa aagtccagat tagggcagat tttagttta	660

aaataggaaa tataaacaata tataaccgca aaaaggtttg tttatagctt ttcgcctggt 720
 gccgtacggt ataaatacat actctctctc cccccctggt tctctttttc tttgttact 780
 tacattttac cgttcctgca ctgccttcac tcaacaacaa aa 822
 <210> 10
 <211> 424
 <212> DNA
 <213> Pichia pastoris
 <400> 10
 ataactgtcg cctctttttat ctgccgcaact gcatgagggtg tccccttagt gggaaagagt 60
 actgagccaa ccctggagga cagcaaggga aaaataccta caacttgctt cataatggtc 120
 gtaaaaacaa tccttgtcgg atataagtgt ttagactgt cccttaccct ctgcgatgtt 180
 ctctctctca aagtttgcga tttctctctc tcagaattgc catcaagaga ctccaggacta 240
 atttcgcagt cccacacgca ctgctacatg attggctgaa atttcctaa agaatttctt 300
 tttcacgaaa atttttttt tacacaagat tttcagcaga tataaatgg agagcaggac 360
 ctccgctgtg actcttcttt tttttctttt attctcacta catacatttt agttattcgc 420
 caac 424
 <210> 11
 <211> 1683
 <212> DNA
 <213> Pichia pastoris
 <400> 11
 atggagtttg tgcctcgca gtccatgaat gcctgtccct ttgtcaggtc aacttctacc 60
 caccatttga agaagttggc agcaaacagt tctctagctg ctactgctag tcattgtccc 120
 gtggttggcc ctgctctcca acagcagaga tactactctc aaccttcaa gccagcccaa 180
 gcccaaacct ccgacattgc tactgggatc aagaaggatg tttctccgat ccgtatggac 240
 tctaataaaa ccgcctttga ttacaatgga atgtatgagt ctgatcttgc gaataaacgt 300
 aaagataact cgtatcgta tttcaataac atcaaccgtc tagccaagga gtttccaag 360
 gcacatgcc agaccgaaga tgacaagggtg accgtctggt gctctaacga ctacttagga 420
 atgggtaggc atcctgagat tatcaaac accatgaaggta ccatggacaa gtacgggttc 480
 ggagcaggag gaactaggaa cattgcaggt cataaccacg ccgctatcaa ttggaaagc 540
 gagttggctt gcttgaacaa gaaggaagcg gctctggtgt tttcatcatg tttcatagct 600
 aacgatgcaa tcactctggt gttgggacaa aaaatcaaaa atttggtcat tttctctgac 660

cagtcgaatc atgcttccat gatattgggt gtgcgtaact ccaaagcgaa gaagcacatc	720
ttcaagcaca acaatttgaa ggatctggag tcgcagttag ctcagtacc caagtcgact	780
cctaaactga tcgccttcga gtcagtttac tctatgtgtg gatctgtggc tcccattgag	840
aagatttgcg atttggctaa aaggtacggt gcctcacct tcttggatga agttcatgct	900
gttggaatgt atggctccta tggacagggt gtagctgagc atttggactt tgatctgcat	960
ttacagtctg gaatcgccag tcctagcgtg gtggacaaac gcaccatatt ggatcgtgtc	1020
gacatgatta ctggtacttg cggaaagtca ttggtactg ttggaggtta cgttgctggt	1080
agtgccaacc taattgattg gttaagatcc tatgcgccag gtttcatttt cactaccaca	1140
cttctcctg ctatcatggc tggtagagcc acttctgttc gtattgttag ggccgacatt	1200
gaggcccgta tcaagcaaca gcttaatact cgctacgtca aagactcatt tgaaaacctt	1260
ggtattccag tcattccaaa cccaagtcac attgttcttg ttctagttgg aaatgctgca	1320
gatgccaaga aggcatccga tatgttaatg aacaaacacc gtatttatgt tcaagctatt	1380
aactacccta ctgtgcctgt cgggtgaagaa cgactaagga ttactcctac tccaggtcat	1440
ggaaaggaga tttgtgacca gctgatcagc gctgtcgacg atgtttttac tgagcttaat	1500
ttaccaagaa tcaacaaatg gcagtcctaa ggtggtcatt gcggtgttgg tgatgctaat	1560
tacgtaccag aacccaatct gtggactcag gaccagctca gcttgacaaa ccaagacttg	1620
cactccaatg tgcacaaccc agtgattgag cagatcgaac cctcatcagg agtcagattg	1680
tag	1683
<210> 12	
<211> 1026	
<212> DNA	
<213> Pichia pastoris	
<400> 12	
atggtgcata aggtgaata cttggacgac cacccaactc agatttccag cattctttca	60
ggaggttaca accaccatt acttcgtgaa tggcaacatg aacgtcaact caacaaaaac	120
atgttcattt ttccctgtt tgtcacagat cgaccagacg aagaagaact tttcctagt	180
ctacctaata tcaagaggtt tggcgtaaac aagttgattc cttatgtagg aggtttggtt	240
tccaaaggat tgagggcggt gatctatttt ggtgttcttc tgaagcccg tgtgaaagat	300
gaagaaggaa cgccgcctga tgatccagag ggacctgtta tccaagccat caaacacttg	360
agaaagaact ttctgacct gtatatcatt accgatgtct gtctatgtga gtacaccagc	420
catggacatt gtggaatact atataggat ggcaactatc acagagagct ctgagtcggt	480

cgtattgctg ctgtagctgt caaatatgct caagctggag ccaactctgt ggctccttct 540
 gatatgactg acggcagaat aagagatatt aaagaaggct tactaagtgc aggactggca 600
 cataaaacgt ttgttatgtc ctacgctgca aaattctctg gtaatttcta tggccctttc 660
 agagatgctg caggttcctg tccatctcaa ggggacagaa aatgttacca gcttccttct 720

ggaggaaaag ggttggccca tcgtgctctg attcgtgata tgaatgaagg cactgatgga 780
 attattgtca aaccatctac attctatttg gacattgtcg ctgatgctta tcagctttgt 840
 aaagactatc ctatctgctg ttaccagggt tctggagagt acgccatgct acatgcagcg 900
 gcagagaaga atattgttga tctgaaatca atcgtctttg aagctcatca aggattcttg 960
 cgggctggag ctcgtttaat cattagttac tttaccctg aattcctgga gtggttatct 1020
 gaatga 1026

<210> 13

<211> 1035

<212> DNA

<213> Pichia pastoris

<400> 13

atgaacaaaa tcgaacagag cggacccatt gattgcagtt ccttgaaatt ggggtcccga 60
 aagtcgcgtc tggtataaat ccaggcagaa atcgtccgcc aattgatatt gaaagaatac 120
 cctgaattgg agacgaagtt ggtcagtggt tccaccctgg gggaccaagt ccagaataaa 180
 gcacttttca cgtttggagg aaaatctttg tggaccaaag aacttgagat gttgttgttg 240
 gagagtgtgg gaggatttga ccaaatagac atgattgtac actcgttgaa agacatgcca 300
 actcatttac cagacgaatt tgagctgggt tgcattattg aaagagaaga ccctagagac 360
 gctttggtcg tgcaagatgg tttatcttac aagtcattgg ccgaccttcc agaggagct 420

gtggtcggta cgtcttcggt tagaagatcg gctcaactac tgaagaattt ccctcatctg 480
 aaattcaaat ctgttagagg aaaccttcag accagactaa gaaaattaga tgatccagat 540
 tccgagtact gctgtctcct ccttgcagca gccggtttta tcaggacagg cttacaacac 600
 agaatttcaa tgtatttgaa cgacgatgtg atgtaccact ccgtcggaca aggagcatta 660
 ggagtagaga tcagaaaagg tgaccaattc atgaaaaata tctgtgaaaa gattgggcat 720
 agaaccacca cccttcgttg tcttgcagag agagcactgc tgagatatct agaggaggc 780
 tgctcgtgac caattggggt ctccactatt tatagcgagg atacgaagga acttaccatg 840

aactccctag tcgtcagttg taacgggtcgt gactcggtaa cagaatcaat gactgaagtc 900
 gtgactactg aagagcaagc tgaagatttc ggtgaaaggc tggcccagaa gctcatagat 960

caaggtgcga aacgcattct tgacgagatc aacttcaaca agatcaaaga gattaaggaa	1020
gagggtttac attaa	1035
<210> 14	
<211> 798	
<212> DNA	
<213> Pichia pastoris	
<400> 14	
atgccaaaag ccattcttct gaagaataaa actacaccga aggatcctta tctggagaac	60
tctgtaagta gtggctactc gaccgatttc gtaccacttt tagatcatat tcacatggag	120
aaatctgaga tcatcgcat tctcaagact gactactttt tgcataaaac tttggcgttt	180
attattacgt cccaaagagc tgtagaaatg ctgaatgagt gtatgcaaact actgagacgt	240
actgatcctg aaattacaca aatcatctat agtaaacctg tctatacagt tggccctgcc	300
acctacagaa tacttgcgga tgctggcttc gtggatctac gaggcggaga taaggcagga	360
aacggatcca ttctagccca gataattttg aatgatgaca ttactactgg aattgaagat	420
tctgacaagc atataacgtt ttctacggga gaaacaagga gagacataat tcccaaatgt	480
ttactctcta acaactttca actttacgaa aagattgtct acaagactct tcctagggat	540
gatatcgtga ctagattcaa gtctgccgtt gacagcatcg accaatcgca aagaagtcc	600
agttgggtgg tcttcttttc gcctcaagga acagaggaca ttgtaacgta tcttcaacac	660
accaaagacc aatttaatat tgcattctatc gggccaacca cagaaaaata ctttctaagc	720
aaaaacctga aacaaaagt tgtggcacct aagccagagc ctatctcttt actattgtct	780
atacaaaaag tgcactaa	798
<210> 15	
<211> 1086	
<212> DNA	
<213> Pichia pastoris	
<400> 15	
atgagtagat ttccagaact gaagaatgac cttattttta gggcagctcg tggtgaaaaa	60
gttgaacgtc ccccaatatg gattatgaga caggccggaa gatatcttcc ggagtaccat	120
gaggtcaaag gaggtaggga cttctttgaa acttgcaggg atgctgagat tgcttctgaa	180
attactatcc agccgattac gcattttgac ggtctgatcg atgcagctat tatcttcagt	240
gatatcttgg tgattcctca agctatgggc atggaagtta agatgggtgga caaagtggc	300
ccacagttcc ccaatccgct aagaaaaccg tctgacttgg atcatttgaa aaaagacgtt	360

gacgttttga aggaactcga ttgggccttc aaagctatct cattgaccag aaaaaaactc 420
aatggacgag tgccttttgc ttgatattgt ggtgctcctt ggactctact ggtttatatg 480

actgaaggag gcggtaccaa gatgtttcga tttgcaaaag agtggatcta caagtttacc 540
aaggaatctc atcaattact ccaacagatc actgacgttg cagtgaatt cttagctcag 600
caagttgttg caggtgcccc aatgttacia gtttttgaat cttggggcgg tgaattgggg 660
cctgatgaat tcgatgagtt ttctttgcct tatttgagac agatttcctc taaacttccc 720
ctgaggttga aggaacttgg aatcacagag aatgttccca taactgtctt tgctaaaggc 780
tcttggtacg ctttgagca attgtgcgac agtggttatg atgtgtctc gttggatttg 840
ttattccgct caagtgatgc tgtccagatt gctaacggaa gaatcgatt gcaaggtaat 900

cttgaccctg gaaccatgta cggctccaaa gaaaccattt ccaagaaagt ggacaaaatg 960
atcaagggtt ttggtggagg aaagcaaac tacataatta attttgaca cggcactcat 1020
ccattcatgg atccagaaca gatcagatgg ttcttacaag aatgtcatcg cattggatct 1080
caatag 1086

<210> 16
<211> 978
<212> DNA
<213> Pichia pastoris
<400> 16

atggccatcg actctgatat caatctaagc tctcccaatg attccatccg tcaaaggatg 60
ttcgagctta tccagcggaa gcaactcgaa attgtcgtg cattggaggc aattgaagga 120

aacgatacca aatttcgttc tgattcttgg gaaagaggag ccgaaggtgg aggaggaaga 180
tctatgctta ttcaagatgg aagagtgttt gaaaaggctg gtgtaaatat ttccaaggtt 240
catggcgtat tgctcctca agctgtgagc cagatgagaa atgaccactc caagctagat 300
ctgcctgcgg gaacctctct gaagtctttt gcctgtgggc ttctgttggc cattcatccc 360
cataatcccc atgctccaac taccatctg aattatcgct acttcgaaac ttgggatgaa 420
actggaaagc ctcacacctg gtggtttggg ggcggtgctg atttaacgcc ttcgtacctg 480
tatcccagg atgccaagca attccatcaa gcccataagg atgcctgga caaacacgat 540

gttagcttgt acccgagatt caaaaagtgg tgtgatgaat actttctgat caaacatcga 600
aatgaaacta gaggtattgg gggatatttc ttgatgatt ttgacgagtt tgatgctgag 660
aggtccctga agttggttga agattgtttc aatgctttct tggaatctta tcccgtatc 720
actcgaaaaa ggatggacac ccttcaact gatgctgaga agaactggca acaaattaga 780

agaggaagat atgtcgaatt caacttagta ttggatagag gtactcaatt tggtttgaga 840
acgcctggat ctctgttga aagtattttg atgtcgttgc caagaacagc tggttgggtc 900
tatgatcatc atccagagcc tggctccaga gaagaggagt tattgcaggt actacaaaat 960

cctattgaat gggatatga 978

<210> 17

<211> 1683

<212> DNA

<213> Pichia pastoris

<400> 17

atgctgaaaa gtcttgcacc aaattcctca attgccgttt taggttcagg gatatctgga 60
ttgactttca gctttttttt gaatcggttg cgtcccgatg ttaagatcca tatctttgaa 120
aaatccaagc aggttggagg atggatcaga tcagaagagc atgaaacttt tcattttgaa 180
aagggaccca gaactttgag aggcacaaaat acgggtacct tgatgttgtt ggatcttctt 240
accaagatag gagcaaatga caaggtcctg ggactgcaca aagattctct tgctaataaa 300

aagtatctgt tgtccccgtt ctcatagttt cacggaaaca acgcaaagct tcttcaagtg 360
ccacaggatt tcagctcttt tgtaaagttc atgtttgacc cgttgtctaa ggatctcatt 420
ctcgggtcttt tgaaagaacc atggcaacca aaattaaagt attcagatga gtcggttgac 480
cattttttca acagaagatt tgctacaaa ctatcagaga atatcgtcag cgcaattgtg 540
catggaatct atgcgggcga cgtgaagaag ttaagtgtga aagccatctt ccctaggctc 600
cctgagatgg aacaggaaaag tggctctatt ataaggtata tgatcgccca atacaggaca 660
aaaaagaacg tcaaacaaaa agttgaccct tttttggcag attatgaaaa attgatcggg 720

acatctttga gtttcaaaaa tatttctttg tttctgaaaa actttcccat gctgagtttt 780
caggttggac tacagaaact tcccatctca ttgaagaacc atttatcaca gattgaaaac 840
atcaagtttc attttgacag caaaatcaaa aacattgctt tggagagcgg taaggtggca 900
ttgactgacc atgatcaggt ttatcttgtt gaccatgtga gatctacat taataccaac 960
gaattggcca aaatcatttc acccgttgtt ccaagttcta ctaagaaaa atccgttttc 1020
aaatccaag cgaatggccc agggctgggtc aaatgtttga gctggctaca ctatacaaat 1080
atactaattg gcaacattta tatacctaag cacgtctcaa aatctatcac cggatttggg 1140

tacttgggtc ctcatcaat gtcttctcag gcatccaaac ttctcgggtg catatttgac 1200
tcagacatcg agactgcaat gactcctaatt ttacagagg ccaacattac ggcgataaac 1260
agtaactctg catctcccaa gcaactccaa aagttttctg accaattcgt caataatgat 1320

ctccctaaat acaccaagtt gacgctaattg cttggaggtc attatctcaa gtcggaggca	1380
gacatgcccc gtccgcaga gagtaacat gctgtcaagg cgattctgtc aaatcacctg	1440
aatattgata tagatgagtt tgcattcttg ccagacttca agatggaaat caccaagatc	1500
cccaactgca ttccccaata tgaagtggg tatcttgatc tcaagagaaa gggtcagaat	1560
gcagcctcca aagagttcaa cgaccaaata agttttggag gcatggcatt tggatgaggt	1620
gtggggatcc ctgactgtgt ccagaatgca ttcaaagatt cggtaccct cagtggcatt	1680
taa	1683
<210> 18	
<211> 1125	
<212> DNA	
<213> Pichia pastoris	
<400> 18	
atgcttaacc gtcgtttcca atctaccgtg tcctcgagtc tgaacaagg cactggaata	60
gtgttcata atagggtgg tccctccact gtcaaggaaa cctatgactt tttatttcgt	120
ctttctcgg acggagattt aatcccggtt ggcagatttc agaacatcct ggcccccttc	180
attgcaagta gaagaacacc caaaattgaa tcctactaca aagctatcgg aggtgggtct	240
cctatccgaa agtggctga ataccagat tctaaactat gtgaaaaatt agacattatc	300
agtccacaat cggctcctca taagccttat gttgccttca gatacgttaa tcctctcact	360
gaagatactt taaaaagat gaaaaatgat ggaattacta aggccattgc cttttctcaa	420
tatccgcaat ttagttattc aaccaccgga tcatcgatta acgaacttta caggcaatcg	480
aaaattttgg accctgatca atctattaaa tggacagtta tagatcgctg gcctgaccac	540
ccagccttag ttaaaacttt cgcagctcat atcaaagata ctctaaacag attcaaaact	600
gaaaatggac tgactgacac aaaagacgtc gtctccaat tcagtgtca ttcctttacca	660
atggatattg tcaataaagg agattcgtat cctgcagaag tcgcagcgag tgtctttgcc	720
attatgaaag aacttaactt ctcaaatect tataaattaa cctggcaatc acaggttggc	780
ccaaagcctt ggctgggtgc tcaactgaa aaaattacca agcagctagc atccagtgat	840
gttctcggag tcgttttggg tcctattgcc ttacctctg atcatattga aactctccat	900
gaactggata ttgaactgat tcaagaacta cctaatectt caaaagtaaa gcgagttgaa	960
tcgttgaacg gagaccaaac ttctattgac tccttggcag aactagttaa ggtcacatt	1020
gattcgaagg ttgtattttc caaccagttg ccattggatt ccatgctggg agtagtgtca	1080
gataattccc tcacagatcc aaaagagttt tcagagccc attga	1125

<210> 19
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 19
 gaggggtctcg gatggagttt gtcgcccgtc 30
 <210> 20
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 20
 gaggggtctcg attacaatct gactcctgat gagg 34
 <210> 21
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 21
 gaggggtctcg gatgggtgcat aaggctgaat acttg 35
 <210> 22
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 22
 gaggggtctcg attattcaga taaccactcc agg 33
 <210> 23
 <211> 35
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 23

gagggtctcg gatgccaaaa gccattcttc tgaag 35

<210> 24

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 24

gagggtctcg attagtgcac tttttgtata gac 33

<210> 25

<211> 35

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 25

gagggtctcg gatgagtaga tttccagaac tgaag 35

<210> 26

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 26

gagggtctcg attattgaga tccaatgcg 29

<210> 27

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 27

gagggtctcg gatggccatc gactctgata tc 32

<210> 28

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide
 <400> 28
 gaggggtctcg attataccca ttcaatagga t 31
 <210> 29
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 29
 gaggggtctcg gatgctgaaa agtcttgac caaa 34
 <210> 30
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 30
 gaggggtctcg attaaatgcc actgagggtg gc 32
 <210> 31
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 31
 caatcgctag catccaacat ccaaagacga aagg 34
 <210> 32
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 32
 ggatagcatg caccttatca agatagctag aaatagaaat gg 42
 <210> 33
 <211> 37

<212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 33
 caatagcatg caacatccaa agacgaaagg ttgaatg 37
 <210> 34
 <211> 43
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 34
 catggtaccg gtaccttata aagatagcta gaaatagaaa tgg 43
 <210> 35
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 35
 caatcgctag catccaacat ccaaagacga aagg 34
 <210> 36
 <211> 43
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 36
 gatattgctc gagaccttat caagatagct agaaatagaa atg 43
 <210> 37
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 37
 caatctcgag aacatccaaa gacgaaaggt tg 32
 <210> 38

<211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 38
 caaccatttc tatttctagc tatcttgata aggtcttaag tcca 44

<210> 39
 <211> 43
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 39
 catggtaccg gtaccttadc aagatagcta gaaatagaaa tgg 43

<210> 40
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 40
 ttacttaagt ccaacatcca aagacgaaag gttg 34

<210> 41
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 41
 tcacagacgc gttgaattgt cc 22

<210> 42
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220><223> Synthetic oligonucleotide
 <400> 42

ttgctcctca gcttagaaga actcgtccaa catcaagtg 39

<210> 43

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 43

ttagtcttgc tcctcagctt agcc 24

<210> 44

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 44

tcacagacgc gttgaattgt cc 22

<210> 45

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 45

ttttgcggcc gcatgagcaa tctaccccca acttttg 37

<210> 46

<211> 36

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 46

aaaagcggcc gcctagacac caccatctag tcggtt 36

<210> 47

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 47

actagatggt ggtgtctagt caagaggatg tcagaatgcc atttg 45

<210> 48

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 48

tctgacatcc tcttgactag acaccacat ctagtcggtt ttctag 46

<210> 49

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 49

aaacgtgtc ttggaaccta atatgac 27

<210> 50

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 50

aaactgtcag ttttgggcca ttg 24

<210> 51

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 51

acgtgtctt ggaaccta atgac 25

<210> 52

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 52

taccattca ataggatttt gtagtacctg c

31

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 53

gagcttcttc tacggccccc

20

<210> 54

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 54

tccagcagag taaaatttcc tagggac

27

<210> 55

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 55

ctcttttagg ttttaagttg tgggaacagt aaca

34

<210> 56

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 56

gtgggtgctt cttgcgtgg

20

<210> 57

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 57

agaattgccca tcaagagact caggact

27

<210> 58

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Synthetic oligonucleotide

<400> 58

gatagagaga aatcgcaaac ttgagagga ag

32