(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0099817
(43) 공개일자 2012년09월12일

(51) 국제특허분류(Int. Cl.)
C12N 1/21 (2006.01) *C12N 15/52* (2006.01)
C12N 15/62 (2006.01) *C12N 15/70* (2006.01)
(21) 출원번호 10-2011-0010772
(22) 출원일자 2011년02월07일
심사청구일자 2011년02월16일

(71) 출원인
경상대학교산학협력단
경상남도 진주시 진주대로 501 (가좌동)
(72) 발명자
김선원
경상남도 사천시 사남면 조동길 61, 대우푸르지오 아파트 102동 102호
왕충룡
경상남도 진주시 가좌동 900번지
(뒷면에 계속)
(74) 대리인
리앤목특허법인

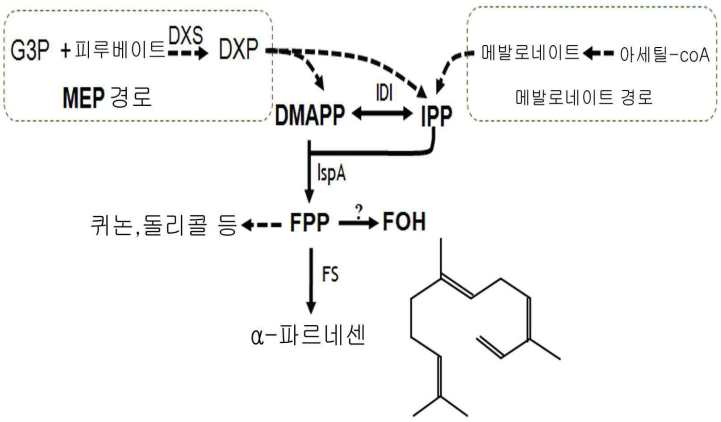
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 **α-파르네센 생산성이 향상된 에세리키아 속 미생물 및 그를 이용하여 α-파르네센을 생산하는 방법**

(57) 요 약

본 발명은 에세리키아속 미생물에 코돈-최적화된 α-파르네센 신타제로 형질전환된 에세리키아속 미생물 및 그를 이용한 α-파르네센 생산 방법에 관한 것이다. 본 발명의 미생물을 이용하면 α-파르네센을 높은 효율로 생산할 수 있다.

대 표 도 - 도1a



(72) 발명자

윤상활

충청남도 부여군 부여읍 서동로 71, 신안빌라 B동 202호

장희정

경상남도 진주시 진주대로829번길 11-13, 305호 (주약동)

최의성

대전광역시 유성구 노은로 353, 송림마을 307동 1703호 (하기동)

특허청구의 범위

청구항 1

에세리키아속 미생물에 코돈-최적화(codon-optimized)된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드로 형질전환된 에세리키아속 미생물.

청구항 2

제1항에 있어서, 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드 및 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드를 포함하는 융합 폴리뉴클레오티드로 형질전환된 에세리키아속 미생물.

청구항 3

제2항에 있어서, 상기 융합 폴리뉴클레오티드는 상기 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드와 상기 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드 사이에 하나 이상의 임의의 아미노산을 코딩하는 폴리뉴클레오티드 링커(linker)가 더 결합된 것인 에세리키아속 미생물.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 에세리키아속 미생물은 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트란스퍼라제 및 히드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 폴리뉴클레오티드, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 유래의 포스포메발로네이트 키나제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 폴리뉴클레오티드 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 폴리뉴클레오티드로 더 형질전환된 것인 미생물.

청구항 5

제1항에 있어서, 상기 에세리키아속 미생물에 코돈-최적화된 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 1의 염기 서열을 갖는 것인 미생물.

청구항 6

제2항에 있어서, 상기 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 2의 아미노산 서열, 및 상기 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 1의 염기 서열을 갖는 것인 미생물.

청구항 7

제3항에 있어서, 상기 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 2의 아미노산 서열, 상기 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 1의 염기 서열, 및 상기 하나 이상의 임의의 아미노산을 코딩하는 폴리뉴클레오티드 링커는 서열번호 3의 아미노산 서열을 갖는 것인 미생물.

청구항 8

제4항에 있어서, 상기 엔테로코커스 패칼리스 유래의 아세틸-CoA 아세틸트란스퍼라제 및 히드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 폴리뉴클레오티드는 서열번호 4, 상기 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 5, 상기 스트렙토코커스 뉴모니애 유래의 메발로네이트 키나제를 코딩하는 폴리뉴클레오티드는 서열번호 6, 상기 스트렙토코커스 뉴모니애 유래의 포스포메발로네이트 키나제를 코딩하는 폴리뉴클레오티드는 서열번호 7, 상기 스트렙토코커스 뉴모니애 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 폴리뉴클레오티드는 서열번호 8, 및 상기 대장균 유래의 이소펜테닐 디포스페이트

(IPP) 이소머라제를 코딩하는 폴리뉴클레오티드는 서열번호 9의 아미노산 서열을 갖는 것인 미생물.

청구항 9

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 미생물은 대장균인 것인 미생물.

청구항 10

제1항 내지 제3항 중 어느 한 항에 따른 미생물을 배양하는 단계; 및

배양물로부터 α-파르네센을 분리하는 단계를 포함하는 미생물로부터 α-파르네센을 생산하는 방법.

청구항 11

대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드 및 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드를 포함하는 융합 폴리뉴클레오티드로서, 서열번호 36의 염기 서열을 갖는 것인 융합 폴리뉴클레오티드.

명 세 서

기 술 분 야

[0001] 본 발명은 α-파르네센 생산성이 향상된 에세리키아 속 미생물 및 이를 이용하여 α-파르네센을 생산하는 방법에 관한 것이다.

배 경 기 술

[0002] 이소프레노이드는 살아있는 유기체에서 중요한 역할을 하는 물질로서 세포의 유동성, 전자 이동 및 기타 대사 작용을 유지하는데 사용된다. 이소프레이드는 40000개 이상의 생산물들로 이루어진 다양한 군을 포함하는데 다수의 천연 이소프레노이드 및 합성 이소프레노이드는 약품, 화장품, 향수, 색소 및 발색제, 살진균제, 방부제, 기능성 식품 및 정밀 화학 중간 물질로서 사용된다.

[0003] 천연의 상태에서, 이소프레노이드는 이의 전구체인 이소펜테닐 디포스페이트(IPP)와 이성체인 디메틸알릴 피로포스페이트(DMAPP)의 연속적인 축합 반응에 의해 합성된다. 상기 전구체에 대한 경로로서는 2가지가 알려져 있다. 식물을 제외한 진핵 생물은 일반적으로, 아세틸 조효소 A(아세틸-CoA)를 IPP로 전환시키기 위해 메발로네이트-의존 경로(MVA)를 이용하는데, 이 때, 상기 IPP는 추후 DMAPP로 이성체와 된다. 일부 예외도 있지만, 원핵생물은 통상적으로 IPP 및 DMAPP를 생산하기 위해 메발로네이트-독립 경로 또는 데옥시자일룰로스-5-포스페이트 경로(MEP)만을 이용한다. 식물은 MVA 경로와 MEP 경로 둘 다를 이용한다.

[0004] 통상적으로, 이소프레노이드는 천연 공급원 예를 들어 식물, 미생물 및 동물로부터 추출하여 제조되어 왔다. 그러나, 추출에는 다수의 심각한 한계점이 있기 때문에, 그로 인한 수율은 보통 매우 낮다. 첫째, 대부분의 이소프레노이드는 천연의 환경 하에서는 소량만이 축적된다. 둘째, 공급 유기체는 일반적으로 원하는 이소프레노이드를 상업적으로 유용한 양만큼 생산하는데 필요한 대규모 배양 공정에 적합하지 않다. 셋째, 이소프레노이드 추출을 위해서는 임의의 독성 용매가 필요하고, 이 용매의 취급과 처리 과정에 특별히 주의해야 하므로, 이소프레노이드를 상업적으로 생산해 내는데 어려움이 따른다.

[0005] 특히 식물로부터 추출되는 이소프레노이드의 클래스인 세스퀴테르펜은 중요한 의학적, 산업적 특성을 갖는다(Berger, 2009; Dhingra et al., 2000; Muntendam et al., 2009). α-파르네센(farnesene) (3,7,11-트리메틸도데카-1,3E,6E,10-테트라엔)은 사과 껍질에서 처음 발견된 가장 간단한 아시클릭 세스퀴테르펜의 하나로서, 식물 방어에 있어 역할을 하고 있다 (Huelin and Murray, 1966; Pare and Tumlinson, 1999). 파르네센은 파르네산(farnesane)으로의 수소화 반응으로 인해 최근 생물연료 전구체로서 개발되고 있으나 (Renneger and McPhee, 2008), 파르네센을 포함한 세스퀴테르펜은 천연적으로 아주 작은 양만큼만 생산된다. 따라서, 대사 공학(metabolic engineering)은 대장균 및 효모로부터 이러한 희박하고 가치있는 화합물을 생산을 위한 대안적이고 매력적인 경로가 될 수 있을 것이다.

[0006] 파르네센 형성은 파네실 피로포스페이트 (FPP)의 디피로포스포릴레이션을 촉매하는 파르네센 신타제(FS)에 의해 최종적으로 수행된다. 그러나 사과, 말루스 엑스 도메스티카(Malus x domestica) 로부터 분리된 것을 포함하는 단지 소수의 파르네센 신타제만이 현재까지 규명되었다. 또한 대장균으로부터의 파르네센 생산은 대부분의 경우

에 있어 대장균에서의 파르네센 신타제의 낮은 단백질 발현으로 인해 제한된다.

[0007] MVA 및 MEP 대사 경로가 밝혀짐으로써 이소프레노이드를 생합성 방법에 의해 생산할 수 있게 되었다. 그러나, 다수의 상업적 분야에서 요망되는 다량의 이소프레노이드 생성물이 생산되었음에도 불구하고, 현재의 기술로써 생산할 수 있는 이소프레노이드보다 더욱 많은 양의 이소프레노이드를 생산하는 발현 시스템과 발효법의 필요성이 대두되고 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 α-파르네센을 높은 효율로 생산할 수 있는 에세리키아 속 미생물을 제공하는 것이다.

[0009] 본 발명의 목적은 α-파르네센을 높은 효율로 생산하는 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 본 발명의 일 양태는 에세리키아속 미생물에 코돈-최적화(codon-optimized)된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드로 형질전환된 에세리키아속 미생물을 제공한다.

[0011] 상기 "코돈-최적화(codon-optimized)"는 표적 핵산 서열로 형질전환시키는 숙주 세포 내에서 표적 핵산 서열을 더욱 높은 수준으로 발현시키기 위해 표적 핵산 서열을 선택된 숙주 세포의 코돈 선호도를 반영하도록 변형시키는 것을 의미한다. 예를 들어 본 발명의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드의 염기 서열은 에세리키아속 미생물의 코돈 선호도에 맞게 변형시킬 수 있다. "코돈 선호도"란 유기체에 따라 각 아미노산을 코딩하는데 자주 사용하는 핵산 서열을 의미한다. 다수의 유기체에 대한 코돈 선호도 표를 입수할 수 있는데, 이 표는 본 발명의 서열을 디자인하는데 참고 기준으로서 사용할 수 있다(Gouy and Gautier (1982) Nucleic Acids Res 10(22) 7055-7074; Eyre-Walker (1996) Mol. Biol Evol 13(6) 864-872 참조). 소정의 숙주 미생물의 코돈 중 가장 많이 사용되는 코돈을 사용하면, 일반적으로 번역 가능성이 증가하게 되므로 원하는 서열의 발현 수준도 증가할 수 있다.

[0012] 상기 "에세리키아속 미생물"은 바람직하게는 에세리키아 콜라이 (Eschrichia coli), 더 구체적으로는 E. coli DH5α가 될 수 있으나 이에 제한되지 않는다.

[0013] 본 발명의 일 구체예에서, 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드 및 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드를 포함하는 융합 폴리뉴클레오티드로 형질전환된 에세리키아속 미생물을 제공한다.

[0014] 상기 "대장균 유래의 FPP 신타제"는 에세리키아 콜리 유래의 파네실 피로포스페이트((E,E)-farnesyl pyrophosphate) 합성효소를 의미하고 이를 코딩하는 폴리뉴클레오티드는 IspA 유전자의 프로모터와 전사 종결 부분을 제외한 실질적으로 FPP 신타제 단백질을 코딩하는 부분을 의미한다.

[0015] 상기 "융합 폴리뉴클레오티드"는 2 이상의 단백질을 코딩하는 폴리뉴클레오티드 핵산 서열을 당업자에게 잘 알려진 분자생물학적 방법으로 연결시킨 폴리뉴클레오티드 핵산 서열을 의미한다. 융합 폴리뉴클레오티드는 근접한 시간대에 근접한 장소에서 융합된 폴리뉴클레오티드들이 코딩하는 단백질들이 같이 발현될 수 있게 한다. 따라서 융합된 폴리뉴클레오티드가 코딩하는 단백질들이 서로 상호작용할 수 있다. 인공 및 천연 융합 효소는 근접한 연속적인 반응을 강제하거나 경로의 순차적인 효소들이 서로 작용하게 하여 중간체가 자유 발산되지 않고 대사체로 직접적으로 전환되게 하여 기질 채널링(substrate channeling) 현상을 나타낸다. 단백질 융합은 동일한 중간체를 이용하는 다른 효소와의 경쟁을 제거함으로써 대사적 중간체 전달의 효과적이고, 배타적이고 보호된 수단으로 표적 대사체 경로를 제공할 수 있다.

[0016] 본 발명은 FPP 신타제와 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드들을 융합시킴으로써 생성된 FPP가 근접한 거리에 생성된 α-파르네센 신타제에 의해 사용될 수 있어 α-파르네센의 생산성을 향상시킬 수 있다. FPP는 파르네센의 전구체로서 사용될 뿐만 아니라 퀴논 및 지질 담체와 같은 세포 구성요소의 생합성에서 사용되므로, FPP는 생성된 후 세포 구성 요소의 생합성에 사용되거나 α-파르네센 신타제에 의해 사용될 수 있다. 하지만 본 발명과 같이 FPP 신타제와 α-파르네센 신타제를 근접하게 발현되도록 함으로써 생성된 FPP가 세포 구성요소의 생합성에 이용되기보다 근접하게 발현된 α-파르네센 신타제에 의해 보다 더 쉽게 이용되어 α-파르네센

의 생산성을 향상시킬 수 있다.

[0017] 본 발명의 다른 구체예에서, 상기 융합 폴리뉴클레오티드는 상기 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드와 상기 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드 사이에 하나 이상의 임의의 아미노산을 코딩하는 폴리뉴클레오티드 링커(linker)가 더 결합된 것인 에세리키아속 미생물을 제공한다.

[0018] 상기 "하나 이상의 임의의 아미노산"은 상기 융합되는 두개 단백질 각각의 기능이 융합된 단백질의 형태에서도 유지되는 채로 발현될 수 있도록 하는 것이면 어떠한 것이라도 포함할 수 있다.. 예를 들면, 상기 하나 이상의 임의의 아미노산은 GGGGSGGGGS, $(Gly_4-Ser)_3$, $(Gly_2-Ser)_2$, 및 $Gly_4-Ser-Gly_5-Ser$로 이루어진 군으로부터 선택된 것일 수 있다. FPP 신타제를 코딩하는 폴리뉴클레오티드와 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드 사이에 하나 이상의 임의의 아미노산으로 이루어진 링커가 결합됨으로 인하여 FPP 신타제와 α-파르네센 신타제의 발현이 용이하게 될 수 있다. 링커 펩타이드 없는 직접적인 단백질 융합은 융합 파트너 간에 입체적 장애를 유발하여 각자의 구조적 배열 및 효소 활성에 영향을 미칠 수 있다.

[0019] 본 발명의 다른 구체예에서, 상기 에세리키아속 미생물은 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트란스퍼라제 및 히드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 폴리뉴클레오티드, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 유래의 포스포메발로네이트 키나제를 코딩하는 폴리뉴클레오티드, 스트렙토코커스 뉴모니애 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 폴리뉴클레오티드 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 폴리뉴클레오티드로 더 형질전환될 수 있다.

[0020] 본래 에세리키아속 미생물은 이소펜테닐 디포스페이트 (IPP) 및 디메틸알릴 디포스페이트(DMAPP)를 생산하기 위해 내부 메틸 메틸-에리트리톨 포스페이트(MEP) 경로를 이용하나 IPP 및 DMAPP 생산량을 증가시키기 위해 메발로네이트(MVA) 경로를 이용하도록 형질전환시킬 수 있다. 상기 엔테로코커스 패칼리스 유래의 유전자들과 스트렙토코커스 뉴모니애 유래의 유전자들은 MVA 경로를 진행하는데 필요한 효소들의 유전자들이다. 도 1에 α-파르네센을 생산하기 위한 MEP 경로와 MVA 경로를 간단하게 도식화하였다. 상기 에세리키아속 미생물이 MVA 경로를 사용하도록 형질전환되면 IPP 및 DMAPP 생산량이 증가하고 이에 따라 파네실 피로포스페이트(FPP) 생산량도 증가하여 α-파르네센 생산을 증가시킬 수 있다.

[0021] 본 발명의 다른 양태는 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드 및 에세리키아속 미생물에 코돈-최적화(codon-optimized)된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드를 포함하는 융합 폴리뉴클레오티드를 제공한다. 상기 융합 폴리뉴클레오티드는 FPP 신타제를 코딩하는 폴리뉴클레오티드 및 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드 사이에 하나 이상의 임의의 아미노산을 코딩하는 폴리뉴클레오티드 링커가 더 결합된 융합 폴리뉴클레오티드일 수 있다. 상기 링커가 더 결합된 융합 폴리뉴클레오티드는 서열번호 36의 염기 서열을 갖는 것일 수 있다.

[0022] 본 발명의 일 구체예에 따르면, 상기 에세리키아속 미생물에 코돈-최적화된 말루스 엑스 도메스티카(Malus x domestica) 유래의 α-파르네센 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 1의 염기 서열, 상기 대장균 유래의 FPP 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 2의 아미노산 서열, 상기 하나 이상의 임의의 아미노산을 코딩하는 폴리뉴클레오티드 링커는 서열번호 3의 아미노산 서열을 갖는 것일 수 있다.

[0023] 본 발명의 일 구체예에 따르면, 상기 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트란스퍼라제 및 히드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 폴리뉴클레오티드는 서열번호 4, 상기 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 폴리뉴클레오티드는 서열번호 5, 상기 스트렙토코커스 뉴모니애 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 폴리뉴클레오티드는 서열번호 6, 상기 스트렙토코커스 뉴모니애 유래의 포스포메발로네이트 키나제를 코딩하는 폴리뉴클레오티드는 서열번호 7, 상기 스트렙토코커스 뉴모니애 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 폴리뉴클레오티드는 서열번호 8, 및 상기 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 폴리뉴클레오티드는 서열번호 9의 아미노산 서열을 갖는 것일 수 있다.

[0024] 본 발명의 다른 양태는 상기 미생물을 배양하는 단계; 및 배양물로부터 α-파르네센을 분리하는 단계를 포함하는 미생물로부터 α-파르네센을 생산하는 방법을 제공한다.

[0025] 상기 방법에서 상기 미생물을 배양하는 것은 배지 위에 데칸 상(decane phase)을 위치시켜 배양하는 것일 수 있

다. 따라서 휘발성이 있는 α-파르네센을 손실 없이 수득할 수 있고, 데칸에 용해되어 있는 α-파르네센을 수득함으로써 수확 및 정제 방법을 모두 단순화하여 생산 비용을 줄일 수 있다.

[0026] 본 발명의 에세리키아 속 미생물로부터 α-파르네센을 생산하는 방법은 본 발명의 미생물을 배양하는 단계를 포함한다. 본 발명의 방법에 있어서, 배양은 합성, 반합성, 또는 복합 배양 배지에서 배양할 수 있다. 배양 배지로는 탄소원, 질소원, 비타민 및 미네랄로 구성된 배지를 사용할 수 있다. 예를 들어, MRS (Man-Rogosa-Sharp) 액체 배지 및 우유가 첨가된 액체 배지를 사용할 수 있다. 탄소원으로는 전분, 포도당, 자당, 갈락토스, 과당, 글리세롤 및 이들의 혼합물로 이루어진 군으로부터 선택된 것을 사용할 수 있으며, 바람직하게는 글리세롤을 사용할 수 있다. 질소원으로는 황산암모늄, 질산암모늄, 질산나트륨, 글루탐산, 카사미노산, 효모추출물, 펩톤, 트립톤, 대두박 및 이들의 혼합물로 이루어진 군으로부터 선택된 것을 사용할 수 있으며, 미네랄은 염화나트륨, 인산제이칼륨, 황산마그네슘 및 이들의 혼합물로 이루어진 군으로부터 선택된 것을 사용할 수 있다. 미생물 배양 배지 내 상기 탄소원, 질소원 및 미네랄 각각은 리터당 10 내지 100 g, 5 내지 40 g 및 0.5 내지 4 g을 이용할 수 있다.

[0027] 상기 배양 배지에 첨가되는 비타민은 비타민 A, 비타민 B, 비타민 C, 비타민 D, 비타민 E 및 이들의 혼합물로 이루어진 군으로부터 선택될 수 있다. 비타민은 상기에서 언급한 탄소원, 질소원, 미네랄과 등과 동시에 배지에 첨가하거나, 멸균하여 준비된 배지에 첨가할 수 있다.

[0028] 배양은 통상의 대장균 배양 조건으로 수행할 수 있으며, 예를 들어 약 15-45?에서 배양할 수 있다. 배양액 중의 배양 배지를 제거하고 농축된 균체만을 회수하기 위해 원심분리 또는 여과과정을 거칠 수 있으며 이러한 단계는 당업자의 필요에 따라 수행할 수 있다. 농축된 균체는 통상적인 방법에 따라 냉동하거나 냉동건조하여 그 활성을 잃지 않도록 보존할 수 있다.

[0029] 상기 배양의 일 예에 있어서, 탄소원으로서 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 예를 들면, 0.1 % (v/v) 내지 5 % (v/v)의 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 또한 배양의 일 예에 있어, MVA 경로의 중간 산물, 예를 들면 메발로네이트를 포함하는 배지에서 이루어지는 것일 수 있다. 미생물이 하부 MVA 경로만을 이용하도록 형질전환된 경우, 예를 들면 0.1mM 내지 15mM의 메발로네이트가 부가된 배지에서 배양이 이루어질 수 있다. 또한 배양의 일 예에 있어, 상기 배양은 IPTG 유도와 함께 이루어질 수 있다. 예를 들면, 0.01mM 내지 1mM의 IPTG 유도하에 배양이 이루어질 수 있다. 배지는 2YT 배지 (1L 당 16 g 트립톤, 5 g 소듐 클로라이드, 및 10 g 효모 추출물)일 수 있고 항생제를 필요로 하는 만큼 첨가할 수 있으며 예를 들면, 앰피실린 100 ㎍/mL, 클로람페니콜 50 ㎍/mL, 및 카나마이신 50 ㎍/mL을 첨가할 수 있다. 또한 배양의 일 예에 있어서, 항생제, 글리세롤, 및/또는 메발로네이트가 첨가된 2YT 배지 위에 데칸 상이 위치할 수 있다.

[0030] 본 발명의 미생물에 대하여는 상기한 바와 같다.

[0031] 본 발명의 방법은 또한, 배양물로부터 α-파르네센을 분리하는 단계를 포함한다. 상기 α-파르네센을 분리하는 것은 당업계에 잘 알려져 있다. 구체적으로, 2-상 배양으로부터 데칸상을 수집하고 12,000 rpm에서 10분 동안 원심분리시켜 세포 파편을 제거하고, 가스 크로마토 그래피 (GC) 및 가스 크로마토그래피-질량 분석 (GC-MS)을 수행하여 분리될 수 있다.

발명의 효과

[0032] 본 발명의 에세리키아 속 미생물을 이용하면 α-파르네센을 높은 효율로 생산할 수 있다.

[0033] 본 발명의 α-파르네센을 생산하는 방법에 의하면, α-파르네센을 높은 효율로 생산할 수 있다.

도면의 간단한 설명

[0034] 도 1a는 α-파르네센 합성 경로에 대한 간략한 다이어그램을 나타낸다. IPP 및 DMAPP는 MEP 경로 또는 MVA 경로에 의해 합성된다. MEP 경로는 글리세르알데하이드 3-포스페이트 및 피루베이트의 축합에 의해 시작된다. MVA 경로는 출발 물질로서 아세틸-CoA를 사용한다. FPP는 1 개의 DMAPP 및 2 개의 IPP의 축합을 통해 IspA에 의해 합성된다. α-파르네센 신타제는 전구체 FPP로부터 α-파르네센 형성시키는 촉매 작용을 한다. 보통 화살표 및 점선 화살표는 각각 단일 및 다중 효소 반응을 나타낸다.

도 1b는 본 발명에서 사용된 합성 플라스미드의 유전자 요약도를 나타낸다. 오각형 및 직사각형은 각각 프로모터 및 유전자를 나타낸다. 채워진 직사각형은 ispA 및 aFS 사이의 융합 연결을 나타낸다. RBS는 둥근 직사각형으로 나타낸다. 도 1a 및 도 1b에 사용된 약어의 의미는 다음과 같고 이 약어들은 본 명세서의 다른 부분에서도

동일한 의미로 사용된다. G-3-P는 글리세르알데히드 3-포스페이트이고; DXP는 1-데옥시-D-자일룰로즈-5-포스페이트(xylulose-5-phosphate)이고; MEP는 메틸-에리트리톨 포스페이트이고; DMAPP는 디메틸알릴 피로포스페이트이고; IPP는 이소펜테닐 피로포스페이트이고; FPP는 (E,E)-파르네실 피로포스페이트이고; FOH는 파르네솔이다. DXS는 DXP 신타제이고, IDI는 IPP 이소머라제이고; IspA는 FPP 신타제이고; (?)는 대장균 내부 비특이적 포스파타제를 나타내고; 및 FS는 α-파르네센 신타제이다. aFS는 인공적으로 코돈-최적화된 α-파르네센 신타제 유전자를 나타낸다.

도 2는 α-파르네센 생산 및 세포 성장에 있어서 α-파르네센 신타제 유전자의 코돈 최적화 효과를 나타낸다. 도 2a에서 pTwFS 또는 pTaFs를 갖는 재조합 대장균의 α-파르네센 생산량이 나타난다. 흰색 막대는 24시간, 검은색 막대는 48시간 배양 후 얻은 α-파르네센 생산량을 나타낸다. 도 2b는 빈 벡터 pTrc99A, pTwFS 또는 pTaFS를 갖는 재조합 대장균의 세포 성장을 각각 정사각형, 원 또는 삼각형으로 나타낸다. 배양은 29도, 1.0% (v/v) 글리세롤을 포함하는 2YT 배지에서 IPTG 없이 또는 0.1 mM IPTG 하에서 24시간 동안 수행되었다.

도 3은 IPP 및 DMAPP 합성의 증가가 α-파르네센 생산 및 세포 성장에 미치는 영향을 나타낸다. 대장균 aFS-CT, 대장균 aFS-dxs, 대장균 aFS-idi 및 대장균 aFS-SN을 29도에서, 1% (v/v)글리세롤를 포함하는 2YT 배지에서, 0.1 mM IPTG 유도 없이 또는 유도 하에서 배양시켰다. 대장균 aFS-SN의 배지에는 5mM 메발로네이트를 보충하였고, 흰색 막대는 24시간, 검은색 막대는 48시간 배양시킨 후 결과를 나타낸다.

도 4는 α-파르네센 생산 및 세포 성장에 있어 IspA 발현 증진의 효과를 나타낸다. 대장균 ispAaFS-CT 및 대장균 ispAaFS-SN를 29도에서, 1% (v/v) 글리세롤을 포함하는 2YT 배지에서 IPTG 유도 없이 또는 0.1mM IPTG 유도 하에서 배양시켰다. 대장균 ispAaFS-SN의 배양액에는 5mM 메발로네이트를 보충하였고 흰색 막대는 24시간, 검은색 막대는 48시간 배양시킨 후 결과를 나타낸다.

도 5는 α-파르네센 생산 및 세포 성장에 있어 IspA 및 α-파르네센 신타제 융합이 미치는 영향을 나타낸다. 대장균 ispANaFS-SN 및 ispALaFS-SN을 29℃에서, 1.0% (v/v) 글리세롤 및 5mM 메발로네이트를 포함하는 2YT 배지에서 IPTG 유도 없이 또는 0.1 mM IPTG 유도하에서 배양시켰다. 흰색 막대는 24시간, 검은색 막대는 48시간 배양시킨 후 결과를 나타낸다.

도 6은 첨가한 메발로네이트 양에 따른 대장균 ispALaFS-SN의 α-파르네센 생산 및 세포 성장을 나타낸다. 배양은 29℃, 0.1 mM IPTG 유도 하에서, 1.0% (v/v) 글리세롤 및 0 내지 15 mM 메발로네이트를 포함한 2YT 배지에서 수행되었다. 흰색 막대는 24시간, 검은색 막대는 48시간 배양시킨 후 결과를 나타낸다.

도 7은 글리세롤 농도에 따른 대장균 ispALaFS-NA의 α-파르네센 생산 및 세포 성장을 나타낸다. 배양은 29℃, 0.1 mM IPTG 유도 하에서, 0 내지 4.0% (v/v) 글리세롤을 포함하는 2YT 배지에서 수행되었다. 흰색 막대는 24시간, 검은색 막대는 48시간 배양시킨 후 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0035] 이하 본 발명을 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 본 발명을 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다. 이하의 실시예에서 다음의 실험 재료 및 방법을 사용하였다.

[0036] (1) 배지 및 배양 조건

[0037] E. coli DH5α (F^-,Φ80dlacZ△M15, △(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(r_K^- m_K^+), phoA, supE44, λ^-,thi-1)를 2YT 배지 (1 L 당 16 g 트립톤, 10 g 효모 추출물, 및 5 g 소듐 클로라이드)에서 플라스미드 제조를 위해서는 37℃에서, α-파르네센 생산을 위해서는 29℃에서 배양시켰다. 37℃에서 밤새 배양시킨 종자 배양액(seed culture)을 1% (v/v) 글리세롤을 포함하는 2YT 배지에서 OD600, 0.1이 될 때까지 접종하였고, 다른 특별한 언급이 없으면 α-파르네센 생산을 위해 29℃에서 배양시켰다. pSSN12Didi 플라스미드를 포함하는 대장균의 배양액에 메발로놀락톤으로부터 제조된 메발로네이트를 다른 언급이 없으면 5 mM의 농도로 보충하였다. 항생제 앰피실린 (100 ㎍/mL) 및 클로람페니콜 (50 ㎍/mL), 및 0.1 mM의 IPTG가 필요한 경우 첨가되었다. α-파르네센의 휘발성때문에, 배양 동안 생산되는 α-파르네센의 회수를 위해 초기에 5mL의 배양액 위에 1mL의 데칸을 덮었다. 세포 성장은 600 nm (OD600)에서 광학적 밀도를 측정하는 것에 의해 측정되었고 모든 실험은 삼중으로 수행되었다.

[0038] (2) 플라스미드의 제조

[0039] 제한 효소 절단 및 형질 전환을 포함하는 기본적인 분자 생물학 기술을 문헌(Sambrook and Russell, 2001)에 기술된 대로 수행하였다. 제조자의 지시에 따라 Pfu DNA 폴리머라제(SolGent, Daejeon, Korea)를 사용하여 PCR에 의해 DNA 단편을 증폭시켰다. 본 발명에서 사용되는 플라스미드 및 프라이머는 표 1에 나열된다.

표 1

[0040]

이름	특성	참조
플라스미드		
pSTV28	Plac 발현 벡터, pACYC184 origin, lacZ, Cmr	Takara Co., Ltd.
pSdxs	E. coli 유래 dxs를 포함하는 pSTV28 벡터	
pSidi	E. coli 유래 idi를 포함하는 pSTV28 벡터	
pSSN12Didi	E. coli 유래의 idi, S. pneumonia 유래 mvaK1, mvaK2 및 mvaD를 포함하는 pSTV28 벡터	
pSNA	E. faecalis 유래 mvaE 및 mvaS, S. pneumoniae유래 mvaK1, mvaK2, 및 mvaD 및 E. coli 유래 idi 를 포함하는 pSTV28 벡터	
pET-30a /MdAFS1	wFS 를 포함하는 pET-30a 벡터	(Green et al., 2007)
pTrc99A	Ptrc 발현 벡터, pBR322 유래, lacIq, Ampr	(Amann et al., 1988)
pTwFS	wFS를 포함하는 pTrc99A 벡터	
pUC57aFS	aFS를 포함하는 pUC57 벡터	
pTaFS	aFS를 포함하는 pTrc99A 벡터	
pTispAaFS	aFS 및 E. coli 유래 ispA 를 포함하는 pTrc99A 벡터	
pTispANaFS	aFS 및 ispA 사이에 링커가 없는 aFS 및 ispA의 융합 유전자를 포함하는 pTrc99A 벡터	
pTispALaFS	aFS 및 ispA 사이에 $(GGGGS)_2$ 링커를 포함하는 aFS 및 ispA의 융합 유전자를 포함하는 pTrc99A 벡터	
프라이머		
aFS-F1	GGGATCC**ATG**GAATTTCGTGTCCACCTG	서열번호 28
aFS-F2	GGGATCCAGGAGGTAATAATA**ATG**GAATTTCGTGTCCACCTG	서열번호 29
aFS-R	CCCAAGC**TTA**GTTGACCAGCGGCTGAAACAG	서열번호 30
ispA-F	GCGAATTC**ATG**GACTTTCCGCAGCAAC	서열번호 31
ispA-R1	GGGATCC**TTA**TTTATTACGCTGGATGTAG	서열번호 32
ispA-R2	GCGGATCCTTTATTACGCTGGATGTAGTC	서열번호 33
ispA-R3	GCGGATCCGCCGCCACCCGAGCCACCGCCACCTTTATTACGCTGGATGTAGTC	서열번호 34

[0041] 참고: 시작 코돈과 종결 코돈은 두꺼운 글씨체로 표시하고 제한효소 사이트는 밑줄을 그어 표시하였다.

표 2

[0042]

효소명	유전자	유전자 서열 (Genebank 허가번호)
엔테로코커스 패칼리스 (Enterococcus faecalis) 유래의 아세틸-CoA 아세틸트란스퍼라제 및 히드록시메틸글루타릴 (HMG)-CoA 리덕타제	mvaE	서열번호 12 (AF290092)
엔테로코커스 패칼리스 유래의 HMG-CoA 신타제	mvaS	서열번호 13 (AF290092)
스트렙토코커스 뉴모니애 (Streptococcus pneumoniae) 유래의 메발로네이트 키나제	mvaK1	서열번호 14(AF290099)
스트렙토코커스 뉴모니애 유래의 포스포메발로네이트 키나제	mvaK2	서열번호 15(AF290099)
스트렙토코커스 뉴모니애 유래의 메발로네이트 디포스페이트 데카르복실라제	mvaD	서열번호 16 (AF290099)
대장균 유래의 이소펜테닐 디포스페이트(IPP) 이소머라제	idi	서열번호 17 (U00096)

표 3

[0043]

유전자	프라이머서열		제한효소
mvaE	F	서열번호 20	Sac I
	R	서열번호 21	Sma I
mvaS	F	서열번호 22	Sma I
	R	서열번호 23	BamH I
mvaK1, mvaK2, mvaD	F	서열번호 24	Kpn I
	R	서열번호 25	Xba I
idi	F	서열번호 26	Sma I
	R	서열번호 27	Sph I

[0044] 1) pSdxs, pSidi, pSSN12Didi 및 pSNA의 제조

[0045] ① pSdxs, pSidi의 제조

[0046] 대장균 유래 dxs 유전자의 DNA 주형 및 프라이머(서열번호 37, 38)를 이용하여 증폭시킨 PCR 산물을 EcoRI과 BamHI 제한효소를 이용하여 pSTV28 벡터에 삽입하여 pSdxs를 제조하였다. 대장균 유래 idi 유전자의 DNA 주형 및 프라이머(서열번호 39, 40)를 이용하여 증폭시킨 PCR 산물을 BamHI과 SalI 제한효소를 이용하여 pSTV28 벡터에 삽입하여 pSidi를 제조하였다.

[0047] ② pSSN12Didi의 제조

[0048] 먼저, 스트렙토코커스 뉴모니애의 염색체 DNA 주형 및 서열번호 24 및 25의 염기서열을 갖는 프라이머를 이용하여 mvaK1, mvaD 및 mvaK2 오페론을 증폭하였으며(mvaK1, mvaK2, mvaD는 하나의 오페론으로 염색체상에 존재해서 각각의 유전자를 PCR 클로닝하지 않고 오페론을 통째로 한번에 PCR 클로닝을 하였다. 증폭된 유전자 단편을 KpnI 및 XbaI으로 처리한 후 동일 효소로 절단된 pBAD24 벡터(Guzman, L. M. 등, J. Bacteriol., 177, 4121-4130, 1995)에 삽입하여 플라스미드 pDSN12D를 제조하였다. 또한, idi 유전자 단편을 대장균 염색체 DNA 주형 및 서열번호 26 및 27의 염기서열을 갖는 프라이머를 이용하여 증폭한 후, 이를 SmaI과 SphI으로 처리하여 동일 효소로 절단된 플라스미드 pDSN12D에 삽입하여 플라스미드 pDSN12Didi를 제조하였다. 그 후, 플라스미드 pDSN12Didi를 BamHI과 SalI으로 절단하여 상기 4개의 유전자가 포함된 유전자 단편을 얻은 후 이를 동일 효소로 처리된 pSTV28 벡터(Takara Korea, 한국) (서열번호 25)에 삽입하여 플라스미드 pSSN12Didi를 제작하였다.

[0049] ③ pSNA의 제조

[0050] 표 2의 유전자들을 표 3에 열거된 프라이머들을 사용하여, 해당 유전자를 포함하고 있는 균주의 염색체 DNA를 주형으로 한, PCR을 통하여 증폭하였다. 증폭된 산물을 표 3에 열거된 제한 효소를 이용하여 pSTV28 벡터(Takara Korea, 한국) (서열번호 18)에 도입하여, 벡터 pSNA를 제조하였다. 벡터 pSNA는 아세틸-CoA로부터 IPP를 생산할 수 있는 메발로네이트 경로의 효소를 코딩하는 유전자를 모두 포함하고 있다.

[0051] 2) pTwFS, pTaFS, pTispAaFS, pTispAN, pTispAL, pTispANaFS 및 pTispALaFS 의 제조

[0052] Green 박사(Horticultural and Food Research Institute, Auckland, New Zealand)가 pET-30a/MdAFS1 플라스미드를 기증하였고 (Green et al., 2007), 이를 EcoRI 및 XhoI로 절단하여 사과 α-파르네센 신타제 유전자의 야생형(wFS: 서열번호 35, GenBank accession number : AY563622.1)을 얻었다. 생성된 단편을 pTrc99A(Amann et al., 1988)의 EcoRI-SalI 부위에 삽입하여 pTwFS를 생산하였다. pUC57-aFS의 플라스미드는 GenScript Corp(NJ, US)의 최적 대장균 코돈 사용에 따라 합성된 인공 α-파르네센 신타제 유전자 (aFS)를 포함한다. aFS 유전자는 프라이머 aFS-F2 및 aFS-R를 사용하여 PCR에 의해 pUC57-aFS로부터 증폭되었다. 생성된 PCR 생산물을 BamHI 및 HindIII로 절단하고 pTrc99A의 상응하는 제한 효소 부위에 삽입하여 pTaFS를 생성하였다. pTispAaFS 플라스미드의 제조를 위해, 대장균 게놈 DNA로부터 프라이머 ispA-F 및 ispA-R1를 사용하여 PCR에 의해 증폭된 ispA 유전자를 EcoRI 및 BamHI으로 절단된 pTaFS 플라스미드에 삽입하였다. ispA 및 aFS를 융합하기 위해, ispA 유전자를 ispA-F 및 ispA-R2 프라이머를 사용하여 PCR에 의해 증폭시켜 정지 코돈을 제거하고 EcoRI 및 BamHI의 절단 효소 사이트를 도입하여 ispAN으로 이름지었다. 동일하게 ispA-F 및 ispA-R3 프라이머로 PCR 생산물 ispAL을 생성시켰고, 이는 ispAN과 비교할 때 유연한 $(GGGGS)_2$ 링커(서열번호 11)를 BamHI 사이트 전에 포함한다. 이 두 개

의 PCR 생산물을 pTrc99A 벡터의 EcoRI-BamHI 사이트에 클로닝시켜 각각 pTispAN 및 pTispAL를 생성시켰다. aFS-F1 및 aFS-R 프라이머로 증폭시킨 aFS 유전자를 BamHI 및 HindIII로 절단하여 pTispAN 및 pTispAL에 삽입하고 각각 pTispANaFS 및 pTispALaFS를 생성시켰다. α-파르네센 신타제 유전자를 포함하는 개략적인 플라스미드 구조가 도 1b에 나타난다. 본 발명에 사용된 PCR 프라이머는 표 1에 나열된다.

[0053] (3) α-파르네센의 식별 및 정량

[0054] 2-상 배양으로부터 데칸상을 수집하고 12,000 rpm에서 10분 동안 원심분리시켜 세포 파편을 제거하고, 가스 크로마토 그래피 (GC) 및 가스 크로마토그래피-질량 분석 (GC-MS)을 수행하였다. α-파르네센 및 FOH의 생산은 불꽃 이온화 검출기(flame ionization detector (FID))를 구비한 Agilent Technologies 7890A Gas Chromatograph를 사용하여 정량하였다. 샘플을 1:10의 분리(split) 비율에서 주입하였다. 1 ㎕ 샘플로부터 각 분석물이 19091J HP-5 컬럼 (길이, 30 m; 내부 지름, 0.32 mm; 필름 두께, 250 ㎛)상에서 분리되었다. 오븐 온도는 처음에 1분 동안 80℃로 유지되었고 10℃/분의 속도로 250℃까지 증가되어 1분 동안 유지되었다. 39 psi를 입구 압력으로 하여 질소를 담체 가스로서 사용하였다. 검출기 온도를 260℃에서 유지하였다. GC-MS 분석을 Double-Focusing Mass Spectrometer GC Mate II (JEOL, Tokyo, Japan) 상에서 수행하였다.

[0055] (4) 메발로네이트 및 글리세롤의 정량

[0056] 메발로네이트 농도는 GC 분석에 의해 결정되었다. 배양 여과액은 3 M HCl를 사용하여 pH 2로 조정하고 45℃에서 1 시간 동안 배양시키고, 무수 Na_2SO_4로 포화시키고, 최종적으로 에틸 아세테이트로 추출하였다. 생성된 샘플들을 메발로네이트 정량을 위해 Agilent Technologies 7890A Gas Chromatograph에 의해 분석하였다. GC의 분석 온도는 1분 동안 180℃의 초기 온도로 통제되고 2.5℃/분의 기울기에서 200℃까지 증가시키고 2분 동안 유지시켰다. 검출기 온도는 260℃에서 유지되었다. 글리세롤 농도는 굴절률 검출기 및 Kromasil KR100-5NH2 column (250 mm×4.6 mm, Eka Chemicals, Bohus, Sweden)을 갖는 고성능 액체 크로마토 그래피 (HPLC) (Shimadzu, Kyoto, Japan)에 의해 측정하였다. 배양액을 원심분리한 후 얻은 상층액을 잔여 글리세롤의 정량을 위해 주입하였다. 이동 상은 1.5 mL/분의 유속인 물 중의 75% 아세토니트릴이었다.

[0057] **실시예 1: 대장균으로부터 생산된 α-파르네센의 식별**

[0058] 사과 α-파르네센 신타제가 인 비트로 분석에서 FPP로부터 α-파르네센을 생산할 수 있다고 보고되었다. 대장균으로부터 α-파르네센을 생산하기 위해, 사과 α-파르네센 신타제 유전자의 야생형(wFS)을 pTrc99A 벡터 내에 클로닝시켜 pTwFS을 생성하고 pTwFS을 포함하는 재조합 대장균(대장균 wFS)을 29℃, 1mL 데칸이 덮어 씌워진 5mL의 2YT 배지에서 24시간 동안 배양시켰다. 데칸 층을 수집하여 GC 및 GC-MS에 의해 분석하였다. GC 크로마토그램에서 7.96분의 보유 시간(retention time)에서의 큰 피크가 대장균 wFS의 배양액에서만 관찰되었다(도 1c). 상기 피크는 그 후 GC-MS에 의해 식별되었다(도 1d). 이 피크의 질량 스펙트럼은 EI 질량 스펙트럼 라이브러리 및 발행된 문헌(Pechous and Whitaker, 2004)에 있는 α-파르네센 프로파일과 일치했다. 따라서, α-파르네센 형성은 대장균 내에서 사과 α-파르네센 신타제가 기능적으로 발현되었다는 것을 나타낸다.

[0059] **실시예 2: 대장균 내에서 높은 수준의 발현을 위한 α-파르네센 신타제의 코돈 최적화**

[0060] 코돈 사용 바이어스(bias)는 생물마다 다양해서, 이종 유래 발현의 장벽이 된다 (Gustafsson et al., 2004; Kane, 1995). wFS 유전자의 개개의 유전자 코돈의 빈도는 대장균의 것과 상당히 상이한 것으로 밝혀졌다 (http://www.kazusa.or.jp/codon/). wFS 유전자의 코돈 바이어스 분석 (http://pro.genomics.purdue.edu/emboss/)은 53.9의 Nc (Effective number of condon) 및 0.58의 CAI (Codon adaptation index)를 나타낸다. 높은 Nc 및 낮은 CAI 값은 대장균 내에서 사과 α-파르네센의 낮은 발현의 가능성을 나타낸다. 따라서, 인공 α-파르네센 신타제(aFS) 유전자를 대장균 코돈 사용에 따라 디자인하고 wFS 유전자의 378 코돈(전체 유전자 코돈의 65.5%)을 변형하여 코돈-최적화된 aFS 유전자의 NC 및 CAI 값이 각각 39.7 및 0.84이 나왔다.

[0061] aFS 유전자를 pTrc99A 벡터에 클로닝하여 pTaFS를 생성하였다. pTaFS를 포함한 대장균(대장균 aFS)을 대장균 wFS의 배양과 같은 조건으로 배양시켰다. SDS-PAGE 분석에서, α-파르네센 신타제의 눈에 보이는 밴드(Mw, 66

kDa)가 0.1 mM IPTG 유도된 대장균 aFS의 배양액에서만 관찰되었다. 이는 α-파르네센 신타제의 발현이 코돈-최적화에 의해 향상되었다는 것을 나타낸다. 또한, α-파르네센이 대장균 wFS와 비교할 때 대장균 aFS에서 더 높게 관찰되었지만, 두 균주 사이에 세포 성장의 차이는 관찰할 수 없었다 (도 2a). 0.1 mM IPTG 유도에 의해 두 균주 모두에서 극심한 세포 성장의 감소를 관찰하였다 (도 2b).

[0062] 비록 그들의 산출은 비-유도된 배양액보다 0.1 mM IPTG 유도된 배양액에서 상대적으로 더 낮지만, 그들의 성장 차이를 생각할 때, 두 균주의 세포 특이적 α-파르네센 산출은 0.1 mM IPTG 의 유도와 함께 급격하게 증가하였다. α-파르네센 생산의 증가는 α-파르네센 신타제의 높은 발현으로 야기된 것으로 결론지을 수 있었다.

[0063] 0.1 mM IPTG 유도 하에서 대장균 wFS 및 대장균 aFS의 세포 성장 억제는 단백질 과발현 또는 희박한 코돈 고갈로 인한 것이 아닐 것이다. α-파르네센 신타제의 발현 수준이 대사적 부담을 야기할 만큼 높지 않고 희박한 코돈은 코돈-최적화된 aFS 유전자에 존재하지 않기 때문이다. 성장 억제를 유발하는 이유는 세포 성장을 위한 필수 대사물 (IPP, DMAPP 및 FPP)의 부족으로 인한 것일 수 있다. α-파르네센 신타제의 증가된 발현은 α-파르네센 형성을 위한 그들의 소비를 더 심하게 하기 때문이다. 그러므로 대장균 aFS의 세포 성장 및 α-파르네센 생산에 있어 IPP, DMAPP 및 FPP 합성의 증가의 효과를 조사해볼 필요가 있었다.

[0064] **실시예 3: α-파르네센 생산을 위한 IPP 및 DMAPP 공급의 향상**

[0065] 대장균 내부 MEP 경로의 향상을 위해, 이 경로와 관련된 dxs 및 idi 유전자를 pSTV28에 클로닝하여 pSdxs 및 pSidi를 생성하고 pTaFS와 함께 같이 형질전환시켜 각각 대장균 aFS-dxs 및 대장균 aFS-idi를 얻었다. 또한 pSTV28 빈 벡터를 공동 형질전환시켜 대장균 aFS-CT를 대조군으로 얻었다 (도 3). 유도되지 않은 배양액에서 대장균 aFS-CT, 대장균 aFS-dxs 및 대장균 aFS-idi 사이에서 α-파르네센 생산 및 세포 성장에 있어 큰 차이는 없었다. 그러나, 0.1mM IPTG 유도된 배양액에서, 대장균 aFS-CT에 비해 대장균 aFS-idi로부터 7배 증가된 α-파르네센 생산을 얻었다. 또한 IPTG 유도에 의해 대장균 aFS에서 관찰되었던 세포 성장 억제는 대장균 aFS-dxs 및 대장균 aFS-idi의 배양에서 발견할 수 없었으나, 대장균 aFS-CT의 배양에서는 발견되었다. 이는 α-파르네센 생산을 위한 빌딩 블록의 소비로 인해 대장균 성장을 위한 IPP 및 DMAPP가 제한되었다는 것을 나타낸다. DXS 및 IDI와 같은 MEP 경로의 속도-결정 효소들의 과발현, 특히 IDI의 과발현은 성장 억제를 완화시키고 α-파르네센 생산을 향상시켰다. α-파르네센 합성을 위한 하나의 분자는 두 개의 IPP 및 하나의 DMAPP를 필요로 하나, MEP 경로는 IPP 및 DMAPP를 5:1의 비율로 생산한다. IPP 및 DMAPP의 소비와 형성 사이의 불균형은 IPP 및 DMAPP의 IDI 촉매 아이소머화의 과발현에 의해 제거할 수 있고, 이는 대장균 aFS-idi에서 α-파르네센 생산의 상당한 증가를 가져왔다.

[0066] α-파르네센 생산에 있어 MVA 경로의 효과를 메발로네이트의 보충에 의해 MVA 하부 경로를 포함하는 pSSN12Didi 플라스미드로 조사하였다. pTaFS 및 pSSN12Didi를 갖는 재조합 대장균 (대장균 aFS-SN)을 상기 기술된 동일한 배양 조건에서 5mM 메발로네이트를 추가하고 배양시켰다 (도 3). IPTG 유도 없이 대장균 aFS-SN에서 α-파르네센 생산의 두드러진 증가를 관찰하였다. 48시간 후, 비-유도 및 0.1 mM IPTG 유도된 배양에서 각각, 17.4 mg/L 및 17.8 mg/L의 α-파르네센 생산을 얻었다. 그러나, IPTG 유도된 배양의 세포 성장은 비-유도된 것에 비해 현저하게 낮았다. 48시간에서 IPTG 유도된 배양으로부터의 6.5mg/L/OD600의 세포 특이적 수득은 비-유도된 배양의 2.1mg/L/OD600 보다 3.2배 더 높았다. 또한 IPTG 유도된 배양에서 α-파르네센 생산의 2배의 현저한 증가는 24시간 후 세포 성장의 정체 상태에서도 유지되었다. 이러한 결과는 도입된 경로에 의한 IPP 및 DMAPP 합성의 증가가 α-파르네센 생산의 증가에 기여한다는 것을 의미한다. 세포 내 IPP 및 DMAPP의 축적에 의해 세포 성장이 억제된다고 보고된다 (Martin et al., 2003). IPTG 유도된 배양에서 대장균 aFS-SN의 성장 억제는 대장균 aFS-CT에서 관찰되는 IPP 및 DMAPP의 부족에 의한 것이기 보다는 IPP 및 DMAPP의 높은 수준으로 인한 것일 것이다. FPP 신타제 (IspA, ispA에 의해 인코드됨)은 두 개의 IPP 및 하나의 DMAPP의 압축에 의해 대장균에서 FPP 합성을 촉매한다 (Fujisaki et al., 1990). ispA 유전자의 단일 염색체 카피는 재조합 대장균에서 α-파르네센 생산을 위한 FPP의 많은 양을 합성하기에 충분하지 않을 것이다. 따라서, FPP 신타제의 증대는 IPP 및 DMAPP의 세포 내 축적을 방지하고 α-파르네센 생산을 위한 FPP 합성을 촉진시킬 것이다.

[0067] **실시예 4: α-파르네센 생산에 있어 IspA 과발현의 효과**

[0068] ispA 유전자를 pTaFS 플라스미드에 클로닝시켜 pTispAaFS를 얻었다 (도 1b). pTispAaFS 및 pSSN12Didi를 포함하는 재조합 대장균 ispAaFS-SN 을 상기 기술된 것과 같이 5mM 메발로네이트를 포함하는 2YT 배지에서 배양시켰

다. 또한 pTispAaFS를 pSTV28 빈 벡터와 함께 형질전환시켜 대조균 균주로 대장균 ispAaFS-CT를 얻었다. IPTG 유도된 48시간 배양에서 대장균 ispAaFS-SN의 α-파르네센 생산은 57.6 mg/L (도 4)로, 도 3에서 얻은 대장균 aFS-SN의 생산보다 3.2배 높았다. IspA의 과발현은 pSSN12Didi 플라스미드를 포함하는 균주에서 α-파르네센 생산을 효과적으로 촉진시켰으나, 대장균 ispAaFS-CT에서는 대장균 aFS-CT(도 3)와 동등한 α-파르네센 생산만을 관찰할 수 있었다. 이는 FPP 합성을 위한 IPP 및 DMAPP의 충분한 공급이 없다면, IspA 과발현에 의한 것만으로는 α-파르네센 합성이 증가될 수 없다는 것을 나타낸다. 또한 대장균 ispAaFS-SN는 IPTG 유도된 및 비-유도된 배양에서 부산물로서 FOH를 생산하였다. FOH 생산은 24시간 배양으로부터 현저하게 증가한 반면, 동일한 시간에 α-파르네센 생산의 증가는 한계점에 다다른다.

[0069] **실시예 5: α-파르네센 및 FOH 생산에 있어 IspA 및 α-파르네센 신타제의 융합 단백질의 효과**

[0070] α-파르네센 신타제의 FPP 이용가능성을 증가시키기 위해, IspA 및 α-파르네센 신타제를 인프레임(in-frame) 단백질 융합에 의해 공동 위치시켰고, 이는 α-파르네센 근처에 FPP 농도를 증가시킬것으로 예상되었다. α-파르네센 신타제로 FPP의 전달에 있어 단백질 융합은 α-파르네센 합성을 위해 세포질 내 대량의 액체 중에서 중간체를 자유 분산시키는 것보다 우수한 동역학적 특성을 나타낼 것이다. 기질 채널링(substrate channeling)이라고 하는 효소 촉매된 반응에 있어 기질 및 중간체의 채널링은 피루베이트 디카르복실라제와 같은 천연 효소 복합체에서 관찰된다. 기질 채널링은 통상적인 중간체의 경쟁하는 대사적 경로로의 불필요한 흐름을 감소시킬 수 있다. 기질 채널링은 또한 단백질 융합에 의해 수행될 수 있다. 따라서, ispANaFS 및 ispALaFS의 두 개의 융합 유전자를 각각 ispA 및 aFS 유전자의 직접적인 융합 또는 그들 사이에 $(GGGGS)_2$ 링커를 삽입하여 생성시켰다. ispANaFS 및 ispALaFS를 pTrc99A 벡터로 클로닝시켜 pTispANaFS 및 pTispALaFS (도 1b)를 생성시키고, pSSN12Didi와 함께 각각 공동 형질전환시켜 각각 대장균 ispANaFS-SN 및 대장균 ispALaFS-SN를 생성시켰다. 이 두 개의 균주는 상기 기재된 것과 같이 5mM 메발로네이트를 포함하는 2YT 배지에서 배양시켰다 (도 5). 대장균 ispANaFS-SN 및 대장균 ispALaFS-SN으로부터의 α-파르네센 생산은 비-유도된 배양에 비교하여 0.1 mM IPTG 유도된 배양에서 훨씬 높게 나타났다. 대장균 ispALaFS-SN의 배양에서, 단백질 융합을 갖지 않는 대장균 ispAaFS-SN(도 4)보다 세포 성장은 거의 증가되지 않은 반면 급격한 α-파르네센 생산의 증가를 관찰할 수 있었다. 대장균 ispALaFS-SN는 48시간 동안 IPTG 유도된 배양에서 86.8 mg/L의 α-파르네센을 생산하였다. 흥미롭게도 배양액 중에 FOH 부생성물 형성은 관찰되지 않았다. 이는 IspALaFS의 융합 단백질이 효과적으로 IPP 및 DMAPP를 α-파르네센 합성으로 이끌었다는 것을 나타낸다. 대장균 ispANaFS-SN의 배양에서, α-파르네센 생산 및 세포 성장은 대장균 ispAaFS-SN의 것과 유사하였고, 24시간까지 FOH의 형성이 없었음에도 불구하고 동등한 양의 FOH가 48 시간에서 여전히 수득되었다. IspANaFS 융합 단백질은 α-파르네센 생산에 있어 긍정적인 효과를 나타내는데 실패하였다. 링커 펩타이드 없는 직접적인 단백질 융합은 융합 파트너 간에 입체적 장애를 유발하여 각자의 구조적 배열 및 효소 활성에 영향을 미칠 수 있다. $(GGGGS)_2$ 링커를 사용하는 효소 융합 전략은 FPP 흐름을 FOH가 아닌 α-파르네센의 형성으로 돌릴 수 있다. IPTG 유도된 배양 조건에서 대장균 ispALaFS-SN으로부터의 더 높은 α-파르네센 생산하였고 pTispALaFS 및 0.1 mM IPTG 유도가 추가적인 연구를 위해 적용되었다.

[0071] **실시예 6: α-파르네센 생산에 있어 메발로네이트 농도의 효과**

[0072] 대장균 ispALaFS-SN의 α-파르네센 생산에 있어 메발로네이트 농도의 효과를 조사하기 위해 0 내지 15 mM의 다양한 농도의 메발로네이트를 첨가하였다 (도 6). 세포 성장 및 α-파르네센 생산이 첨가 농도가 증가함에 따라 증가하였다. 15mM 메발로네이트를 첨가한 48시간에서 153.6 mg/L의 가장 높은 α-파르네센 생산량을 얻었고 5 mM 메발로네이트를 첨가하였을 때 얻은 86.8 mg/L의 α-파르네센 생산과 비교할 때 1.8배 증가한 것이다. 메발로네이트를 첨가하지 않은 배양에서 단지 9.8 mg/L의 α-파르네센이 얻어졌고 세포 성장 또한 메발로네이트를 첨가한 배양에 비해 훨씬 낮았다. α-파르네센 합성과 세포 성장이 내부 MEP 경로에 의해서만 제한적인 양으로 보충되는 IPP 및 DMAPP를 사용하는데 경쟁하였기 때문이다. 남은 메발로네이트 분석은 많은 양의 메발로네이트가 모든 메발로네이트 첨가된 배양에 남아있고 15 mM 메발로네이트 첨가 배양에서 48시간 동안 처음 첨가된 메발로네이트의 28%만이 소비되었다는 것을 보여준다. 메발로네이트 소비 및 α-파르네센 생산은 24 시간에서 설명할 수 없이 낮은 α-파르네센 생산이 관찰되는 15 mM 메발로네이트 첨가 배양을 제외하고, 배양의 처음 24시간 동안 거의 진행된다. 비록 α-파르네센 생산 및 세포 성장이 메발로네이트를 높은 농도로 첨가하는 것에 의해 증진되지만, 메발로네이트는 비싼 물질이고 많은 양의 잔존 메발로네이트가 낭비되기 때문에 이는 α-파르네센 생산을 위한 경제적인 방법이 아니다. 배양 배지에 있어 높은 메발로네이트 농도의 이익은 메발로네이트의

농도 구배에 의해 결정되는 확산 수송에 의한 메발로네이트의 세포적 흡수의 증가 때문인 것으로 보인다. 그러므로, α-파르네센의 대량 생산을 위해 기질로서 메발로네이트를 요구하는 하부 경로(pSSN12Didi) 대신 내부적으로 풍부한 대사 산물인 아세틸-CoA로부터 시작하는 전체 MVA 경로를 사용하는 것이 필요하다.

[0073] **실시예 7: 전체 MVA 경로를 사용하는 α-파르네센 생산**

[0074] pTispALaFS 플라스미드를 pSNA와 함께 같이 형질전환시켜 대장균 ispALaFS-NA를 형성시켜, 0.1 mM IPTG 및 다양한 농도의 글리세롤 (0 - 4%, (v/v))을 포함하는 2YT 배지에서 배양시켰다. 그의 풍부함, 낮은 가격 및 높은 정도의 적합성(reduction) 때문에 탄소 및 에너지 소스로서 글리세롤을 선택하였다 (Murarka et al., 2008; Yazdani and Gonzalez, 2007). 또한 글리세롤은 바이오디젤 생산 과정에서 부산물로서 생산되어 새로운 용도가 탐색되고 있다 (da Silva et al., 2009; Fukuda et al., 2001). 결론적으로 α-파르네센의 생산은 대장균 ispALaFS-NA의 배양에 글리세롤의 추가에 의해 급격하게 증가하였다. 2.0% (v/v)까지는 글리세롤 농도에 따라 증가하였고 3.0% (v/v)의 글리세롤 첨가에서 증가의 한계점에 도달하였다 (도 7). 380.0 mg/L의 가장 높은 α-파르네센 생산이 3.0% (v/v) 글리세롤이 첨가된 48시간 배양에서 얻어졌다. 따라서, 메발로네이트의 첨가를 요구하는 하부 MVA 경로를 갖는 대장균보다 효율적으로 전체 MVA 경로를 포함하는 재조합 대장균은 값싼 탄소 소스로부터 α-파르네센을 생산할 수 있었다. 대장균 ispALaFS-NA로부터의 α-파르네센 생산은 글리세롤의 존재하에서 24시간 후 현저하게 증가하였고, 2.0% (v/v) 글리세롤이 첨가된 배양에서 24시간에 비해 48시간에서 10배 증가가 관찰되었다. 24시간 후의 이러한 급격한 증가는 메발로네이트 중간체의 생산에서도 관찰되었다. 24시간에서 2.5 mM 미만의 메발로네이트가 글리세롤 농도에 따른 큰 차이 없이 모든 배양액에서 관찰된 반면, 2.0% (v/v)보다 많은 글리세롤이 첨가된 배양에서 상당한 양의 메발로네이트가 48시간에 축적되었다. 23.8 mM의 가장 높은 메발로네이트 축적은 4.0% (v/v) 글리세롤 첨가된 배양에서 관찰되었다. 2.0% (v/v)보다 많은 글리세롤의 첨가는 α-파르네센 생산에 있어 증가를 나타내지 않았지만 글리세롤 농도가 증가함에 따라 메발로네이트 축적은 증가하였다. 2.0% (v/v) 글리세롤 첨가된 배양에서 생산된 메발로네이트 양은 MVA 하부 경로에 포화된 농도로써 대장균 ispALaFS-NA의 α-파르네센 생산에 충분하였다. 배양액에서 메발로네이트의 높은 축적은 MVA 상부 경로(아세틸-CoA로부터 메발로네이트까지)에 비해 MVA 하부 경로의 낮은 효율을 의미한다. 그러므로, 메발로네이트의 과생산으로부터 기인하는 대사물의 낭비를 최소화하기 위해 균형잡힌 MVA 경로를 개발하는 것이 필요하다(예, 낮은 효율의 효소를 더 좋은 효율의 효소로 대체함으로써 속도를 제한하는 단계를 제거하는 것).

[0075] 모든 글리세롤 첨가된 배양에서 24시간 후 세포 성장에 있어 현저한 차이는 없었으나, 48시간 후에는 글리세롤 첨가의 증가에 따라 성장도 증가하였다. 글리세롤 첨가 없는 배양에서, α-파르네센 생산 및 세포 성장은 탄소 및 에너지원의 제한으로 인해 극심하게 제한되었다. 잔여 글리세롤은 1.0% (v/v) 미만으로 글리세롤이 첨가된 배양에서 24시간 후 제한되었고, 0.5% (v/v) 글리세롤 첨가된 배양에서 48시간에는 고갈되었다. 그러나, 2.0% (v/v)보다 많은 글리세롤이 첨가된 배양에서는 글리세롤이 제한되지 않았다. 세포 성장 및 메발로네이트와 α-파르네센의 생산 모두 24시간 후의 배양에서도 현저하게 증가하였다. 특히, α-파르네센 및 메발로네이트의 생산은 세포 성장에 비해 24 시간 후 두드러지게 증가하였다. 전반적인 세포 대사의 중요한 대사물인 아세틸-CoA은 지수 증식기(logarithmic phase)에 주로 세포 성장을 위해 사용되고 세포가 활발하게 자라지 않은 때인 24시간 후에는 메발로네이트 경로를 위해 사용 가능하게 되는 것으로 보인다. 표 4는 대사 공학된 대장균의 단계적 α-파르네센 생산의 증가를 나타낸다.

표 4

[0076]

균주	배양 조건	α-파르네센 생산 (mg/L)	증가 정도(배)
대장균 wFS	2YT medium, 29℃, 1.0% (v/v) glycerol	1.2	1
대장균 aFS	2YT medium, 29℃, 1.0% (v/v) glycerol	1.6	1.3
대장균 aFS-dxs	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG	1.9	1.6
대장균 aFS-idi	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG	6.2	5.2
대장균 aFS-SN	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG, 5mM mevalonate	17.8	14.8
대장균ispAaFS-SN	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG, 5mM mevalonate	57.6	48.0
대장균ispANaFS-SN	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG, 5mM mevalonate	62.2	51.8

대장균ispALaFS-SN	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG, 5mM mevalonate	86.8	72.3
	2YT medium, 29℃, 1.0% (v/v) glycerol, 0.1mM IPTG, 15mM mevalonate	153.6	128.0
대장균ispALaFS-NA	2YT medium, 29℃, 3.0% (v/v) glycerol, 0.1mM IPTG	380.0	316.7

[0077] 참고

[0078] Amann, E., et al., 1988. Gene. 69, 301-15.

[0079] Berger, R. G., 2009. Biotechnol Lett. 31, 1651-9.

[0080] da Silva, G. P., et al., 2009. Biotechnol Adv. 27, 30-9.

[0081] Dhingra, V., et al., 2000. Life Sci. 66, 279-300.

[0082] Fujisaki, S., et al., 1990. J Biochem. 108, 995-1000.

[0083] Fukuda, H., et al., 2001. J Biosci Bioeng. 92, 405-16.

[0084] Green, S., et al., 2007. Phytochemistry. 68, 176-88.

[0085] Gustafsson, C., et al., 2004. Trends Biotechnol. 22, 346-53.

[0086] Huelin, F. E., Murray, K. E., 1966. Nature. 210, 1260-1.

[0087] Kane, J. F., 1995. Curr Opin Biotechnol. 6, 494-500.

[0088] Martin, V. J., et al., 2003. Nat Biotechnol. 21, 796-802.

[0089] Muntendam, R., et al., 2009. Appl Microbiol Biotechnol. 84, 1003-19.

[0090] Murarka, A., et al., 2008. Appl Environ Microbiol. 74, 1124-35.

[0091] Pare, P. W., Tumlinson, J. H., 1999. Plant Physiol. 121, 325-32.

[0092] Pechous, S. W., Whitaker, B. D., 2004. Planta. 219, 84-94.

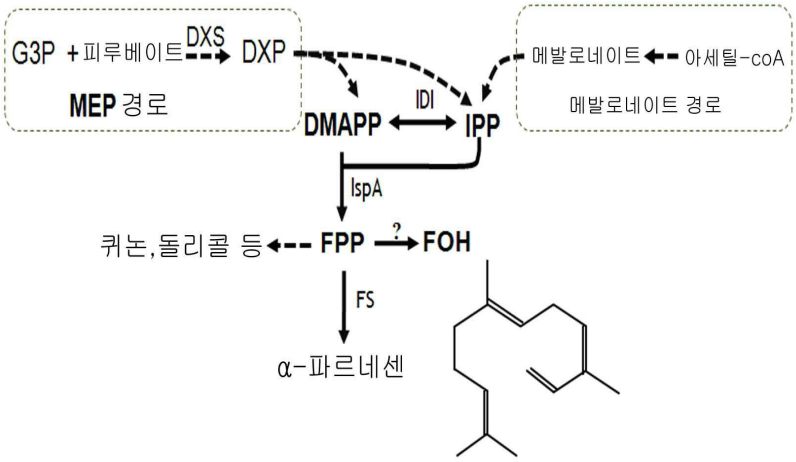
[0093] Renneger, N. S., McPhee, D. J., Vol. 7399323. AMYRIS BIOTECHNOLOGIES, INC US 2008.

[0094] Sambrook, J., Russell, D. Eds.), 2001. Cold Spring Harbor Laboratory Press, New York.

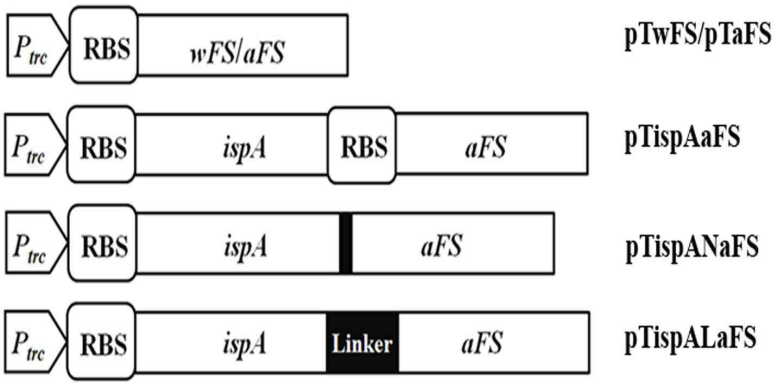
[0095] Yazdani, S. S., Gonzalez, R., 2007. Curr Opin Biotechnol. 18, 213-9.

도면

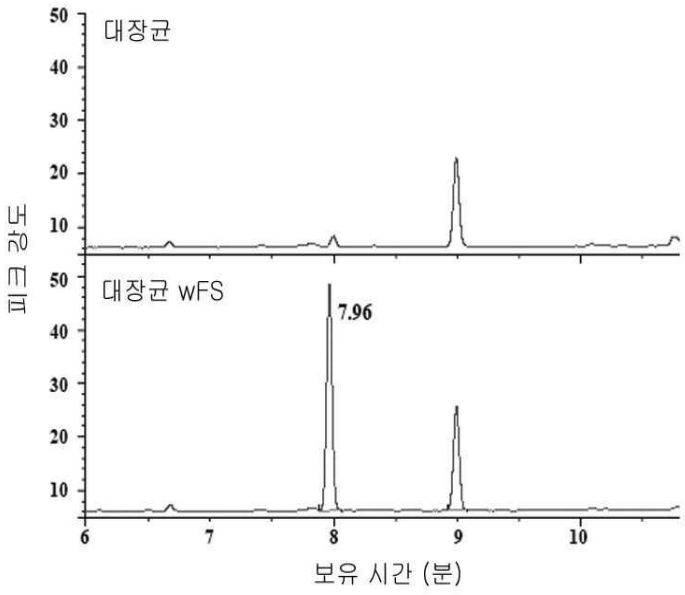
도면1a



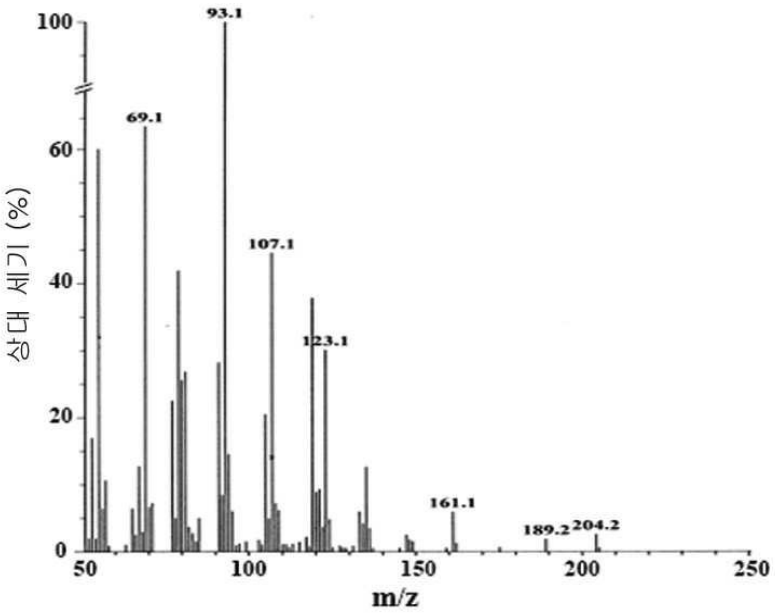
도면1b



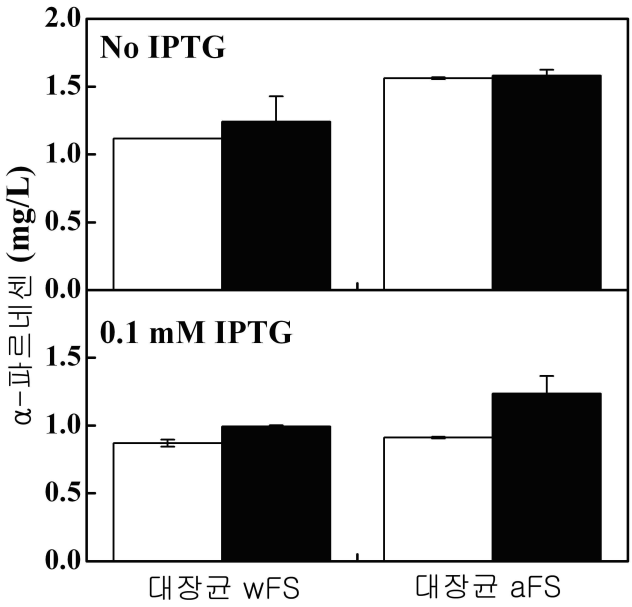
도면1c



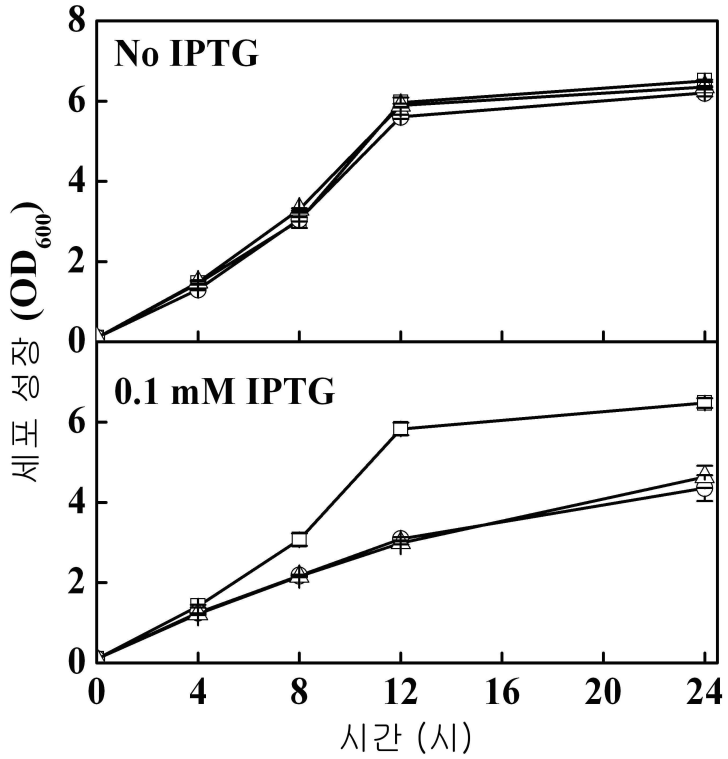
도면1d


100
60
40
20
0
상대 세기 (%)
93.1
69.1
107.1
123.1
161.1
189.2
204.2
50
100
150
200
250
m/z

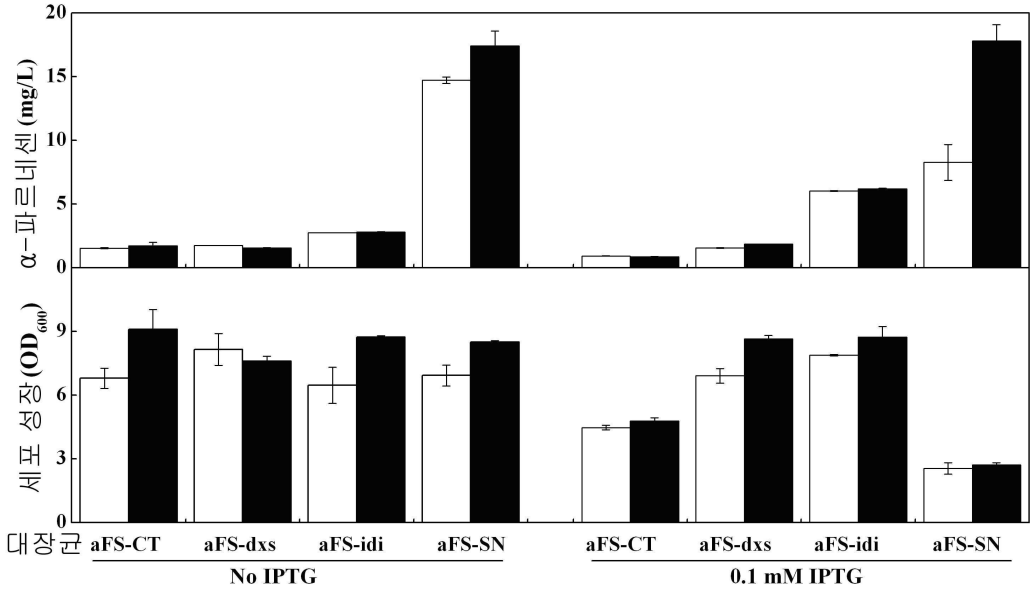
도면2a


No IPTG
0.1 mM IPTG
α-파르네센 (mg/L)
2.0
1.5
1.0
0.5
0.0
대장균 wFS
대장균 aFS

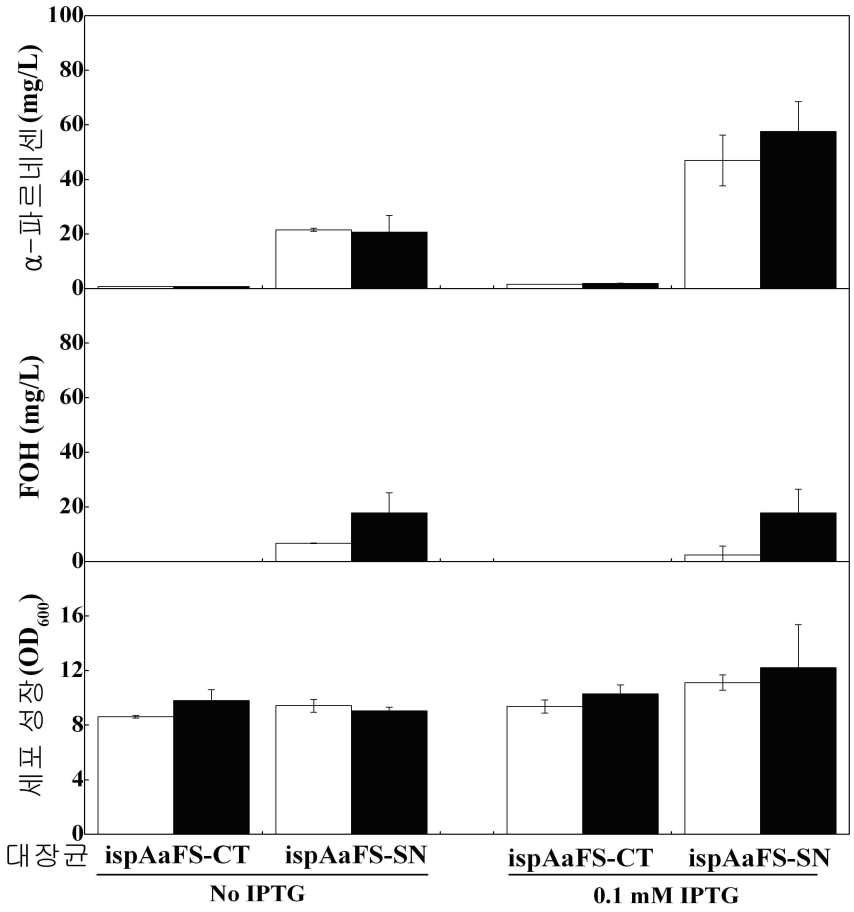
도면2b


No IPTG
0.1 mM IPTG
세포 성장 (OD600)
시간 (시)
0
2
4
6
8
0
4
8
12
16
20
24

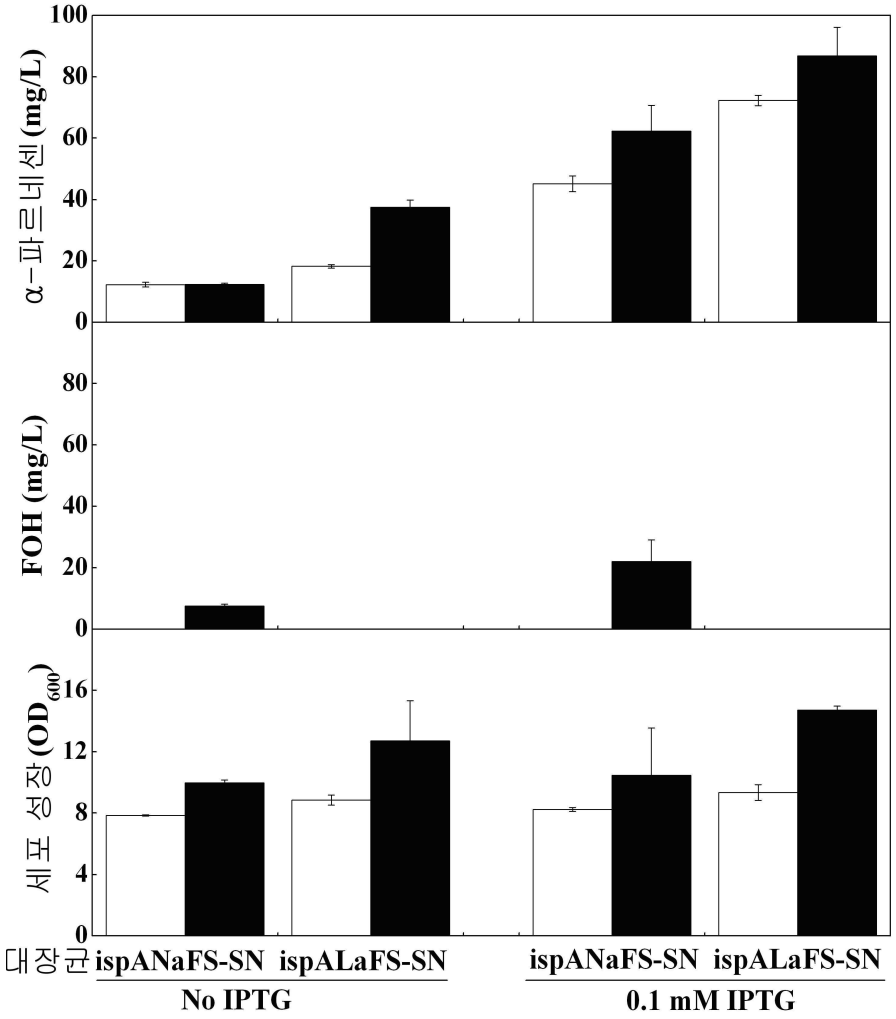
도면3


α-파르네센 (mg/L)
세포 성장 (OD600)
대장균
aFS-CT
aFS-dxs
aFS-idi
aFS-SN
No IPTG
0.1 mM IPTG

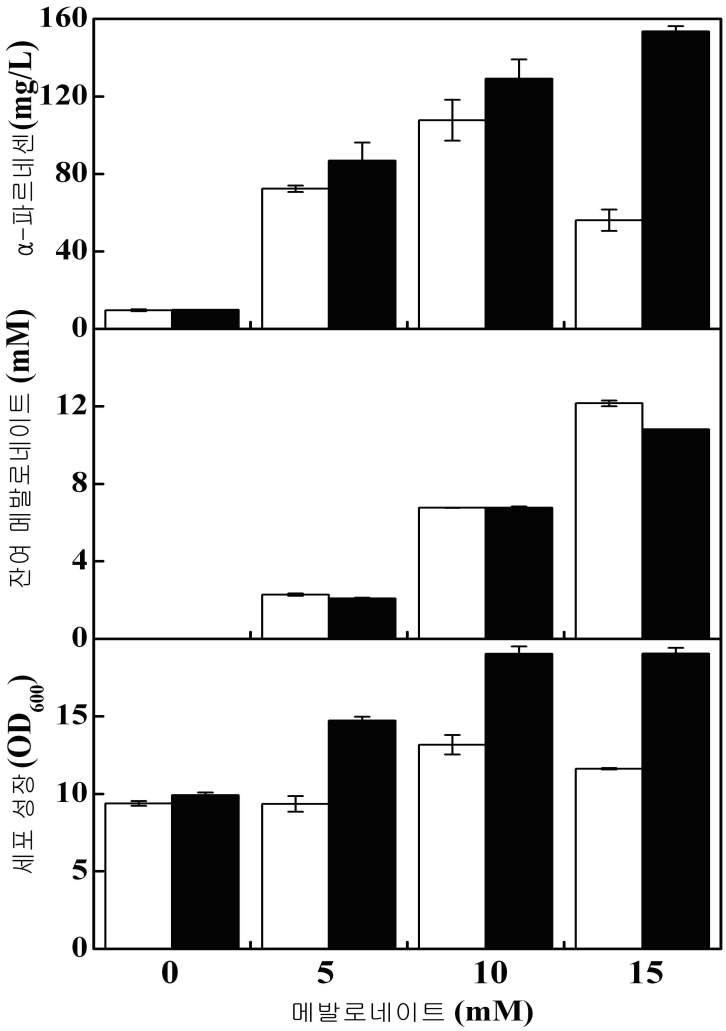
도면4


α-파르네센(mg/L)
100
80
60
40
20
0
FOH (mg/L)
80
60
40
20
0
세포 성장(OD600)
16
12
8
4
0
대장균
ispAaFS-CT
ispAaFS-SN
ispAaFS-CT
ispAaFS-SN
No IPTG
0.1 mM IPTG

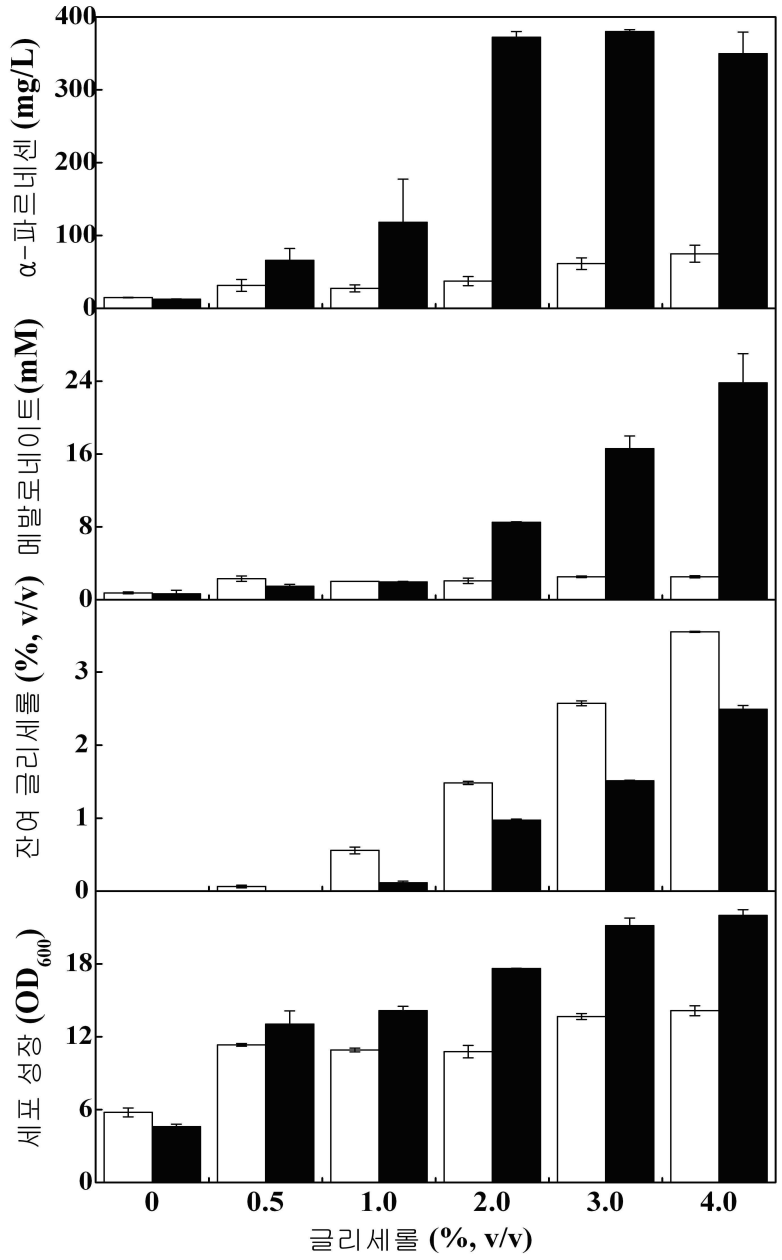
도면5


α-파르네센 (mg/L)
FOH (mg/L)
세포 성장 (OD600)
100
80
60
40
20
0
80
60
40
20
0
16
12
8
4
0
대장균
ispANaFS-SN
ispALaFS-SN
ispANaFS-SN
ispALaFS-SN
No IPTG
0.1 mM IPTG

도면6


160
120
80
40
0
α-파르네센(mg/L)
12
8
4
0
잔여 메발로네이트 (mM)
15
10
5
0
세포 성장 (OD600)
0
5
10
15
메발로네이트 (mM)

도면7



서 열 목 록

<110> INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYEONGSANG NATIONAL UNIVERSITY

<120> A microorganism of Escherichia genus having enhanced productivity of alpha-farnesene and method of producing alpha-farnesene using the same

<130> PN092240

<160> 40

<170> KopatentIn 1.71

<210> 1

<211> 1731

<212> DNA

<213> Artificial Sequence

<220><223> condon-optimized alpha-farnesene synthase

<400> 1

atggaatttc gtgtccacct gcaagcagac aatgaacaga aaattttcca aaaccaaatg 60

aaacctgaac cggaggctag ttacctgatt aaccagcgcc gttcagcgaa ttacaaaccg 120

aacatctgga aaaacgattt tctggatcag tccctgatct cgaaatatga tggtgacgaa 180

tatcgtaaac tgagcgaaaa actgatcgaa gaggtcaaaa tctatatttc ggctgaaact 240

atggacctgg tggcaaagct ggaactgatt gatagtgttc gtaaactggg gctggccaat 300

ctgttcgaaa aagaaatcaa agaagcactg gattcgattg ctgcaattga aagcgataac 360

ctgggtaccc gcgatgatct gtatggcacc gcgctgcact tcaaaatcct gcgccaacat 420

ggctataaag tgtcgcagga catttttgga cgttttatgg atgaaaaggg taccctggag 480

aatcatcatt ttgctcatct gaaaggaatg ctggaactgt ttgaagcctc gaacctgggt 540

tttgaaggcg aggacattct ggatgaagcc aaggcgtcgc tgacgctggc gctgcgcgat 600

agcggtcaca tttgctatcc tgattcgaat ctgagccgtg atgtcgtaca ctccctggag 660

ctgccgagcc atcgtcgcgt ccagtggttc gatgtgaaat ggcagatcaa cgcatatgaa 720

aaagatattt gtcgcgttaa cgcgacgctg ctggagctgg ccaaactgaa ttttaacgtg 780

gtgcaggcac agctgcagaa aaacctgcgc gaagcctcgc gttggtgggc caatctgggg 840

tttgcagata atctgaaatt tgcgcgtgat cgtctggtgg aatgtttttc atgcgccgtg 900

ggcgtggcct tcgaaccgga acatagttca tttcgtattt gtctgaccaa agtcattaac 960

ctggtcctga tcatcgatga tgtgtacgat atttacggtt cggaagaaga actgaaacat 1020

tttaccaacg cggtggatcg ttgggattct cgcgaaactg aacagctgcc ggaatgtatg 1080

aaaatgtgct tccaggttct gtataacaca acttgtgaaa ttgcccgtga aattgaggag 1140

gagaacggct ggaatcaggt gctgccgcag ctgacgaaag tttgggcgga cttttgtaaa 1200

gcgctgctgg tagaagcaga atggtataac aagagtcaca ttccgacact ggaggagtat 1260

ctgcgcaatg gctgcattag cagctccgtg agtgttctgc tggtgcatag ctttttctca 1320

atcacgcatg aagggacaaa agaaatggcg gattttctgc ataaaaacga agatctgctg 1380

tataatatta gcctgatcgt gcgtctgaac aacgacctgg ggaccagtgc ggcggagcag 1440

gaacgcggcg acagtcctag cagtattgtc tgttatatgc gtgaagtgaa tgcctccgaa 1500

gaaacggcgc gtaaaaatat caaaggtatg atcgacaatg cgtggaaaaa agtgaacggt 1560

aaatgtttca ccacgaacca ggttccgttc ctgtcctctt ttatgaataa cgcgaccaac 1620

atggcgcgcg tcgctcactc gctgtataaa gacggagatg gttttggaga tcaggaaaaa 1680

ggtccacgta cccacattct gtcgctgctg tttcagccgc tggtcaacta a 1731

<210> 2

<211> 299

<212> PRT

<213> E. coli

<400> 2

Met Asp Phe Pro Gln Gln Leu Glu Ala Cys Val Lys Gln Ala Asn Gln
1 5 10 15

Ala Leu Ser Arg Phe Ile Ala Pro Leu Pro Phe Gln Asn Thr Pro Val
20 25 30

Val Glu Thr Met Gln Tyr Gly Ala Leu Leu Gly Gly Lys Arg Leu Arg
35 40 45

Pro Phe Leu Val Tyr Ala Thr Gly His Met Phe Gly Val Ser Thr Asn

50 55 60

Thr Leu Asp Ala Pro Ala Ala Ala Val Glu Cys Ile His Ala Tyr Ser
65 70 75 80

Leu Ile His Asp Asp Leu Pro Ala Met Asp Asp Asp Asp Leu Arg Arg
85 90 95

Gly Leu Pro Thr Cys His Val Lys Phe Gly Glu Ala Asn Ala Ile Leu
100 105 110

Ala Gly Asp Ala Leu Gln Thr Leu Ala Phe Ser Ile Leu Ser Asp Ala
115 120 125

Asp Met Pro Glu Val Ser Asp Arg Asp Arg Ile Ser Met Ile Ser Glu
130 135 140

Leu Ala Ser Ala Ser Gly Ile Ala Gly Met Cys Gly Gly Gln Ala Leu
145 150 155 160

Asp Leu Asp Ala Glu Gly Lys His Val Pro Leu Asp Ala Leu Glu Arg
165 170 175

Ile His Arg His Lys Thr Gly Ala Leu Ile Arg Ala Ala Val Arg Leu
180 185 190

Gly Ala Leu Ser Ala Gly Asp Lys Gly Arg Arg Ala Leu Pro Val Leu

195 200 205

Asp Lys Tyr Ala Glu Ser Ile Gly Leu Ala Phe Gln Val Gln Asp Asp

210 215 220

Ile Leu Asp Val Val Gly Asp Thr Ala Thr Leu Gly Lys Arg Gln Gly

225 230 235 240

Ala Asp Gln Gln Leu Gly Lys Ser Thr Tyr Pro Ala Leu Leu Gly Leu

245 250 255

Glu Gln Ala Arg Lys Lys Ala Arg Asp Leu Ile Asp Asp Ala Arg Gln

260 265 270

Ser Leu Lys Gln Leu Ala Glu Gln Ser Leu Asp Thr Ser Ala Leu Glu

275 280 285

Ala Leu Ala Asp Tyr Ile Ile Gln Arg Asn Lys

290 295

<210> 3

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> linker

<400> 3

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 4

<211> 803

<212> PRT

<213> Enterococcus faecalis

<400> 4

Met Lys Thr Val Val Ile Ile Asp Ala Leu Arg Thr Pro Ile Gly Lys

1 5 10 15

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

20 25 30

Val Thr Thr Gln Leu Leu Lys Arg His Ser Thr Ile Ser Glu Glu Ile

35 40 45

Asp Gln Val Ile Phe Gly Asn Val Leu Gln Ala Gly Asn Gly Gln Asn

50 55 60

Pro Ala Arg Gln Ile Ala Ile Asn Ser Gly Leu Ser His Glu Ile Pro

65 70 75 80

Ala Met Thr Val Asn Glu Val Cys Gly Ser Gly Met Lys Ala Val Ile

85 90 95

Leu Ala Lys Gln Leu Ile Gln Leu Gly Glu Ala Glu Val Leu Ile Ala

100 105 110

Gly Gly Ile Glu Asn Met Ser Gln Ala Pro Lys Leu Gln Arg Phe Asn

115 120 125

Tyr Glu Thr Glu Ser Tyr Asp Ala Pro Phe Ser Ser Met Met Tyr Asp

130 135 140

Gly Leu Thr Asp Ala Phe Ser Gly Gln Ala Met Gly Leu Thr Ala Glu

145 150 155 160

Asn Val Ala Glu Lys Tyr His Val Thr Arg Glu Glu Gln Asp Gln Phe

165 170 175

Ser Val His Ser Gln Leu Lys Ala Ala Gln Ala Gln Ala Glu Gly Ile

180 185 190

Phe Ala Asp Glu Ile Ala Pro Leu Glu Val Ser Gly Thr Leu Val Glu

195 200 205

Lys Asp Glu Gly Ile Arg Pro Asn Ser Ser Val Glu Lys Leu Gly Thr

210 215 220

Leu Lys Thr Val Phe Lys Glu Asp Gly Thr Val Thr Ala Gly Asn Ala

225 230 235 240

Ser Thr Ile Asn Asp Gly Ala Ser Ala Leu Ile Ile Ala Ser Gln Glu

245 250 255

Tyr Ala Glu Ala His Gly Leu Pro Tyr Leu Ala Ile Ile Arg Asp Ser

260 265 270

Val Glu Val Gly Ile Asp Pro Ala Tyr Met Gly Ile Ser Pro Ile Lys

275 280 285

Ala Ile Gln Lys Leu Leu Ala Arg Asn Gln Leu Thr Thr Glu Glu Ile

290 295 300

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

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

435 440 445

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

Val Arg Glu Thr Glu Ile Phe Gln Gln Ala Glu Leu Ser Tyr Pro Ser
515 520 525
Ile Val Lys Arg Gly Gly Gly Leu Arg Asp Leu Gln Tyr Arg Ala Phe

530 535 540

Asp Glu Ser Phe Val Ser Val Asp Phe Leu Val Asp Val Lys Asp Ala

545 550 555 560

Met Gly Ala Asn Ile Val Asn Ala Met Leu Glu Gly Val Ala Glu Leu

565 570 575

Phe Arg Glu Trp Phe Ala Glu Gln Lys Ile Leu Phe Ser Ile Leu Ser

580 585 590

Asn Tyr Ala Thr Glu Ser Val Val Thr Met Lys Thr Ala Ile Pro Val

595 600 605

Ser Arg Leu Ser Lys Gly Ser Asn Gly Arg Glu Ile Ala Glu Lys Ile

610 615 620

Val Leu Ala Ser Arg Tyr Ala Ser Leu Asp Pro Tyr Arg Ala Val Thr

625 630 635 640

His Asn Lys Gly Ile Met Asn Gly Ile Glu Ala Val Val Leu Ala Thr

645 650 655

Gly Asn Asp Thr Arg Ala Val Ser Ala Ser Cys His Ala Phe Ala Val

660 665 670

Lys Glu Gly Arg Tyr Gln Gly Leu Thr Ser Trp Thr Leu Asp Gly Glu

675 680 685

Gln Leu Ile Gly Glu Ile Ser Val Pro Leu Ala Leu Ala Thr Val Gly

690 695 700

Gly Ala Thr Lys Val Leu Pro Lys Ser Gln Ala Ala Ala Asp Leu Leu

705 710 715 720

Ala Val Thr Asp Ala Lys Glu Leu Ser Arg Val Val Ala Ala Val Gly

725 730 735

Leu Ala Gln Asn Leu Ala Ala Leu Arg Ala Leu Val Ser Glu Gly Ile

740 745 750

Gln Lys Gly His Met Ala Leu Gln Ala Arg Ser Leu Ala Met Thr Val

755 760 765

Gly Ala Thr Gly Lys Glu Val Glu Ala Val Ala Gln Gln Leu Lys Arg

770 775 780

Gln Lys Thr Met Asn Gln Asp Arg Ala Leu Ala Ile Leu Asn Asp Leu

785 790 795 800

Arg Lys Gln

<210> 5

<211> 383

<212> PRT

<213> Enterococcus faecalis

<400> 5

Met Thr Ile Gly Ile Asp Lys Ile Ser Phe Phe Val Pro Pro Tyr Tyr
1 5 10 15

Ile Asp Met Thr Ala Leu Ala Glu Ala Arg Asn Val Asp Pro Gly Lys
20 25 30

Phe His Ile Gly Ile Gly Gln Asp Gln Met Ala Val Asn Pro Ile Ser
35 40 45

Gln Asp Ile Val Thr Phe Ala Ala Asn Ala Ala Glu Ala Ile Leu Thr

50 55 60

Lys Glu Asp Lys Glu Ala Ile Asp Met Val Ile Val Gly Thr Glu Ser
65 70 75 80

Ser Ile Asp Glu Ser Lys Ala Ala Ala Val Val Leu His Arg Leu Met
85 90 95

Gly Ile Gln Pro Phe Ala Arg Ser Phe Glu Ile Lys Glu Ala Cys Tyr
100 105 110

Gly Ala Thr Ala Gly Leu Gln Leu Ala Lys Asn His Val Ala Leu His
115 120 125

Pro Asp Lys Lys Val Leu Val Val Ala Ala Asp Ile Ala Lys Tyr Gly

130 135 140

Leu Asn Ser Gly Gly Glu Pro Thr Gln Gly Ala Gly Ala Val Ala Met
145 150 155 160

Leu Val Ala Ser Glu Pro Arg Ile Leu Ala Leu Lys Glu Asp Asn Val
165 170 175

Met Leu Thr Gln Asp Ile Tyr Asp Phe Trp Arg Pro Thr Gly His Pro
180 185 190

Tyr Pro Met Val Asp Gly Pro Leu Ser Asn Glu Thr Tyr Ile Gln Ser

195 200 205

Phe Ala Gln Val Trp Asp Glu His Lys Lys Arg Thr Gly Leu Asp Phe

210 215 220

Ala Asp Tyr Asp Ala Leu Ala Phe His Ile Pro Tyr Thr Lys Met Gly

225 230 235 240

Lys Lys Ala Leu Leu Ala Lys Ile Ser Asp Gln Thr Glu Ala Glu Gln

245 250 255

Glu Arg Ile Leu Ala Arg Tyr Glu Glu Ser Ile Ile Tyr Ser Arg Arg

260 265 270

Val Gly Asn Leu Tyr Thr Gly Ser Leu Tyr Leu Gly Leu Ile Ser Leu

275 280 285

Leu Glu Asn Ala Thr Thr Leu Thr Ala Gly Asn Gln Ile Gly Leu Phe

290 295 300

Ser Tyr Gly Ser Gly Ala Val Ala Glu Phe Phe Thr Gly Glu Leu Val

305 310 315 320

Ala Gly Tyr Gln Asn His Leu Gln Lys Glu Thr His Leu Ala Leu Leu

325 330 335

Asp Asn Arg Thr Glu Leu Ser Ile Ala Glu Tyr Glu Ala Met Phe Ala

340 345 350

Glu Thr Leu Asp Thr Asp Ile Asp Gln Thr Leu Glu Asp Glu Leu Lys

355 360 365

Tyr Ser Ile Ser Ala Ile Asn Asn Thr Val Arg Ser Tyr Arg Asn

370 375 380

<210> 6

<211> 292

<212> PRT

<213> Streptococcus pneumoniae

<400> 6

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

1 5 10 15

Ile Gly Glu His Ala Val Val Tyr Gly Tyr Pro Ala Ile Ser Leu Pro

20 25 30

Leu Leu Glu Val Glu Val Thr Cys Lys Val Val Ser Ala Glu Ser Pro

35 40 45

Trp Arg Leu Tyr Glu Glu Asp Thr Leu Ser Met Ala Val Tyr Ala Ser

50 55 60

Leu Glu Tyr Leu Asp Ile Thr Glu Ala Cys Val Arg Cys Glu Ile Asp

65 70 75 80

Ser Ala Ile Pro Glu Lys Arg Gly Met Gly Ser Ser Ala Ala Ile Ser

85 90 95

Ile Ala Ala Ile Arg Ala Val Phe Asp Tyr Tyr Gln Ala Asp Leu Pro

100 105 110

His Asp Val Leu Glu Ile Leu Val Asn Arg Ala Glu Met Ile Ala His

115 120 125

Met Asn Pro Ser Gly Leu Asp Ala Lys Thr Cys Leu Ser Asp Gln Pro

130 135 140

Ile Arg Phe Ile Lys Asn Val Gly Phe Thr Glu Leu Glu Met Asp Leu

145 150 155 160

Ser Ala Tyr Leu Val Ile Ala Asp Thr Gly Val Tyr Gly His Thr Arg

165 170 175

Glu Ala Ile Gln Val Val Gln Asn Lys Gly Lys Asp Ala Leu Pro Phe

180 185 190

Leu His Ala Leu Gly Glu Leu Thr Gln Gln Ala Glu Val Ala Ile Ser

195 200 205

Gln Lys Tyr Ala Glu Gly Leu Gly Leu Ile Phe Ser Gln Ala His Leu

210 215 220

His Leu Lys Glu Ile Gly Val Ser Ser Pro Glu Ala Asp Phe Leu Val

225 230 235 240

Glu Thr Ala Leu Ser Tyr Gly Ala Leu Gly Ala Lys Met Ser Gly Gly

245 250 255

Gly Leu Gly Gly Cys Ile Ile Ala Leu Val Thr Asn Leu Thr His Ala

260 265 270

Gln Glu Leu Ala Glu Arg Leu Glu Glu Lys Gly Ala Val Gln Thr Trp

275 280 285

Ile Glu Ser Leu

290

<210> 7

<211> 336

<212> PRT

<213> Streptococcus pneumoniae

<400> 7

Met Ile Ala Val Lys Thr Cys Gly Lys Leu Tyr Trp Ala Gly Glu Tyr

1 5 10 15

Ala Ile Leu Glu Pro Gly Gln Leu Ala Leu Ile Lys Asp Ile Pro Ile

20 25 30

Tyr Met Arg Ala Glu Ile Ala Phe Ser Asp Ser Tyr Arg Ile Tyr Ser

35 40 45

Asp Met Phe Asp Phe Ala Val Asp Leu Arg Pro Asn Pro Asp Tyr Ser

50 55 60

Leu Ile Gln Glu Thr Ile Ala Leu Met Gly Asp Phe Leu Ala Val Arg

65 70 75 80

Gly Gln Asn Leu Arg Pro Phe Ser Leu Lys Ile Cys Gly Lys Met Glu

85 90 95

Arg Glu Gly Lys Lys Phe Gly Leu Gly Ser Ser Gly Ser Val Val Val

100 105 110

Leu Val Val Lys Ala Leu Leu Ala Leu Tyr Asn Leu Ser Val Asp Gln

115 120 125

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

130 135 140

Asn Gly Ser Met Gly Asp Leu Ala Cys Ile Val Ala Glu Asp Leu Val

145 150 155 160

Leu Tyr Gln Ser Phe Asp Arg Gln Lys Ala Ala Ala Trp Leu Glu Glu

165 170 175

Glu Asn Leu Ala Thr Val Leu Glu Arg Asp Trp Gly Phe Phe Ile Ser
180 185 190

Gln Val Lys Pro Thr Leu Glu Cys Asp Phe Leu Val Gly Trp Thr Lys
195 200 205

Glu Val Ala Val Ser Ser His Met Val Gln Gln Ile Lys Gln Asn Ile
210 215 220

Asn Gln Asn Phe Leu Ser Ser Ser Lys Glu Thr Val Val Ser Leu Val
225 230 235 240

Glu Ala Leu Glu Gln Gly Lys Ala Glu Lys Val Ile Glu Gln Val Glu
245 250 255

Val Ala Ser Lys Leu Leu Glu Gly Leu Ser Thr Asp Ile Tyr Thr Pro
260 265 270

Leu Leu Arg Gln Leu Lys Glu Ala Ser Gln Asp Leu Gln Ala Val Ala
275 280 285

Lys Ser Ser Gly Ala Gly Gly Gly Asp Cys Gly Ile Ala Leu Ser Phe
290 295 300

Asp Ala Gln Ser Ser Arg Asn Thr Leu Lys Asn Arg Trp Ala Asp Leu
305 310 315 320

Gly Ile Glu Leu Leu Tyr Gln Glu Arg Ile Gly His Asp Asp Lys Ser
325 330 335

<210> 8

<211> 317

<212> PRT

<213> Streptococcus pneumoniae

<400> 8

Met Asp Arg Glu Pro Val Thr Val Arg Ser Tyr Ala Asn Ile Ala Ile
1 5 10 15

Ile Lys Tyr Trp Gly Lys Lys Lys Glu Lys Glu Met Val Pro Ala Thr
20 25 30

Ser Ser Ile Ser Leu Thr Leu Glu Asn Met Tyr Thr Glu Thr Thr Leu
35 40 45

Ser Pro Leu Pro Ala Asn Val Thr Ala Asp Glu Phe Tyr Ile Asn Gly

```
     50                  55                  60
Gln Leu Gln Asn Glu Val Glu His Ala Lys Met Ser Lys Ile Ile Asp
 65                  70                  75                  80
Arg Tyr Arg Pro Ala Gly Glu Gly Phe Val Arg Ile Asp Thr Gln Asn

                 85                  90                  95
Asn Met Pro Thr Ala Ala Gly Leu Ser Ser Ser Ser Ser Gly Leu Ser
            100                 105                 110
Ala Leu Val Lys Ala Cys Asn Ala Tyr Phe Lys Leu Gly Leu Asp Arg
        115                 120                 125
Ser Gln Leu Ala Gln Glu Ala Lys Phe Ala Ser Gly Ser Ser Ser Arg
    130                 135                 140
Ser Phe Tyr Gly Pro Leu Gly Ala Trp Asp Lys Asp Ser Gly Glu Ile
145                 150                 155                 160

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

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

Val Ser Lys Thr Lys Asp Leu Ser Gln Asp Asp Cys Cys

305 310 315

<210> 9

<211> 182

<212> PRT

<213> E. coli K12 MG1655

<400> 9

Met Gln Thr Glu His Val Ile Leu Leu Asn Ala Gln Gly Val Pro Thr

1 5 10 15

Gly Thr Leu Glu Lys Tyr Ala Ala His Thr Ala Asp Thr Arg Leu His

20 25 30

Leu Ala Phe Ser Ser Trp Leu Phe Asn Ala Lys Gly Gln Leu Leu Val

35 40 45

Thr Arg Arg Ala Leu Ser Lys Lys Ala Trp Pro Gly Val Trp Thr Asn

50 55 60

Ser Val Cys Gly His Pro Gln Leu Gly Glu Ser Asn Glu Asp Ala Val

65 70 75 80

Ile Arg Arg Cys Arg Tyr Glu Leu Gly Val Glu Ile Thr Pro Pro Glu

85 90 95

Ser Ile Tyr Pro Asp Phe Arg Tyr Arg Ala Thr Asp Pro Ser Gly Ile

100 105 110

Val Glu Asn Glu Val Cys Pro Val Phe Ala Ala Arg Thr Thr Ser Ala

115 120 125

Leu Gln Ile Asn Asp Asp Glu Val Met Asp Tyr Gln Trp Cys Asp Leu

130 135 140

Ala Asp Val Leu His Gly Ile Asp Ala Thr Pro Trp Ala Phe Ser Pro

145 150 155 160

Trp Met Val Met Gln Ala Thr Asn Arg Glu Ala Arg Lys Arg Leu Ser

165 170 175

Ala Phe Thr Gln Leu Lys

180

<210> 10

<211> 900

<212

> DNA

<213> E. coli

<400> 10

atggactttc cgcagcaact cgaagcctgc gttaagcagg ccaaccaggc gctgagccgt 60

tttatcgccc cactgccctt tcagaacact cccgtggtcg aaaccatgca gtatggcgca 120

ttattaggtg gtaagcgcct gcgacctttc ctggtttatg ccaccggtca tatgttcggc 180

gttagcacaa acacgctgga cgcacccgct gccgccgttg agtgtatcca cgcttactca 240

ttaattcatg atgatttacc ggcaatggat gatgacgatc tgcgtcgcgg tttgccaacc 300

tgccatgtga agtttggcga agcaaacgcg attctcgctg gcgacgcttt acaaacgctg 360

gcgttctcga ttttaagcga tgccgatatg ccggaagtgt cggaccgcga cagaatttcg 420

atgatttctg aactggcgag cgccagtggt attgccggaa tgtgcggtgg tcaggcatta 480

gatttagacg cggaaggcaa acacgtacct ctggacgcgc ttgagcgtat tcatcgtcat 540

aaaaccggcg cattgattcg cgccgccgtt cgccttggtg cattaagcgc cggagataaa 600

ggacgtcgtg ctctgccggt actcgacaag tatgcagaga gcatcggcct tgccttccag 660

gttcaggatg acatcctgga tgtggtggga gatactgcaa cgttgggaaa acgccagggt 720

gccgaccagc aacttggtaa aagtacctac cctgcacttc tgggtcttga gcaagcccgg 780

aagaaagccc gggatctgat cgacgatgcc cgtcagtcgc tgaaacaact ggctgaacag 840

tcactcgata cctcggcact ggaagcgcta gcggactaca tcatccagcg taataaataa 900

900

<210> 11

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> linker

<400> 11

ggtggcggtg gctcgggtgg cggcggatcc 30

<210> 12

<211> 2412

<212> DNA

<213> Enterococcus faecalis

<400> 12

ttgaaaacag tagttattat tgatgcatta cgaacaccaa ttggaaaata taaaggcagc 60

ttaagtcaag taagtgccgt agacttagga acacatgtta caacacaact tttaaaaaga 120

cattccacta tttctgaaga aattgatcaa gtaatctttg gaaatgtttt acaagctgga 180

aatggccaaa atcccgcacg acaaatagca ataaacagcg gtttatctca tgaaattccc 240

gcaatgacag ttaatgaggt ctgcggatca ggaatgaagg ccgttatttt ggcgaaacaa 300

ttgattcaat taggagaagc ggaagtttta attgctggcg ggattgagaa tatgtcccaa 360

gcacctaaat tacaacgatt taattacgaa acagaaagct atgatgcgcc tttttctagt 420

atgatgtacg atgggttaac ggatgccttt agtggtcaag caatgggctt aactgctgaa 480

aatgtggccg aaaagtatca tgtaactaga gaagagcaag atcaattttc tgtacattca 540

caattaaaag cagctcaagc acaagcagaa gggatattcg ctgacgaaat agccccatta 600

gaagtatcag gaacgcttgt ggagaaagat gaagggattc gccctaattc gagcgttgag 660

aagctaggaa cgcttaaaac agtttttaaa gaagacggta ctgtaacagc agggaatgca 720

tcaaccatta atgatggggc ttctgctttg attattgctt cacaagaata tgccgaagca 780

cacggtcttc cttatttagc tattattcga gacagtgtgg aagtcggtat tgatccagcc 840

tatatgggaa tttcgccgat taaagccatt caaaaactgt tagcgcgcaa tcaacttact 900

acggaagaaa ttgatctgta tgaaatcaac gaagcatttg cagcaacttc aatcgtggtc 960

caaagagaac tggctttacc agaggaaaag gtcaacattt atggtggcgg tatttcatta 1020

ggtcatgcga ttggtgccac aggtgctcgt ttattaacga gtttaagtta tcaattaaat 1080

caaaaagaaa agaaatatgg agtggcttct ttatgtatcg gcggtggctt aggactcgct 1140

atgctactag agagacctca gcaaaaaaaa aacagccgat tttatcaaat gagtcctgag 1200

gaacgcctgg cttctcttct taatgaaggc cagatttctg ctgatacaaa aaaagaattt 1260

gaaaatacgg ctttatcttc gcagattgcc aatcatatga ttgaaaatca aatcagtgaa 1320

acagaagtgc cgatgggcgt tggcttacat ttaacagtgg acgaaactga ttatttggta 1380

ccaatggcga cagaagagcc ctcagtgatt gcggctttga gtaatggtgc aaaaatagca 1440

caaggattta aaacagtgaa tcaacaacgt ttaatgcgtg gacaaatcgt tttttacgat 1500

gttgcagacg ccgagtcatt gattgatgaa ctacaagtaa gagaaacgga aatttttcaa 1560

caagcagagt taagttatcc atctatcgtt aaacgcggcg gcggcttaag agatttgcaa 1620

tatcgtgctt ttgatgaatc atttgtatct gtcgactttt tagtagatgt taaggatgca 1680

atgggggcaa atatcgttaa cgctatgttg gaaggtgtgg ccgagttgtt ccgtgaatgg 1740

tttgcggagc aaaagatttt attcagtatt ttaagtaatt atgccacgga gtcggttgtt 1800

acgatgaaaa cggctattcc agtttcacgt ttaagtaagg ggagcaatgg ccgggaaatt 1860

gctgaaaaaa ttgttttagc ttcacgctat gcttcattag atccttatcg ggcagtcacg 1920

cataacaaag ggatcatgaa tggcattgaa gctgtcgttt tagctacagg aaatgataca 1980

cgcgctgtta gcgcttcttg tcatgctttt gcggtgaagg aaggtcgcta ccaaggtttg 2040

actagttgga cgctggatgg cgaacaacta attggtgaaa tttcagttcc gcttgcgtta 2100

gccacggttg gcggtgccac aaaagtctta cctaaatctc aagcagctgc tgatttgtta 2160

gcagtgacgg atgcaaaaga actaagtcga gtagtagcgg ctgttggttt ggcacaaaat 2220

ttagcggcgt tacgggcctt agtctctgaa ggaattcaaa aaggacacat ggctctacaa 2280

gcacgttctt tagcgatgac ggtcggagct actggtaaag aagttgaggc agtcgctcaa 2340

caattaaaac gtcaaaaaac gatgaaccaa gaccgagcct tggctatttt aaatgattta 2400

agaaaacaat aa 2412

<210> 13

<211> 1152

<212> DNA

<213> Enterococcus faecalis

<400> 13

atgacaattg ggattgataa aattagtttt tttgtgcccc cttattatat tgatatgacg 60

gcactggctg aagccagaaa tgtagaccct ggaaaatttc atattggtat tgggcaagac 120

caaatggcgg tgaacccaat cagccaagat attgtgacat ttgcagccaa tgccgcagaa 180

gcgatcttga ccaaagaaga taaagaggcc attgatatgg tgattgtcgg gactgagtcc 240

agtatcgatg agtcaaaagc ggccgcagtt gtcttacatc gtttaatggg gattcaacct 300

ttcgctcgct ctttcgaaat caaggaagct tgttacggag caacagcagg cttacagtta 360

gctaagaatc acgtagcctt acatccagat aaaaaagtct tggttgtagc agcagatatt 420

gcaaaatatg gattaaattc tggcggtgag cctacacaag gagctggggc ggttgcaatg 480

ttagttgcta gtgaaccgcg catcttggct ttaaaagagg ataatgtgat gctgacgcaa 540

gatatctatg acttttggcg tccaacaggc catccgtatc ctatggtcga tggtcctttg 600

tcaaacgaaa cctacatcca atcttttgcc caagtctggg atgaacataa aaaaagaacc 660

ggtcttgatt ttgcagatta tgatgcttta gcgttccata ttccttacac aaaaatgggc 720

aaaaaagcct tattagcaaa aatctccgac caaactgaag cagaacagga acgaatttta 780

gcccgttatg aagaaagcat catctatagt cgtcgcgtag gaaacttgta tacgggttca 840

ctttatctgg gactcatttc ccttttagaa aatgcaacga ctttaaccgc aggcaatcaa 900

attgggttat tcagttatgg ttctggtgct gtcgctgaat ttttcactgg tgaattagta 960

gctggttatc aaaatcattt acaaaaagaa actcatttag cactgctaga taatcggaca 1020

gaactttcta tcgctgaata tgaagccatg tttgcagaaa ctttagacac agatattgat 1080

caaacgttag aagatgaatt aaaatatagt atttctgcta ttaataatac cgttcgctct 1140

tatcgaaact aa 1152

<210> 14

<211> 879

<212> DNA

<213> Streptococcus pneumoniae

<400> 14

atgacaaaaa aagttggtgt cggtcaggca catagtaaga taattttaat aggggaacat 60

gcggtcgttt acggttatcc tgccatttcc ctgcctcttt tggaggtgga ggtgacctgt 120

aaggtagttt ctgcagagag tccttggcgc ctttatgagg aggatacctt gtccatggcg 180

gtttatgcct cactggagta tttggatatc acagaagcct gcgttcgttg tgagattgac 240

tcggctatcc ctgagaaacg ggggatgggt tcgtcagcgg ctatcagcat agcggccatt 300

cgtgcggtat ttgactacta tcaggctgat ctgcctcatg atgtactaga aatcttggtc 360

aatcgagctg agatgattgc ccatatgaat cctagtggtt tggatgctaa gacctgtctc 420

agtgaccaac ctattcgctt tatcaagaac gtaggattta cagaacttga gatggattta 480

tccgcctatt tggtgattgc cgatacgggt gtttatggtc atactcgtga agccatccaa 540

gtggttcaaa ataagggcaa ggatgcccta ccgtttttgc atgccttggg agaattaacc 600

cagcaagcag aagttgcgat ttcacaaaaa tatgctgaag gactgggact aatcttcagt 660

caagctcatt tacatctaaa agaaattgga gtcagtagcc ctgaggcaga ctttttggtt 720

gaaacggctc ttagctatgg tgctctgggt gccaagatga gcggtggtgg gctaggaggt 780

tgtatcatag ccttggtaac caatttgacg cacgcacaag aactagcaga aagattagaa 840

gagaaaggag ctgttcagac atggatagag agcctgtaa 879

<210> 15

<211> 1011

<212> DNA

<213> Streptococcus pneumoniae

<400> 15

atgattgctg ttaaaacttg cggaaaactc tattgggcag gtgaatatgc tattttagag 60

ccagggcagt tagctttgat aaaggatatt cccatctata tgagggctga gattgctttt 120

tctgacagct accgtatcta ttcagatatg tttgatttcg cagtggactt aaggcccaat 180

cctgactaca gcttgattca agaaacgatt gctttgatgg gagacttcct cgctgttcgc 240

ggtcagaatt taagaccttt ttccctaaaa atctgtggca aaatggaacg agaagggaaa 300

aagtttggtc taggttctag tggcagcgtc gttgtcttgg ttgtcaaggc tttactggct 360

ctctataatc tttcggttga tcagaatctc ttgttcaagc tgactagcgc tgtcttgctc 420

aagcgaggag acaatggttc catgggcgac cttgcctgta ttgtggcaga ggatttggtt 480

ctttaccagt catttgatcg ccagaaggcg gctgcttggt tagaagaaga aaacttggcg 540

acagttctgg agcgtgattg gggatttttt atctcacaag tgaaaccaac tttagaatgt 600

gatttcttag tgggatggac caaggaagtg gctgtatcga gtcacatggt ccagcaaatc 660

aagcaaaata tcaatcaaaa ttttttaagt tcctcaaaag aaacggtggt ttctttggtc 720

gaagccttgg agcaggggaa agccgaaaaa gttatcgagc aagtagaagt agccagcaag 780

cttttagaag gcttgagtac agatatttac acgcctttgc ttagacagtt gaaagaagcc 840

agtcaagatt tgcaggccgt tgccaagagt agtggtgctg gtggtggtga ctgtggcatc 900

gccctgagtt ttgatgcgca atcttctcga aacactttaa aaaatcgttg ggccgatctg 960

gggattgagc tcttatatca agaaaggata ggacatgacg acaaatcgta a 1011

<210> 16

<211> 954

<212> DNA

<213> Streptococcus pneumoniae

<400> 16

atggatagag agcctgtaac agtacgttcc tacgcaaata ttgctattat caaatattgg 60

ggaaagaaaa aagaaaaaga gatggtgcct gctactagca gtatttctct aactttggaa 120

aatatgtata cagagacgac cttgtcgcct ttaccagcca atgtaacagc tgacgaattt 180

tacatcaatg gtcagctaca aaatgaggtc gagcatgcca agatgagtaa gattattgac 240

cgttatcgtc cagctggtga gggctttgtc cgtatcgata ctcaaaacaa tatgcctacg 300

gcagcgggcc tgtcctcaag ttctagtggt ttgtccgccc tggtcaaggc ttgtaatgct 360

tatttcaagc ttggattgga tagaagtcag ttggcacagg aagccaaatt tgcctcaggc 420

tcttcttctc ggagttttta tggaccacta ggagcctggg ataaggatag tggagaaatt 480

taccctgtag agacagactt gaaactagct atgattatgt tggtgctaga ggacaagaaa 540

aaaccaatct ctagccgtga cgggatgaaa ctttgtgtgg aaacctcgac gacttttgac 600

gactgggttc gtcagtctga gaaggactat caggatatgc tgatttatct caaggaaaat 660

gattttgcca agattggaga attaacggag aaaaatgctc tggctatgca tgctacgaca 720

aagactgcta gtccagcctt ttcttatctg acggatgcct cttatgaggc tatggccttt 780

gttcgccagc ttcgtgagaa aggagaggcc tgctacttta ccatggatgc tggtcccaat 840

gttaaggtct tctgtcagga gaaagacttg gagcatttgt cagaaatttt cggtcagcgt 900

tatcgcttga ttgtgtcaaa aacaaaggat ttgagtcaag atgattgctg ttaa 954

<210> 17

<211> 546

<212> DNA

<213> E. coli K12 MG1655

<400> 17

atgcaaacgg aacacgtcat tttattgaat gcacagggag ttcccacggg tacgctggaa 60

aagtatgccg cacacacggc agacacccgc ttacatctcg cgttctccag ttggctgttt 120

aatgccaaag gacaattatt agttacccgc cgcgcactga gcaaaaaagc atggcctggc 180

gtgtggacta actcggtttg tgggcaccca caactgggag aaagcaacga agacgcagtg 240

atccgccgtt gccgttatga gcttggcgtg gaaattacgc ctcctgaatc tatctatcct 300

gactttcgct accgcgccac cgatccgagt ggcattgtgg aaaatgaagt gtgtccggta 360

tttgccgcac gcaccactag tgcgttacag atcaatgatg atgaagtgat ggattatcaa 420

tggtgtgatt tagcagatgt attacacggt attgatgcca cgccgtgggc gttcagtccg 480

tggatggtga tgcaggcgac aaatcgcgaa gccagaaaac gattatctgc atttacccag 540

cttaaa 546

<210> 18

<211> 2999

<212> DNA

<213> Artificial Sequence

<220><223> pSTV vector sequence

<400> 18

cgtatggcaa tgaaagacgg tgagctggtg atatgggata gtgttcaccc ttgttacacc 60

gttttccatg agcaaactga aacgttttca tcgctctgga gtgaatacca cgacgatttc 120

cggcagtttc tacacatata ttcgcaagat gtggcgtgtt acggtgaaaa cctggcctat 180

ttccctaaag ggtttattga gaatatgttt ttcgtctcag ccaatccctg ggtgagtttc 240

accagttttg atttaaacgt ggccaatatg gacaacttct tcgcccccgt tttcaccatg 300

ggcaaatatt atacgcaagg cgacaaggtg ctgatgccgc tggcgattca ggttcatcat 360

gccgtctgtg atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgat 420

gagtggcagg gcggggcgta atttttttaa ggcagttatt ggtgccctta aacgcctggt 480

gctacgcctg aataagtgat aataagcgga tgaatggcag aaattcgaaa gcaaattcga 540

cccggtcgtc ggttcagggc agggtcgtta aatagccgct tatgtctatt gctggtttac 600

cggtttattg actaccggaa gcagtgtgac cgtgtgcttc tcaaatgcct gaggccagtt 660

tgctcaggct ctccccgtgg aggtaataat tgacgatatg atcatttatt ctgcctccca 720

gagcctgata aaaacggtta gcgcttcgtt aatacagatg taggtgttcc acagggtagc 780

cagcagcatc ctgcgatgca gatccggaac ataatggtgc agggcgcttg tttcggcgtg 840

ggtatggtgg caggccccgt ggccgggggа ctgttgggcg ctgccggcac ctgtcctacg 900

agttgcatga taaagaagac agtcataagt gcggcgacga tagtcatgcc ccgcgcccac 960

cggaaggagc taccggacag cggtgcggac tgttgtaact cagaataaga aatgaggccg 1020

ctcatggcgt tccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca 1080

gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga 1140

gttagctcac tcattaggca ccccaggctt tacactttat gcttccggct cgtatgttgt 1200

gtggaattgt gagcggataa caatttcaca caggaaacag ctatgaccat gattacgaat 1260

tcgagctcgg tacccgggga tcctctagag tcgacctgca ggcatgcaag cttggcactg 1320

gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact taatcgcctt 1380

gcagcacatc ccccttcgc cagctggcgt aatagcgaag aggcccgcac cgatcgccct 1440

tcccaacagt tgcgcagcct gaatggcgaa tgagcttatc gatgataagc tgtcaaacat 1500

gagaattaca acttatatcg tatggggctg acttcaggtg ctacatttga agagataaat 1560

tgcactgaaa tctagaaata ttttatctga ttaataagat gatcttcttg agatcgtttt 1620

ggtctgcgcg taatctcttg ctctgaaaac gaaaaaaccg ccttgcaggg cggtttttcg 1680

aaggttctct gagctaccaa ctctttgaac cgaggtaact ggcttggagg agcgcagtca 1740

ccaaaacttg tcctttcagt ttagccttaa ccggcgcatg acttcaagac taactcctct 1800

aaatcaatta ccagtggctg ctgccagtgg tgcttttgca tgtctttccg ggttggactc 1860

aagacgatag ttaccggata aggcgcagcg gtcggactga acgggggggtt cgtgcataca 1920

gtccagcttg gagcgaactg cctacccgga actgagtgtc aggcgtggaa tgagacaaac 1980

gcggccataa cagcggaatg acaccggtaa accgaaaggc aggaacagga gagcgcacga 2040

gggagccgcc aggggaaacg cctggtatct ttatagtcct gtcgggtttc gccaccactg 2100

atttgagcgt cagatttcgt gatgcttgtc agggggggcgg agcctatgga aaaacggctt 2160

tgccgcggcc ctctcacttc cctgttaagt atcttcctgg catcttccag gaaatctccg 2220

ccccgttcgt aagccatttc cgctcgccgc agtcgaacga ccgagcgtag cgagtcagtg 2280

agcgaggaag cggaatatat cctgtatcac atattctgct gacgcaccgg tgcagccttt 2340

tttctcctgc cacatgaagc acttcactga caccctcatc agtgccaaca tagtaagcca 2400

gtatacactc cgctagcgct gatgtccggc ggtgcttttg ccgttacgca ccaccccgtc 2460

agtagctgaa caggagggac agctgataga aacagaagcc actggagcac ctcaaaaaca 2520

ccatcataca ctaaatcagt aagttggcag catcacccga cgcactttgc gccgaataaa 2580

tacctgtgac ggaagatcac ttcgcagaat aaataaatcc tggtgtccct gttgataccg 2640

ggaagccctg ggccaacttt tggcgaaaat gagacgttga tcggcacgta agaggttcca 2700

actttcacca taatgaaata agatcactac cgggcgtatt ttttgagtta tcgagatttt 2760

caggagctaa ggaagctaaa atggagaaaa aaatcactgg atataccacc gttgatatat 2820

cccaatggca tcgtaaagaa cattttgagg catttcagtc agttgctcaa tgtacctata 2880

accagaccgt tcagctggat attacggcct ttttaaagac cgtaaagaaa aataagcaca 2940

agttttatcc ggcctttatt cacattcttg cccgcctgat gaatgctcat ccggaattt 2999

<210> 19

<211> 4176

<212> DNA

<213> Artificial Sequence

<220><223> pTrc99A vector

<400> 19

gtttgacagc ttatcatcga ctgcacggtg caccaatgct tctggcgtca ggcagccatc 60

ggaagctgtg gtatggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc 120

gcactcccgt tctggataat gtttttgcg ccgacatcat aacggttctg gcaaatattc 180

tgaaatgagc tgttgacaat taatcatccg gctcgtataa tgtgtggaat tgtgagcgga 240

taacaatttc acacaggaaa cagaccatgg aattcgagct cggtacccgg ggatcctcta 300

gagtcgacct gcaggcatgc aagcttggct gttttggcgg atgagagaag attttcagcc 360

tgatacagat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca 420

gtagcgcggt ggtcccacct gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg 480

atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaaacga 540

aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggt gaacgctctc 600

ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccggaggg 660

tggcgggcag gacgcccgcc ataaactgcc aggcatcaaa ttaagcagaa ggccatcctg 720

acggatggcc tttttgcgtt tctacaaact ctttttgttt atttttctaa atacattcaa 780

atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 840

agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg gcattttgcc 900

ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg 960

gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc 1020

gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat 1080

tatcccgtgt tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg 1140

acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag 1200

aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa 1260

cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc 1320

gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca 1380

cgatgcctac agcaatggca acaacgttgc gcaaactatt aactggcgaa ctacttactc 1440

tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc 1500

tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg 1560

ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta 1620

tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag 1680

gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga 1740

ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc 1800

tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 1860

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 1920

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttc 1980

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt 2040

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 2100

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 2160

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 2220

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 2280

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 2340

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 2400

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat 2460

ggaaaaacgc cagcaacgcg gccttttttac ggttcctggc cttttgctgg ccttttgctc 2520

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt 2580

gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag 2640

cggaagagcg cctgatgcgg tattttctcc ttacgcatct gtgcggtatt tcacaccgca 2700
tatggtgcac tctcagtaca atctgctctg atgccgcata gttaagccag tatacactcc 2760
gctatcgcta cgtgactggg tcatggctgc gccccgacac ccgccaacac ccgctgacgc 2820
gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg 2880
gagctgcatg tgtcagaggt tttcaccgtc atcaccgaaa cgcgcgaggc agcagatcaa 2940
ttcgcgcgcg aaggcgaagc ggcatgcatt tacgttgaca ccatcgaatg gtgcaaaacc 3000
tttcgcggta tggcatgata gcgcccggaa gagagtcaat tcagggtggt gaatgtgaaa 3060

ccagtaacgt tatacgatgt cgcagagtat gccggtgtct cttatcagac cgtttcccgc 3120
gtggtgaacc aggccagcca cgtttctgcg aaaacgcggg aaaaagtgga agcggcgatg 3180
gcggagctga attacattcc caaccgcgtg gcacaacaac tggcgggcaa acagtcgttg 3240
ctgattggcg ttgccacctc cagtctggcc ctgcacgcgc cgtcgcaaat tgtcgcggcg 3300
attaaatctc gcgccgatca actgggtgcc agcgtggtgg tgtcgatggt agaacgaagc 3360
ggcgtcgaag cctgtaaagc ggcggtgcac aatcttctcg cgcaacgcgt cagtgggctg 3420
atcattaact atccgctgga tgaccaggat gccattgctg tggaagctgc ctgcactaat 3480

gttccggcgt tatttcttga tgtctctgac cagacaccca tcaacagtat tattttctcc 3540
catgaagacg gtacgcgact gggcgtggag catctggtcg cattgggtca ccagcaaatc 3600
gcgctgttag cgggcccatt aagttctgtc tcggcgcgtc tgcgtctggc tggctggcat 3660
aaatatctca ctcgcaatca aattcagccg atagcggaac gggaaggcga ctggagtgcc 3720
atgtccggtt ttcaacaaac catgcaaatg ctgaatgagg gcatcgttcc cactgcgatg 3780
ctggttgcca acgatcagat ggcgctgggc gcaatgcgcg ccattaccga gtccgggctg 3840
cgcgttggtg cggatatctc ggtagtggga tacgacgata ccgaagacag ctcatgttat 3900

atcccgccgt caaccaccat caaacaggat tttcgcctgc tggggcaaac cagcgtggac 3960
cgcttgctgc aactctctca gggccaggcg gtgaagggca atcagctgtt gcccgtctca 4020
ctggtgaaaa gaaaaaccac cctggcgccc aatacgcaaa ccgcctctcc ccgcgcgttg 4080
gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg 4140
caacgcaatt aatgtgagtt agcgcgaatt gatctg 4176

<210> 20
<211> 44
<212> DNA
<213> Artificial Sequence

<220><223> Forward primer

<400> 20

cgagctcagg agcatttaga tgttgaaaac agtagttatt attg 44

<210> 21

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer

<400> 21

gcccggggtg gcctgaaacg gctacc 26

<210> 22

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> Forward primer

<400> 22

gcccgggagg agttaaagaa atgacaattg 30

<210> 23

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer

<400> 23

cggatccctt agtttcgata agagcgaac 29

<210> 24

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> Forward primer

<400> 24

cggtaccaat gacaaaaaaa gttggtgtcg g 31

<210> 25

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer

<400> 25

ttctagatta cgatttgtcg tcatgtcc 28

<210> 26

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> Forward primer

<400> 26

ccccgggagg agagaaatta tgcaaacgga ac 32

<210> 27

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer

<400> 27

tgcatgctta tttaagctgg gtaaatg 27

<210> 28

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> aFS-F1

<400> 28

gggatccatg gaatttcgtg tccacctg 28

<210> 29

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> aFS-F2

<400> 29

gggatccagg aggtaataat aatggaattt cgtgtccacc tg 42

<210> 30

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> aFS-R

<400> 30

cccaagctta gttgaccagc ggctgaaaca g 31

<210> 31

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> ispA-F

<400> 31

gcgaattcat ggactttccg cagcaac 27

<210> 32

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> ispA-R1

<400> 32

gggatcctta tttattacgc tggatgtag 29

<210> 33

<211> 29

<212> DNA

<213> Artificial Sequence

<220><223> ispA-R2

<400> 33

gcggatcctt tattacgctg gatgtagtc 29

<210> 34

<211> 53

<212> DNA

<213> Artificial Sequence

<220><223> ispA-R3

<400> 34

gcggatccgc cgccacccga gccaccgcca cctttattac gctggatgta gtc 53

<210> 35

<211> 1731

<212> DNA

<213> Malus domestica

<400> 35

atggaattca gagttcactt gcaagctgat aatgagcaga aaatttttca aaaccagatg 60

aaacccgaac ctgaagcctc ttacttgatt aatcaaagac ggtctgcaaa ttacaagcca 120

aatatttgga agaacgattt cctagatcaa tctcttatca gcaaatacga tggagatgag 180

tatcggaagc tgtctgagaa gttaatagaa gaagttaaga tttatatatc tgctgaaaca 240

atggatttag tagctaagtt ggagctcatt gacagcgtcc gaaaactagg cctcgcgaac 300

ctcttcgaaa aggaaatcaa ggaagcccta gacagcattg cagctatcga aagcgacaat 360

ctcggcacaa gagacgatct ctatggtact gcattacact tcaagatcct caggcagcat 420

ggctataaag tttcacaaga tatatttggt agattcatgg atgaaaaggg cacattagag 480

aaccaccatt tcgcgcattt aaaaggaatg ctggaacttt tcgaggcctc aaacctgggt 540

ttcgaaggtg aagatatttt agatgaggcg aaagcttcct tgacgctagc tctcagagat 600

agtggtcata tttgttatcc agacagtaac ctttccaggg acgtagttca ttccctggag 660

cttccatcac accgcagagt gcagtggttt gatgtcaaat ggcaaatcaa cgcctatgaa 720

aaagacattt gtcgcgtcaa cgccacgtta ctcgaattag caaagcttaa tttcaacgta 780

gttcaggccc aactccaaaa aaacttaagg gaagcatcca ggtggtgggc aaatctgggc 840

ttcgcagaca acttgaaatt tgcaagagat agactggttg aatgtttctc atgtgctgtg 900

ggagtagcat tcgagcctga gcactcatct tttagaatat gtcttaccaa agtcatcaac 960

ttagtactga tcatagacga cgtctatgat atttatggtt cagaggaaga gctaaagcac 1020

ttcaccaatg ctgttgatag gtgggattct agggaaactg agcagcttcc agagtgtatg 1080

aagatgtgtt tccaagtact ctacaacact acttgtgaaa ttgctcgtga aattgaggag 1140

gagaatggtt ggaaccaagt attacctcaa ttgaccaaag tgtgggcaga tttttgtaaa 1200

gcattattgg tggaggcaga gtggtataat aagagccata taccaaccct tgaagagtac 1260

ctaagaaacg gatgcatttc atcatcagtt tcagtgcttt tggttcactc gtttttctct 1320

ataactcatg agggaaccaa agagatggct ggttttcttc acaagaatga agatcttttg 1380

tataatatct ctctcatcgt tcgcctcaac aatgatttgg gaacttccgc ggctgaacaa 1440

gagagagggg attctccttc atcaatcgta tgttacatga gagaagtgaa tgcctctgaa 1500

gaaacagcta ggaagaacat taagggcatg atagacaatg catggaagaa agtaaatgga 1560

aaatgcttca caacaaacca agtgcctttt ctgtcatcat tcatgaacaa tgccacaaac 1620

atggcacgtg tggcgcacag cctttacaaa gatggagatg ggtttggtga ccaagagaaa 1680

gggcctcgga cccacatcct gtctttacta ttccaacctc ttgtaaacta g 1731

<210> 36

<211> 2658

<212> DNA

<213> Artificial Sequence

<220><223> fusion ispALaFS DNA sequence

<400> 36

atggactttc cgcagcaact cgaagcctgc gttaagcagg ccaaccaggc gctgagccgt 60

tttatcgccc cactgccctt tcagaacact cccgtggtcg aaaccatgca gtatggcgca 120

ttattaggtg gtaagcgcct gcgacctttc ctggtttatg ccaccggtca tatgttcggc 180

gttagcacaa acacgctgga cgcacccgct gccgccgttg agtgtatcca cgcttactca 240

ttaattcatg atgatttacc ggcaatggat gatgacgatc tgcgtcgcgg tttgccaacc 300

tgccatgtga agtttggcga agcaaacgcg attctcgctg gcgacgcttt acaaacgctg 360

gcgttctcga ttttaagcga tgccgatatg ccggaagtgt cggaccgcga cagaatttcg 420

atgatttctg aactggcgag cgccagtggt attgccggaa tgtgcggtgg tcaggcatta 480

gatttagacg cggaaggcaa acacgtacct ctggacgcgc ttgagcgtat tcatcgtcat 540

aaaaccggcg cattgattcg cgccgccgtt cgccttggtg cattaagcgc cggagataaa 600

ggacgtcgtg ctctgccggt actcgacaag tatgcagaga gcatcggcct tgccttccag 660

gttcaggatg acatcctgga tgtggtggga gatactgcaa cgttgggaaa acgccagggt 720

gccgaccagc aacttggtaa aagtacctac cctgcacttc tgggtcttga gcaagcccgg 780

aagaaagccc gggatctgat cgacgatgcc cgtcagtcgc tgaaacaact ggctgaacag 840

tcactcgata cctcggcact ggaagcgcta gcggactaca tcatccagcg taataaaggt 900

ggcggtggct cgggtggcgg cggatccatg gaatttcgtg tccacctgca agcagacaat 960

gaacagaaaa ttttccaaaa ccaaatgaaa cctgaaccgg aggctagtta cctgattaac 1020

cagcgccgtt cagcgaatta caaaccgaac atctggaaaa acgattttct ggatcagtcc 1080

ctgatctcga aatatgatgg tgacgaatat cgtaaactga gcgaaaaact gatcgaagag 1140

gtcaaaatct atatttcggc tgaaactatg gacctggtgg caaagctgga actgattgat 1200

agtgttcgta aactggggct ggccaatctg ttcgaaaaag aaatcaaaga agcactggat 1260

tcgattgctg caattgaaag cgataacctg ggtacccgcg atgatctgta tggcaccgcg 1320

ctgcacttca aaatcctgcg ccaacatggc tataaagtgt cgcaggacat ttttggacgt 1380

tttatggatg aaaagggtac cctggagaat catcattttg ctcatctgaa aggaatgctg 1440

gaactgtttg aagcctcgaa cctgggtttt gaaggcgagg acattctgga tgaagccaag 1500

gcgtcgctga cgctggcgct gcgcgatagc ggtcacattt gctatcctga ttcgaatctg 1560

agccgtgatg tcgtacactc cctggagctg ccgagccatc gtcgcgtcca gtggttcgat 1620

gtgaaatggc agatcaacgc atatgaaaaa gatatttgtc gcgttaacgc gacgctgctg 1680

gagctggcca aactgaattt taacgtggtg caggcacagc tgcagaaaaa cctgcgcgaa 1740

gcctcgcgtt ggtgggccaa tctggggttt gcagataatc tgaaatttgc gcgtgatcgt 1800

ctggtggaat gtttttcatg cgccgtgggc gtggccttcg aaccggaaca tagttcattt 1860

cgtatttgtc tgaccaaagt cattaacctg gtcctgatca tcgatgatgt gtacgatatt 1920

tacggttcgg aagaagaact gaaacatttt accaacgcgg tggatcgttg ggattctcgc 1980

gaaactgaac agctgccgga atgtatgaaa atgtgcttcc aggttctgta taacacaact 2040

tgtgaaattg cccgtgaaat tgaggaggag aacggctgga atcaggtgct gccgcagctg 2100

acgaaagttt gggcggactt ttgtaaagcg ctgctggtag aagcagaatg gtataacaag 2160

agtcacattc cgacactgga ggagtatctg cgcaatggct gcattagcag ctccgtgagt 2220

gttctgctgg tgcatagctt tttctcaatc acgcatgaag ggacaaaaga aatggcggat 2280

tttctgcata aaaacgaaga tctgctgtat aatattagcc tgatcgtgcg tctgaacaac 2340

gacctgggga ccagtgcggc ggagcaggaa cgcggcgaca gtcctagcag tattgtctgt 2400

tatatgcgtg aagtgaatgc ctccgaagaa acggcgcgta aaaatatcaa aggtatgatc 2460

gacaatgcgt ggaaaaaagt gaacggtaaa tgtttcacca cgaaccaggt tccgttcctg 2520

tcctctttta tgaataacgc gaccaacatg gcgcgcgtcg ctcactcgct gtataaagac 2580

ggagatggtt ttggagatca ggaaaaaggt ccacgtaccc acattctgtc gctgctgttt 2640

cagccgctgg tcaactaa 2658

<210> 37

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> dxs-F

<400> 37

cgaattcagg aggaaaatta tgagttttga tattgcc 37

<210> 38

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> dxs-R

<400> 38

tggatcctta tgccagccag gccttg 26

<210> 39

<211> 37

<212> DNA

<213> Artificial Sequence

<220><223> idi-F

<400> 39

cggatccagg agagaaatta tgcaaacgga acacgtc 37

<210> 40

<211> 25

<212> DNA

<213> Artificial Sequence

<220><223> idi-R

<400> 40

ctgtcgactt aagctgggta aatgc 25