



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0050178
(43) 공개일자 2013년05월15일

(51) 국제특허분류(Int. Cl.)

C12N 9/00 (2006.01) *C12N 15/52* (2006.01)

C12N 15/70 (2006.01) *C07K 19/00* (2006.01)

(21) 출원번호 10-2011-0115390

(22) 출원일자 2011년11월07일

심사청구일자 2011년11월07일

(71) 출원인

경상대학교산학협력단

경상남도 진주시 진주대로 501 (가좌동)

(72) 발명자

김선원

경상남도 사천시 사남면 월성리 푸르지오아파트
102-102

왕충룡

경상남도 진주시 가좌동 900번지

(뒷면에 계속)

(74) 대리인

리엔텍특허법인

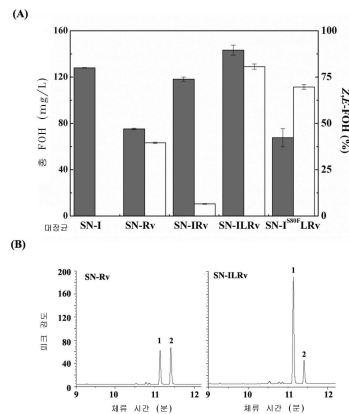
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 Z, E-파네실디포스페이트 신타제와 대장균 파네실디포스페이트 신타제의 융합 효소 및 그 용도

(57) 요약

ispA의 IPP와 DMAPP의 축합을 촉매하여 GPP를 생성하는 활성과 Z,E-파네실디포스페이트 (FPP) 신타제의 활성을 모두 가진 Z,E-파네실디포스페이트 (FPP) 신타제와 대장균 파네실디포스페이트 신타제 (ispA)의 융합 효소 및 그 용도를 제공한다.

대표도 - 도2



(72) 발명자

윤상환

충청남도 부여군 부여읍 서동로 71, 신안빌라 B동
202호

장희정

경상남도 진주시 진주대로829번길 11-13, 305호 (주약동)

최의성

대전광역시 유성구 하기동 송림마을아파트
307-1703

특허청구의 범위

청구항 1

이소펜테닐디포스페이트 (IPP)와 제라닐디포스페이트 (GPP)의 축합을 촉매하여 Z,E-파네실디포스페이트를 생성하는 Z,E-파네실디포스페이트 (FPP) 신타제와 대장균의 파네실디포스페이트 신타제 (ispA)의 융합 효소로서 ispA의 IPP와 DMAPP의 축합을 촉매하여 GPP를 생성하는 활성과 Z,E-파네실디포스페이트 (FPP) 신타제의 활성을 모두 가진 융합 효소.

청구항 2

제1항에 있어서, Z,E-FPP 신타제는 서열번호 1의 아미노산 서열을 갖는 것이고, 상기 ispA는 서열번호 2의 아미노산 서열을 갖는 것인 융합 효소.

청구항 3

제1항에 있어서, Z,E-FPP와 ispA는 유연한 펩티드 링커에 의하여 연결된 것인 융합 효소.

청구항 4

제3항에 있어서, 상기 링커는 서열번호 7의 아미노산 서열을 갖는 것인 융합 효소.

청구항 5

제1항 내지 제4항 중 어느 한 항의 융합 효소의 아미노산 서열을 코딩하는 핵산.

청구항 6

제1항 내지 제4항 중 어느 한 항에 따른 융합 효소의 아미노산 서열을 코딩하는 핵산이 도입된 에세리키아 속 미생물.

청구항 7

제6항에 있어서, 대장균인 미생물.

청구항 8

제6항에 있어서, 메발로네이트(MVA) 경로를 이용하도록 형질전환된 에세리키아 속 미생물로서,

a) 스트렙토코커스 뉴모니에 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니에 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니에 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자, 또는

b) 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트랜스퍼라제 및 하이드록시메틸 글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 유전자, 스트렙토코커스 뉴모니에 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니에 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니에 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자로 형질전환된 미생물.

청구항 9

제8항에 있어서, a)에 있어서, 스트렙토코커스 뉴모니에 유래의 메발로네이트 키나제가 서열번호 8, 스트렙토코커스 뉴모니에 유래의 포스포메발로네이트 키나제가 서열번호 9, 스트렙토코커스 뉴모니에 유래의 메발로네이트 디포스페이트 데카르복실라제가 서열번호 10, 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제가 서열번호 11의 아미노산 서열을 갖는 것인 미생물.

청구항 10

제8항에 있어서, b)에 있어서, 엔테로코커스 패칼리스 유래의 아세틸-CoA 아세틸트랜스퍼라제 및 하이드록시메틸글루타릴 (HMG)-CoA 리덕타제가 서열번호 12, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제가 서열번호 13, 스트렙토코커스 뉴모니아 유래의 메발로네이트 키나제가 서열번호 8, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제가 서열번호 9, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제가 서열번호 10, 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제가 서열번호 11의 아미노산 서열을 갖는 것인 미생물.

청구항 11

제6항에 있어서, 상기 미생물은 대장균 유래의 ispA 유전자로 더 형질전환된 것인 미생물.

청구항 12

제11항에 있어서, 상기 대장균 유래의 ispA 유전자는 서열번호 2의 아미노산 서열을 코딩하는 것인 미생물.

청구항 13

제6항에 따른 미생물을 배양하는 단계; 및

배양물로부터 FPP 또는 파네솔을 분리하는 단계를 포함하는, 미생물로부터 FPP 또는 파네솔을 생산하는 방법.

청구항 14

제13항에 있어서, 상기 배양은 배지 위에 데칸 상을 위치시켜 배양하는 것인 방법.

명 세 서

기술분야

[0001] 본 발명은 Z,E-파네실디포스페이트 (FPP) 신타제와 대장균 파네실디포스페이트 신타제 (ispA)의 융합 효소 및 그 용도에 관한 것이다.

배경 기술

[0002] 이소프레노이드는 2-C-메틸-D-에리트리톨 4-포스페이트 (MEP) 경로 또는 메발로네이트 (MVA) 경로로부터 생성된 5-탄소 (C6) 빌딩 블록 이소펜테닐 디포스페이트 (IPP) 및 디메틸알릴 피로포스페이트 (DMAPP)로부터 합성된다 (Bloch et al. 1959; Rohdich et al. 2002). 이소프레노이드 알릴 디포스페이트의 성장하는 탄화수소 사슬이 IPP에 부가되는, IPP와 DMAPP의 헤드-테일 축합은 긴 이소프레닐 디포스페이트를 생성한다 (Lange et al. 2000).

[0003] 이들 사슬 신장은 새로 형성된 이중결합에 대하여 두 가지 방식, E 및 Z로 진행될 수 있다. 따라서, 입체특이적 반응 (stereospecific reaction)은 네 가지 구조 (configurations) (E,E, E,Z, Z,E and Z,Z)를 가진 파네실 디포스페이트 (FPP) 산물을 생성한다 (Thulasiram et al. 2007). 따라서, 각각 E-FPP 신타제 및 Z-FPP 신타제가 입체 화학 반응에 관여한다. 대장균 IspA를 포함한 대부분의 FPP 신타제는 모두 E-이중 결합 형성 효소로서, 1 DMAPP와 2 IPP의 E-축합을 촉매하여 E,E-FPP를 형성한다. E,E-FPP는 살아 있는 개체에서 콜레스테롤 파네실화 단백질, 유비퀴논 및 돌리콜과 같은 필수 세포 성분의 생합성에 이용되었다 (Ashby and Edwards 1990; Clarke 1992; Matsuoka et al. 1991; Zhang et al. 1993).

[0004] 그러나, FPP 신타제가 GPP 중간체를 방출하는지는 잘 밝혀지지 않았다 (도 1) (Thulasiram and Poulter 2006). 비록 긴-사슬 (C40-C200) cis-이소프레닐 신타제가 대부분의 개체에서 흔하지만 (Bugg and Brandish 1994), Z-FPP 신타제는 오늘날까지 거의 클로닝되고 확인되지 않았다. *Mycobacterium tuberculosis*의 Rv1086은 GPP와 IPP의 Z-축합을 촉매하는 Z,E-FPP 신타제라는 것이 보고되었다 (Schulbach et al. 2000).

[0005] 세스퀴테르펜은 범 알리시클릭 전구체 FPP로부터 유래된 천연 산물의 큰 패밀리이다. 지금까지 발견된 세스퀴테르펜 생합성과 연관된 대부분의 FPP 신타제는 E 타입이다. 따라서, E,E-FPP가 세스퀴테르펜 신타제의 기질인 것으로 당연히 여겨졌으나, E,E-FPP가 실제로 가장 선호하는 기질인지는 입증되지 않았다 (Faraldos et al.). 사실, 세스퀴테르펜 합성에 Z-타입 (E,Z-, Z,E- 및 Z,Z-) FPP의 관여는 이전 연구에서 일부 힌트가 주어졌다. 예

를 들면, 면으로부터의 조 효소 추출물은 E,E-입체이성질체의 촉매화에 비하여, 더 높은 효율로 Z,E- 또는 Z,Z-FPP를 δ -카디넨으로 전환시킨다 (Heinstein et al. 1970). 토마토 유래의 세스퀴테르펜, endo-베르가모텐의 고리화는 입체화학적 제한으로 인하여, E,E-FPP 보다는 Z,E-FPP로부터 더 선호된다고 제안되었다 (Coates and Denissen 1988). 최근에는, 토마토유래의 사트날렌 및 베르가모텐 신타제 (SBS)는 유일한 기질로서 Z,Z-FPP를 사용하는 것이 보고된 바 있다 (Sallaud et al. 2009).

[0006] 상기한 선행기술을 고려하면, 세스퀴테르펜 생산을 위하여 Z-타입 FPP를 효율적으로 생산하는 효소를 만드는 것이 중요하다. 이에 본 발명자들은 흔하지 않은 전구체인를 사용하여 세스퀴테르펜의 과생산을 위한 수단을 제공하는, Z,E-FPP를 우선적으로 합성할 수 있는 에세리키아 속 균주를 개발하였다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 일 양상은 Z,E-파네실디포스페이트 (FPP) 신타제와 대장균 파네실디포스페이트 신타제 (ispA)의 융합 효소 및 그를 코딩하는 핵산을 제공한다.

[0008] 본 발명의 다른 양상은 상기 융합 효소를 코딩하는 핵산을 포함하는 미생물 및 그를 이용하는 방법을 제공한다.

과제의 해결 수단

[0009] 본 발명의 일 양상은 이소펜테닐디포스페이트 (IPP)와 제라닐디포스페이트 (GPP)의 축합을 촉매하여 Z,E-파네실디포스페이트를 생성하는 Z,E-파네실디포스페이트 (FPP) 신타제와 대장균의 파네실디포스페이트 신타제 (ispA)의 융합 효소로서 ispA의 IPP와 DMAPP의 축합을 촉매하여 GPP를 생성하는 활성과 Z,E-파네실디포스페이트 (FPP) 신타제의 활성을 모두 가진 융합 효소를 제공한다.

[0010] Z,E-FPP 신타제는 *Mycobacterium tuberculosis* 유래의 Rv1086 효소일 수 있다. Z,E-FPP 신타제는 서열번호 1의 아미노산 서열을 갖는 것일 수 있다. ispA는 대장균 유래의 것일 수 있다. ispA는 서열번호 2의 아미노산 서열을 갖는 것일 수 있다. Z,E-FPP 신타제와 ispA는 유연한 링커에 의하여 연결된 것일 수 있다. 상기 링커는 펩티드 링커일 수 있다. 용어 "펩티드 링커"란 아미노산이 펩티드 결합에 연결되어 있는 것을 포함한다. 상기 펩티드 링커는 총 아미노산 중 글리신이 50% 이상인 유연한 펩티드 링커인 것일 수 있다. 상기 링커는 (GGGG)₂ (서열번호 7)의 아미노산 서열을 갖는 것일 수 있다. Z,E-FPP 신타제와 ispA는 선택적으로 링커를 매개하여, 한 융합 파트너의 C 말단에 다른 파트너의 N 말단이 연결되거나, 그 반대의 방식으로 연결될 것일 수 있다. 예를 들면, 상기 융합 효소는 서열번호 2의 ispA의 C 말단에 서열번호 7의 링커가 연결되고 여기에 서열번호 1의 Z,E-FPP 신타제가 연결된 형태인, 서열번호 3의 아미노산 서열을 갖는 것일 수 있다.

[0011] 본 발명의 다른 양상은 상기 융합 효소의 아미노산 서열을 코딩하는 핵산을 제공한다. 상기 핵산은 서열번호 4 및 5의 서열을 포함하는 것 또는 서열번호 6의 서열일 수 있다. 상기 핵산은 에세리키아 속 미생물, 예를 들면, 대장균에서 발현 또는 증폭될 수 있는 조절 서열이 포함된 플라스미드 또는 벡터에 포함된 것일 수 있다.

[0012] 본 발명의 다른 양상은 상기한 융합 효소의 아미노산 서열을 코딩하는 핵산이 도입된 에세리키아 속 미생물을 제공한다. 상기 미생물은 대장균일 수 있다.

[0013] 상기 미생물은 메발로네이트 경로의 유전자가 도입되어 탄소원으로부터 IPP 또는 DMAPP를 합성할 수 있는 것일 수 있다. 상기 미생물에 메발로네이트 경로의 유전자를 도입하는 것은 알려져 있다. 예를 들면, 한국특허공개 2009-78113에는 외래 메발로네이트 경로 유전자를 도입하여 이소프레노이드를 생산할 수 있는 대장균 및 그를 이용한 이소프레노이드 생산 방법이 개시되어 있다.

[0014] 상기 미생물은 메발로네이트(MVA) 경로를 이용하도록 형질전환된 에세리키아 속 미생물로서, a) 스트렙토코커스 뉴모니아 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카복실라제를 코딩하는 유전자 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자, 또는 b) 엔테로코커스 패칼리스 (*Enterococcus faecalis*) 유래의 아세틸-CoA 아세틸트랜스퍼라제 및 하이드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자, 엔테로코커스 패칼리스 유래의 HMG-CoA

신타제를 코딩하는 유전자, 스트렙토코커스 뉴모니아 (*Streptococcus pneumoniae*) 유래의 메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제를 코딩하는 유전자, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제를 코딩하는 유전자로 형질전환된 미생물일 수 있다.

[0015] a)에 있어서, 스트렙토코커스 뉴모니아 유래의 메발로네이트 키나제가 서열번호 8, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제가 서열번호 9, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제가 서열번호 10, 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제가 서열번호 11의 아미노산 서열을 갖는 것일 수 있다. 상기 미생물은 벡터 pSSN12Didi를 포함하는 미생물, 예를 들면 대장균일 수 있다.

[0016] b)에 있어서, 엔테로코커스 패칼리스 유래의 아세틸-CoA 아세틸트랜스퍼라제 및 하이드록시메틸글루타릴 (HMG)-CoA 리덕타제가 서열번호 12, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제가 서열번호 13, 스트렙토코커스 뉴모니아 유래의 메발로네이트 키나제가 서열번호 8, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제가 서열번호 9, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제가 서열번호 10, 및 대장균 유래의 이소펜테닐 디포스페이트 (IPP) 이소머라제가 서열번호 11의 아미노산 서열을 갖는 것일 수 있다. 상기 미생물은 벡터 pSNA를 포함하는 미생물, 예를 들면 대장균일 수 있다.

[0017] 상기 미생물은 대장균 유래의 ispA 유전자로 더 형질전환된 것일 수 있다. 상기 대장균 유래의 ispA 유전자는 서열번호 2의 아미노산 서열을 코딩하는 것, 예를 들면, 서열번호 5의 염기 서열을 갖는 것인 미생물일 수 있다.

[0018] 본 발명의 다른 양상은 상기한 미생물을 배양하는 단계; 및 배양물로부터 FPP 또는 파네솔을 분리하는 단계를 포함하는, 미생물로부터 FPP 또는 파네솔을 생산하는 방법을 제공한다.

[0019] 상기 방법에서 상기 미생물을 배양하는 것은 배지 위에 데칸 상(decane phase)을 위치시켜 배양하는 것일 수 있다. 따라서 휘발성이 있는 파네솔을 손실 없이 수득할 수 있고, 데칸에 용해되어 있는 파네솔을 수득함으로써 수확 및 정제 방법을 모두 단순화하여 생산 비용을 줄일 수 있다.

[0020] 본 발명의 에세리키아 속 미생물로부터 FPP 또는 파네솔을 생산하는 방법은 상기한 바와 같은 본 발명의 미생물을 배양하는 단계를 포함한다.

[0021] 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, 바람직하기로는 글리세롤이다. 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.

[0022] 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.

[0023] 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.

[0024] 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.

[0025] 상기 배양의 일 예에 있어서, 탄소원으로서 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 예를 들면, 0.5 % (v/v) 내지 2 % (v/v)의 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 또한 배양의 일 예에 있어, MVA 경로의 중간 산물, 예를 들면 MVA를 포함하는 배지에서 이루어지는 것일 수 있다. 미생물이 하부 MVA 경로만을 이용하도록 형질전환된 경우, 예를 들면 5mM MVA가 부가된 배지에서 배양이 이루어질 것이다.

또한 배양의 일 예에 있어, 상기 배양은 IPTG 유도와 함께 이루어질 수 있다. 예를 들면, 0.1mM IPTG 유도하에 배양이 이루어진다. 배지는 2YT 배지 (1L 당 16 g 트립토판, 5 g 소듐 클로라이드, 및 10 g 효모 추출물)일 수 있고 항생제(엠페실린 100 $\mu\text{g/mL}$, 클로람페니콜 50 $\mu\text{g/mL}$, 및 카나마이신 50 $\mu\text{g/mL}$)를 필요로 하는 만큼 첨가할 수 있다. 또한 배양의 일 예에 있어서, 항생제, 글리세롤, 및/또는 MVA가 첨가된 2YT 배지 위에 데칸 상이 위치할 수 있다.

[0026] 본 발명의 미생물에 대하여는 상기한 바와 같다.

[0027] 본 발명의 방법은 또한, 배양물로부터 FPP 또는 파네솔을 분리하는 단계를 포함한다. 상기 파네솔을 분리하는 것은 당업계에 잘 알려져 있다. 구체적으로 2-상 배양에서, 데칸 상을 수집하여 원심분리, 예를 들면 10분 동안 12,000 rpm으로 원심분리하여 세포 파편을 제거하고 그 후 박층 크로마토그래피 (TLC), 가스 크로마토그래피 (GC) 등의 방법에 의해 분리될 수 있다.

발명의 효과

[0028] 본 발명의 일 양상에 따른 융합 효소는 IPP로부터 GPP 생성을 촉매할 수 있는 동시에 GPP로부터 비대사성 대사산물인 Z,E-FPP 생성을 촉매할 수 있다.

[0029] 본 발명의 일 양상에 따른 핵산은 상기 융합 효소를 코딩할 수 있다.

[0030] 본 발명의 일 양상에 따른 미생물에 의하면 비대사성 대사산물 Z,E-FPP를 효율적으로 합성할 수 있으며, 그로부터 파생하는 세스퀴테르펜 산물을 효율적으로 합성할 수 있다.

[0031] 본 발명의 일 양상에 다른 방법에 의하면, 미생물로부터 FPP 또는 파네솔을 효율적으로 생산할 수 있다.

도면의 간단한 설명

[0032] 도 1은 E,E-FOH 및 Z,E-FOH 합성 경로의 스킴이다. 전구체 IPP 및 DMAPP는 MEP 또는 MVA 경로에 의하여 합성된다. E,E-FPP는 IPP 및 DMAPP의 축합에 의하여 IspA에 의하여 합성된다. Rv1086는 한 분자의 GPP와 한 분자의 IPP의 축합을 촉매하여 Z,E-FPP만을 형성한다. 경로 중간체에 대한 약자는 다음과 같다: DMAPP, 디메틸알릴 피로포스페이트; IPP, 이소펜테닐 피로포스페이트; GPP, 제라닐 피로포스페이트; E,E-FPP, E,E-파네실 피로포스페이트; Z,E-FPP, Z,E-파네실 피로포스페이트; E,E-FOH, E,E-파네솔; Z,E-FOH, Z,E-파네솔. 효소 약자는 다음과 같다: IspA, E,E-FPP 신타제, Rv1086, *M. tuberculosis*의 Z,E-FPP 신타제; 포스파타제, 대장균의 미지 내재적 포스파타제.

도 2는 *M. tuberculosis*의 Rv1086를 발현하는 재조합 대장균으로부터 Z,E-FOH 형성을 나타낸다 (A). 닫힌 바 및 열린 바는 각각 총 FOH 생산 및 Z,E-FOH 백분율을 나타낸다. 대장균 SN-Rv 및 SN-ILRv의 GC 크로마토그램을 나타낸다 (B). 피크 1 및 피크 2는 Z,E-FOH 및 E,E-FOH를 나타낸다.

도 3은 정제된 IspA-Rv1086 (ILRv) 융합 단백질의 SDS-PAGE 및 천연-PAGE 분석을 나타낸다 (A 및 B). 융합 효소의 가능한 구조 (conformations) (C). n=2 및 4는 ILRv 모노머의 수를 나타낸다. 다이머에서 유연한 글리신-세린 링커의 움직임 스킴을 나타낸다 (D).

도 4는 IspA 및 Rv1086의 상보적 발현에 의한 세포 성장 구제를 나타낸다. 균주는 표 1에 나타내었다. 세포 성장은 다양한 시간에서 결정되었다.

도 5는 본 발명의 실시예에 사용된 벡터 지도의 일부를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

[0033] 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.

재료 및 방법

[0035] 이하의 실시예에서는 달리 언급이 없으면, 하기 재료 및 방법을 사용하였다.

(1) 균주 및 생육 조건

[0037] 본 실시예에 사용된 균주는 표 1에 나타내었다. 대장균 DH5 α 를 플라스미드 작성 (플라스미드 construction) 및 FOH 생산을 위한 숙주 균주로서 하였다. 대장균 DH5 α 중 FPP 신타제 유전자 (*ispA*)를 Red/ET 재조합(유전자

Bridges GmbH, Germany)에 의하여 결실시켜 대장균 DH5 α ($\Delta ispA$)를 생성하였으며, 이를 균주를 성장 구제 연구 (growth rescue studies)를 위한 숙주 세포로 사용하였다. *ispA* 결실을 위한 카나마이신 저항성 카세트를 프라이머 DispA-F 및 DispA-R (표 1)를 사용한 PCR에 의하여, 대장균 SF7N 계놈 DNA (Fujisaki et al. 2005)로부터 증폭하였다. 대장균 BL21 (DE3)를 단백질 정제를 위한 숙주로 사용하였다. 대장균 DH5 α 를 플라스미드 작성을 위하여 2YT (16 g/L tryptone, 10 g/L yeast extract, and 5 g/L NaCl) 배지 중에서 37°C에서 생육시켰다.

[0038] FOH 생산 및 성장 구제를 위하여 사용된 대장균 균주들은 1% (v/v) 글리세롤 및 0.1 mM IPTG를 포함한 2YT 배지 중에서 29°C에서 배양되었다. MVA 경로의 하부 부분 (bottom portion)을 코딩하는 pSSN12Didi 플라스미드를 포함하는 균주에 대하여는 5mM 농도의 메탈로네이트를 공급하였다 (Yoon et al. 2009; Yoon et al. 2006). 이전에 보고된 바와 같이, FOH 생산을 위하여 데칸 씌워진 배양 (decane overlaid culture)을 수행하였다 (Wang et al. 2010). 대장균 BL21 (DE3)는 단백질 정제를 위하여 LB (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) 중에서 22°C에서 생육되었다. 600 nm에서의 광밀도 (OD_{600})가 0.6에 도달하였을 때 단백질을 발현을 유도하기 위하여, 0.1mM 농도의 IPTG를 첨가하였다. 암피실린 (100 μ g/ml), 클로람페니콜 (50 μ g/ml) 및 카나마이신 (50 μ g/ml)를 필요에 따라 보충하였다.

[0039] 표 1. 본 실시예에 사용된 균주, 플라스미드 및 프라이머

표 1

명칭	설명	참조
균주		
BL21 (DE3)	F^- , <i>ompT</i> , <i>hsdS_B</i> (r_B^- , m_B^-), <i>dcm</i> , <i>gal</i> , λ (DE3)	Novagen
SF7N	F^- , λ^- , IN(<i>rrnD-rrnE</i>)1, <i>rph</i> -1, $\Delta ispA::FRT-Kan-FRT$	(Fujisaki 등. 2005)
DH5 α	F^- , <i>fhuA2</i> , Δ (<i>argF-lacZ</i>)U169, <i>phoA</i> , <i>glnV44</i> , $\Phi 80\Delta lacZ$ M15, <i>gyrA96</i> , <i>recA1</i> , <i>relA1</i> , <i>endA1</i> , <i>thi</i> -1, <i>hsdR17</i> (r_K^- , m_K^+)	ATCC
SN-Rv	pSSN12Didi 및 pTRv1086를 포함하는 DH5 α	본 발명
SN-I	pSSN12Didi 및 pTispA를 포함하는 DH5 α	본 발명
SN-IRv	pSSN12Didi 및 pTispARv1086를 포함하는 DH5 α	본 발명
SN-ILRv	pSSN12Didi 및 pTispALRv1086를 포함하는 DH5 α	본 발명
SN-I ^{S80F} LRv	pSSN12Didi 및 pTispA ^{S80F} LRv1086를 포함하는 DH5 α	본 발명
DH5 α $\Delta ispA$	DH5 α $\Delta ispA::FRT-Kan-FRT$	본 발명
WT-C	pTrc99A를 포함하는 DH5 α	본 발명
MU-C	pTrc99A를 포함하는 DH5 α $\Delta ispA$	본 발명
MU-I	pTispA를 포함하는 DH5 α $\Delta ispA$	본 발명
MU-Rv	pTRv1086를 포함하는 DH5 α $\Delta ispA$	본 발명
MU-ILRv	pTispALRv1086를 포함하는 DH5 α $\Delta ispA$	본 발명
MU-I ^{S80F} LRv	pTispA ^{S80F} LRv1086를 포함하는 DH5 α $\Delta ispA$	본 발명
플라스미드		
pTrc99A	P_{trc} ; <i>lacI^q</i> ; pBR322 ori; Amp^R	
pTRv1086	$P_{trc}::Rv1086$; <i>lacI^q</i> ; pBR322 ori; Amp^R	본 발명
pTispA	$P_{trc}::ispA$; <i>lacI^q</i> ; pBR322 ori; Amp^R	(Wang et al. 2010)
pTispARv1086	$P_{trc}::ispA$, <i>Rv1086</i> ; <i>lacI^q</i> ; pBR322 ori; Amp^R	본 발명
pTispALRv1086	$P_{trc}::ispA$ LRv1086; <i>lacI^q</i> ; pBR322 ori; Amp^R	본 발명
pTispA ^{S80F} LRv1086	$P_{trc}::ispA^{S80F}$ LRv1086; <i>lacI^q</i> ; pBR322 ori; Amp^R	본 발명
pETispALRv1086	$P_{T7}::his6$ tag- <i>ispA</i> ^{S80F} LRv1086; <i>lacI</i> ; pBR322 ori; Kan^R	본 발명
pSSN12Didi	$P_{lac}::mvaK1$, <i>mvaK2</i> , <i>mvaD</i> , <i>idi</i> ; p15A ori; Cm^R	(Yoon et al. 2006; Yoon et al. 2007)

프라이머		
Rv1086-F1	TAGGATCC ATG GAGATCATCCCGCCGCG (서열번호 12)	본 발명
Rv1086-F2	TGGATCCAGGAGGTAATAAAC CATG GAGATCATCCCGCCGCG (서열번호 13)	본 발명
Rv1086-R	ATCTCGAG T TACCTGCCGTAGCTGCGATGC (서열번호 14)	본 발명
ispA ^{S80F} -F	GAGTGTATCCACGCTTACTTTTAAATTCATGATGATTACCG (서열번호 16)	본 발명
ispA ^{S80F} -R	AAAGTAAGCGTGGATACACTCAACGGCGGCAGC (서열번호 17)	본 발명
DispA-F	GCGTACAAATTCTGCTGTCTGAC	본 발명
DispA-R	CAACAGTCGTAACCTCTGGGTG	본 발명

[0041] 비교: 개시 코돈 및 중지 코돈은 볼드체로 되어 있고, 효소 부위는 밑줄쳐져 있고, 리보솜 결합 부위는 이탤릭체이다.

[0042] (2) 플라스미드 작성

[0043] 플라스미드 작성을 위하여 표준 분자 생물학 기법을 사용하였다 (Sambrook and Russell 2001).

[0044] *Mycobacterium tuberculosis* H37Rv (ATCC 25177D-5)의 게놈 DNA로부터 프라이머 Rv1086-F1 (서열번호 12) 및 Rv1086-R (서열번호 14)를 사용한 PCR에 의하여 *Rv1086* 유전자를 증폭하고, BamHI 및 XhoI로 자르고, pTrc99A 벡터 (서열번호 15)의 BamHI/SalI 부위에 클로닝하여, pTrv1086 플라스미드를 생성하였다.

[0045] pTispArv1086 플라스미드를 프라이머 Rv1086-F2 (서열번호 13) 및 Rv1086-R (서열번호 14)을 사용한 PCR에 의하여 증폭된 *Rv1086* 유전자 (Wang et al. 2010)를 pTispA의 BamHI/SalI 부위에 삽입하여 제조하였다. pTispALrv1086 플라스미드의 제작을 위하여, *Rv1086* 유전자를 *ispA*의 TAA 정지 코돈이 (GGGS)₂ 랩터드 링커를 코딩하는 올리고뉴클레오타이드로 치환된 pTispAL의 BamHI/SalI 부위에 삽입하였다 (Wang et al. 2011b). 플라스미드 pTispA^{S80F}LRv1086는 프라이머 ispA^{S80F}-F (서열번호 16) 및 ispA^{S80F}-R (서열번호 17)를 사용한 부위 지향 돌연변이 (site-directed mutation)에 의하여 생성되었다.

[0046] pTispALrv1086 플라스미드의 작성을 위하여, EcoRI 및 HindIII에 의하여 pTispALrv1086로부터 잘려진 *ispALrv1086* 단편을 pET28a (Novagen: 서열번호 18)의 EcoRI/HindIII에 삽입하였다. 이하의 실시예에 사용된 상기 플라스미드 및 PCR 프라이머는 표 1에 열거하였다.

[0047] (3) FOH 생산의 정량

[0048] FOH의 농도는 기술된 바와 같은 가스 크로마토그래피 (GC)법에 의하여 정량하였다 (Wang et al. 2010). 간단하게 설명하면, 화염 이온화 검출기(flame ionization detector: FID) 및 자동샘플기를 구비한 가스 크로마토그램 (Agilent Technologies 7890A)을 사용하였다. 시료를 HP-INNOWAX column (19091N-133, 길이 30 m, 내경 0.25 mm, 및 피름 250 μ m) 중에 1:10의 분할 비율 (split ratio)로 주입하였다.

[0049] 오븐 온도는 1 분 동안 50℃에서 개시되어 150℃까지 20℃/분의 속도로 상승하였고 250℃까지 10℃/분의 속도로 상승하였고, 250℃에서 1분 동안 유지하였다. 질소 기체를 입구 압력 39 psi의 캐리어 기체로서 사용하였다. 검출기를 260℃에서 유지하였다. GC 크로마토그래프에서 관찰된 11.1분 (피크 1) 및 11.4 분 (피크 2)의 보류 시간의 두 개 피크가 각각 Z,E-FOH 및 E,E-FOH였다. E,E-FOH (Sigma-Aldrich)의 표준곡선을 Z,E-FOH 및 E,E-FOH 모두의 농도를 추정하기 위하여 사용하였다.

[0050] (4) IspA-Rv1086 (ILRv) 융합 단백질의 SDS-PAGE 및 천연-PAGE 분석

[0051] IspA-Rv1086 (ILRv) 융합 단백질을 pTispALrv086 플라스미드를 포함한 대장균 BL21 (DE3)의 0.1 mM IPTG 유도된 배양으로부터 정제하였다. 세포를 원심분리에 의하여 회수하고, 400mM NaCl 및 10mM 이미다졸을 포함한 20mM 포스페이트 버퍼, pH 7.9 중에서 초음파 처리에 의하여 파쇄되었다. 가용성 His-태그된 단백질을 니켈-니크틸로아세트산-아가로즈 칼럼 (Invitrogen)에 의하여 정제하였다.

[0052] 용출된 단백질을 4-12% SDS-PAGE 또는 천연-PAGE에 의하여 분석하였다. 겔을 쿠마시-블루 염색에 의하여 가시화하였다.

[0053] 실시예 1: Rv1086 단독 발현에 의한 *Z,E*-FPP 합성

[0054] 본 발명자들은 세포 내 FPP가 고수준으로 축적된 경우 대장균 중의 일부 내재적 포스파타제가 FPP를 FOH로 가수분해할 것으로 가정하였다. 실제로, *E,E*-FOH 생산은 IspA 발현의 강화 (enhancement) 및 IPP 및 DMAPP의 공급을 위한 잘-수행된 MVA 경로의 도입에 의하여 유도되었다 (Wang et al. 2010).

[0055] 그러한 포스파타제가 넓은 기질 특이성을 갖는 것을 가정하여, MVA 경로의 하부 부분을 코딩하는 pSSN12Didi를 갖는 대장균 중에서 *Z,E*-FPP 신타제 Rv1086을 발현시켰다. 예상한 대로, Rv1086의 기능적 발현은 대장균 SN-Rv에서 *Z,E*-FOH 형성을 유도하였다 (표 1). 반면, 대장균 SN-I는 유일한 산물로서 *E,E*-FOH를 생산하였다. *Z,E*-FOH 생산의 성공은 조작된 균주로부터의 FPP 스테레오타입 (stereotype)을 확인하기 위한 교묘한 방법을 제안한다.

[0056] 더욱이, Rv1086는 단지 GPP로부터 *Z,E*-FPP로의 단계만을 촉매한다. 본 결과는 또한 지정된 (designated) GPP 신타제 없이 대장균 중에서 *Z,E*-FPP를 합성하기 위하여 Rv1086가 GPP를 이용할 수 있다는 것을 나타낸다 (도 1).

[0057] 대장균 IspA는 DMAPP의 GPP, C10 중간체를 거친 *E,E*-FPP로의 연속적 사슬 신장을 촉매한다 (Fujisaki et al. 1990). 본 발명자들은 FPP 합성 동안 중간체로서 Rv1086에 이용가능한 유리 GPP가 IspA로부터 방출되는 것으로 생각하였다.

[0058] 따라서, Rv1086는 총 (*E,E*- 및 *Z,E*-) FOH의 39.5%의 *Z,E*-FOH 생산에 이르게 하는, 공통의 GPP 기질로부터 *Z,E*-FPP 합성을 위하여 IspA와 경쟁할 수 있다 (도 2).

[0059] 본 실시예에서 사용된 pSSN12Didi는 *S. 뉴모니아*의 mvaK1, mvaK2, 및 mvaD, 및 대장균의 idi를 포함하는 pSTV28 벡터이고, pSTV28 벡터 (서열번호 21)는 P_{lac} 발현 벡터, pACYC184 기원(origin), lacZ, Cm^r는 특성을 갖는다. 본 명세서에서 사용된 탄소원으로부터 IPP를 합성하는데 관여하는 메발로네이트 경로의 효소를 코딩하는 유전자는 다음 표 2와 같다.

표 2

효소명	유전자	유전자 서열 (Genbank 허가번호)
엔테로코커스 패칼리스 (<i>Enterococcus faecalis</i>) 유래의 아세틸-CoA 아세틸트랜스퍼라제 및 하이드록시메틸글루타릴 (HMG)-CoA 리덕타제	mvaE	서열번호 22(AF290092)
엔테로코커스 패칼리스 유래의 HMG-CoA 신타제	mvaS	서열번호 23(AF290092)
스트렙토코커스 뉴모니아 (<i>Streptococcus pneumoniae</i>) 유래의 메발로네이트 키나제	mvaK1	서열번호 24(AF290099)
스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제	mvaK2	서열번호 25(AF290099)
스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제	mvaD	서열번호 26 (AF290099)
대장균 유래의 이소펜테닐 디포스페이트(IPP) 이소머라제	idi	서열번호 27 (U00096)

[0061] pSSN12Didi의 제조는 다음과 같이 제조되었다. 먼저, 스트렙토코커스 뉴모니아의 염색체 DNA 주형 및 서열번호 28 및 29의 염기서열을 갖는 프라이머를 이용하여 mvaK1, mvaD 및 mvaK2 오픈론을 증폭하였으며, mvaK1, mvaK2, mvaD는 하나의 오픈론으로 염색체상에 존재해서 각각의 유전자를 PCR 클로닝하지 않고 오픈론을 통째로 한번에 PCR 클로닝을 하였다. 증폭된 유전자 단편을 KpnI 및 XbaI로 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 30 및 31의 염기서열을 갖는 프라이머를 이용하여 증폭한 후, 이를 SmaI과 SphI로 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 19)에 삽입하여 플라스미드 pSSN12Didi를 제작하였다.

[0062] pSNA는 다음과 같이 제조되었다. 표 3은 표 2의 유전자를 증폭하기 위한 프라이머 및 제한 효소를 나타낸다.

표 3

[0063]

유전자	프라이머 서열		제한 효소
mvaE	F	서열번호 32	SacI
	R	서열번호 33	SmaI
mvaS	F	서열번호 34	SmaI
	R	서열번호 35	BamHI
mvaK1, mvaK2, mvaD	F	서열번호 28	KpnI
	R	서열번호 29	XbaI
idi	F	서열번호 30	SmaI
	R	서열번호 31	SphI

[0064]

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 19)에 도입하여, 벡터 pSNA를 제조하였다. 벡터 pSNA는 아세틸-CoA로부터 IPP를 생산할 수 있는 메발로네이트 경로의 효소를 코딩하는 유전자를 모두 포함하고 있다.

[0065]

실시예 2: IspA-Rv1086 융합 키메라에 의한 Z,E-FPP 합성으로의 GPP 흐름의 지시

[0066]

대장균 SN-Rv (75.3 mg/L)의 총 FOH 생산은 대장균 SN-I (128.0 mg/L)의 그것보다 41% 더 낮다는 것은 흥미롭다. 대장균 SN-Rv만이 그 염색체에 한 카피의 *ispA*를 가지고 있기 때문에, GPP 이용가능성이 이 균주에서 FOH 생산을 한정할 수 있다. 추가 IspA 과발현은 예상과 같이 대장균 SN-IRv (118.2 mg/L)에서 대장균 SN-I에 견줄 만큼의 FOH 생산을 유도하였으나, 또한 총 FOH 생산의 단지 6.5%, Z,E-FOH 생산 (7.7 mg/L)의 유의성 있는 감소를 유도하였다.

[0067]

이는 IspA가 GPP 생산자일 뿐만 아니라 Rv1086과 경쟁하는 GPP 소비자임을 확인한다. 이 결과는 또한 진행성 (processive) IspA는 IPP의 DMAPP와 GPP와의 축합은 다른 효율로 촉매한다는 것 즉, 사슬 신장의 C5->C10 단계 (GPP 형성)은 C10->C15 단계 (E,E-FPP 형성) 보다 더 효율적이라는 것을 나타낸다. C10 중간체의 자유 확산은 낮은 IspA 발현의 경우에만 허용되고, 이 경우 제한된 IspA는 우선적으로 E,E-FPP로의 추가적 축합보다는 GPP 합성을 촉매한다.

[0068]

대장균 SN-I 및 대장균 SN-IRv 모두에서 IspA가 과발현되는 경우, 충분한 IspA는 GPP 축합에만 한정되지 않고, E,E-FPP 합성을 처리하기 위하여 GPP를 다시 끌어온다. 그러므로, IspA는 IspA 및 Rv1086 모두를 발현하는 균주 (대장균 SN-IRv)에서 Rv1086에 비하여 GPP 활용 (recruitment)의 타고난 잇점 (innate advantage)를 갖고, Rv1086가 GPP 기질을 활용할 기회는 IspA로부터 GPP 방출의 순간 (transient)에서의 그 위치에 크게 의존한다.

[0069]

이 가정에 의하면, IspA-Rv1086 (ILRv) 인-프레임 (in-frame) 융합 단백질을 IspA에 대하여 Rv1086 상에서 GPP 경쟁을 촉진할지를 결정하기 위하여 제작하였다. 융합 단백질에 기초한 성공적 대사 과정은 키메라 효소의 가용성 발현 및 두 개 효소 활성의 유지를 필요로 한다 (Albertsen et al.).

[0070]

키메라 ILRv의 발현을 SDS-PAGE 분석으로 조사하였다 (도 3A). ILRv의 외견 분자량은 별개의 IspA 및 Rv1086의 거동과 일치하는 이론상 분자량 (63.2 KDa)보다 컸다 (Fujisaki et al. 1990; Schulbach et al. 2001). 이들 두 프레닐트랜스퍼라제 (prenyltransferases) (Hosfield et al. 2004; Wang et al. 2008)사이 구조적 유사성이 없음에도 불구하고, IspA 및 Rv1086 모두 Mg^{2+} -의존성 호모다имер 효소이다. IspA 활성은 호모다имер 형성에 완전하게 의존적이다 (Koyama et al. 2000).

[0071]

본 발명자들은 ILRv 키메라가 천연-PAGE에서 약 130 kDa 및 260 kDa에서 이동하는 것을 관찰하였다 (도 3B). 이는 ILRv가 쉽게 다이머를 형성하고 (도 3C, n = 2), 또한 테트라머를 형성할 것 같다는 것(도 3C, n = 4)을 나타낸다.

[0072]

다음으로, 본 발명자들은 먼저 *ispA* 결실 균주 대장균 DH5α Δ *ispA*를 사용하여 ILRv 키메라의 FPP 합성 활성을 검사하였다. *ispA* 유전자는 대장균 성장에 필수적인 것으로 알려져 있고 (Fujisaki et al. 2005), *ispA*의 결실은 대장균 MU-C의 세포 성장을 심하게 지연시키는 것으로 관찰되었다 (도 4).

[0073]

그러나, 대장균 MU-ILRv에서 ILRv 발현은, 대장균 MU-I에서 ILRv 발현이 하였던 것과 같이, 그러한 성장지체를

회복시켰다.

- [0074] 이는 유연한 글리신-세린 링커 (flexible glycine-serine linker)에 의하여 Rv1086에 융합된 경우, IspA는 FPP 신타제 활성을 보유한다는 것을 나타낸다.
- [0075] Rv1086 도메인의 기능적 발현에 대한 직접적 시험은 *Z,E*-FOH 형성의 측정이었다. 실제로, 대장균 SN-ILRv는 성공적으로, 총 생산 (143.3 mg/L)의 80.7%인 115.6 mg/L의 *Z,E*-FOH를 생산하였다 (도 2).
- [0076] 이는 IspA 및 Rv1086의 융합은 Rv1086 도메인의 *Z,E*-FPP 신타제 활성을 손상시키지 않는다는 것을 나타낸다. 따라서, Rv1086 및 IspA의 공간적 배열 (spatial arrangement)은 *Z,E*-FPP 합성을 위한 Rv1086에게 공평한 GPP 접근과 증가된 GPP 이용을 제공한다.
- [0077] **실시예 3: *Z,E*-FPP는 대장균에서는 비대사성 대사산물 (non-metabolizable metabolite)이다**
- [0078] *E,E*-FPP는 이소프레노이드 경로의 키는 및 지질 캐리어와 같은 필수 대사산물의 합성을 위한 전구체로서 중요한 다분기점 (multiple branch point)에 있다.
- [0079] *E,E*-FPP로부터 유래된 세스퀴테르펜의 생산을 위한 대사 과정에서, *E,E*-FPP 흐름 손실 (flux loss)은 불가피하다.
- [0080] *Z,E*-FPP가 내재적 대사 경로를 통하여 대장균에서 대사되지 않는다면, 예를 들면, *Z,E*-FOH 형성에만 제공되어질 수 있다면, FOH 생산의 증가를 유도한다. 놀랍지않게도, FPP 합성을 위한 알릴 기질(allylic substrate)로서 DMAPP를 이용할 수 없기 때문에, Rv1086의 발현은 *ispA* 돌연변이의 성장 지체를 구제할 수 없었다 (도4, MU-Rv) (Schulbach et al. 2001). 반면 키메라 ILRv는 *E,E*-FPP를 생산하는 능력을 유지한다 (도4, MU-ILRv).
- [0081] *Z,E*-FPP만을 합성하기 위하여, 본 발명자들은 변이체 IspA^{S80F}-Rv1086 (I^{S80F}LRv)를 생성하였으며, S80F 돌연변이는 IspA가 *E,E*-FPP 대신 GPP를 특이적으로 합성하도록 하는 것으로 알려져 있다 (Narita et al. 1999; Reiling et al. 2004).
- [0082] 돌연변이된 단백질 I^{S80F}LRv의 기능적 발현은 대장균 SN- I^{S80F}LRv에서 *Z,E*-FOH 형성에 의하여 입증되었다 (도 2). 대장균 SN- I^{S80F}LRv의 총 FOH 생산은 대장균 SN-Rv의 그것과 비슷하지만, *Z,E*-FOH 생산 (비율: proportion)은 대장균 SN-Rv에 비하여, 1.6(1.7) 배까지 증가하였다. S80F 돌연변이가 IspA의 촉매 활성을 손상시킨다는 것을 고려하면 (Reiling et al. 2004), *Z,E*-FOH 생산의 증가는 I^{S80F}LRv가 IPP 및 DMAPP로부터 *Z,E*-FPP를 생산할 수 있다는 것을 나타낸다.
- [0083] 그러므로, *Z,E*-FPP가 대장균의 대상에 참여하는지를 보기 위하여, 변이체 I^{S80F}LRv를 *ispA* 결실 돌연변이체에 도입하였다. 대장균 MU-C에 나타난 성장 지체가 대장균 SN-I^{S80F}LRv의 이 변이체 발현에 의하여 보충된다면, *E,E*-FPP의 형성 없이 대장균의 이소프레노이드 경로에서 *Z,E*-FPP를 이용한다는 것을 나타내고, 그 반대의 경우도 성립한다. 대장균 MU-I^{S80F}LRv의 세포 성장은 I^{S80F}LRv 발현에 의하여 회복되지 않았으며, 대장균 MU보다 더 악화되었다 (도 4). 그러므로, 본 발명자들은
- [0084] *Z,E*-FPP는 대장균에서 적어도 성장을 위한 필수 이소프레노이드 대사산물의 합성을 위하여 대사될 수 없다는 것으로 가정하였다.
- [0085] 옥타프레닐 디포스페이트 신타제와 운테카프레닐 디포스페이트 신타제는 IPP 및 DMAPP의 축합에 대한 약간의 활성을 보이며, *ispA* 결실 돌연변이의 생존을 유지할 수 있다. (Fujisaki et al. 1990; Fujisaki et al. 1986). I^{S80F}LRv에 의한 "쓸모없는 (useless)" *Z,E*-FPP 합성을 위한 IPP 및 DMAPP의 경쟁적 소비는 대장균 MU-I^{S80F}LRv의 악화된 세포 성장을 아마 야기시켰고, 대장균 MU-Rv에서도 그럴 것이다.
- [0086] 이런 관점에서, 내재적 이소프레노이드 합성으로의 *Z,E*-FPP 흐름의 손실 및 *Z,E*-합성에 대한 피드백 조절이 적기 때문에, 전구체로서 *Z,E*-FPP를 이용하는 것은 *E,E*-FPP를 사용하는 것에 비하여, 세스퀴테르펜 생산에 유리할 것이다 (Martin et al. 2003).
- [0087] **토의**
- [0088] 부가 가치가 있는 화합물질의 생산을 위하여 미생물을 대사 조작하는 것은 산업적 적용에 기대가 크다.

- [0089] 효소적 과정은 화학 합성을 통하여 이루기 어려운 내재적 입체특이성 (intrinsic stereospecificity)을 가진 촉매 반응으로 알려져 있다 (Yu et al. 2005). 본 실시예에서, 본 발명자들은 *Z*-타입 세스퀴테르펜의 합성을 위한 기질인, *Z,E*-FPP를 생산하기 위하여 *M. tuberculosis* Rv1086를 대장균 IspA에 융합하였다. 실제로, 115.6 mg/L 역가의 *Z,E*-FOH가 대장균에서 성공적으로 생산되었다. 적절한 세스퀴테르펜 신타제가 is functionally co-expressed in 대장균에서 공동발현되는 경우, 다른 세스퀴테르펜의 과발현시키는 것도 예상할 수 있다.
- [0090] IspA는 GPP 중간체를 통하여 IPP와 DMAPP로부터 *E,E*-FPP를 합성하는 진행성 효소 (processive enzyme)인 반면, Rv1086는 GPP 기질만로부터 *Z,E*-FPP를 합성한다. IspA 및 Rv1086는 *Z,E*-FPP 합성에서 GPP 기질에 대한 상호 경쟁자라는 것이 밝혀졌다. 흥미롭게도, IspA 및 Rv1086 융합은 *Z,E*-FPP 합성에 유리하게 하였으며, 이들을 별개로 발현하는 대장균 SN-ILRv에 비하여, 대장균 SN-ILRv에서 *Z,E*-FOH 생산이 12 배 증가되었다 (도2). IspA 및 Rv1086의 공간적 위치가 증가된 *Z,E*-FPP 합성에 핵심인 것처럼 보인다. 키메라 ILRv 중에서, Rv1086는 유연한 글리신-세린 링커에 의하여 IspA의 C 말단에 연결되었다. 상기 링커는 완전하게 펼쳐져 ($\Phi=180^\circ$, $\Psi=180^\circ$) 약 3.6 nm가 될 수 있으나, 자연에서는 일어날 것 같지 않지만 경계 지점으로서 역할을 할 수 있다. 이 길이는 대장균에서 평균 단백질 직경 (5nm)에 견줄만하다 (<http://redpoll.pharmacy.ualberta.ca/CCDB/>).
- [0091] ILRv 다이머 경우에서, Rv1086는 고정된 기준으로서 IspA 주변의 제한 공간에서 이동하도록만 되어 있을 것이다 (도 3D, 점선 영역). 이 움직임은 테트라머에서는 더 한정될 것이다. 따라서, 근접성 (close proximity)은 Rv1086가 GPP 기질에 대하여 IspA와 공정하게 경쟁할 수 있도록 한다.
- [0092] 대장균 SN-ILRv에서 *Z,E*-FOH 비율 80.7%는 염색체 발현된 IspA 뿐만 아니라 키메라 ILRv에 의하여, *Z,E*-/*E,E*-FPP가 4:1의 비율로 합성된다는 것을 나타낸다.
- [0093] 그래서, IspA가 Rv1086 (K_m , 38 μ M)에 비하여 더 높은 GPP 친화성 (K_m , 8.7 μ M)을 갖는 것을 고려하여 (Fujisaki et al. 1986; Schulbach et al. 2001), 본 발명자들은 Rv1086가 IspA 보다 더 높은 턴오버 수 (turnover number) (K_{cat})를 갖는 것으로 생각하였다.
- [0094] 예측치 (Back-of-envelop-estimations)는 대사산물의 국지적 농도는 생산 효소로부터의 거리에 의존한다는 것을 나타낸다. 그러므로, GPP 농도는 IspA 및 Rv1086 사이의 근처에서 (도 3D) Rv1086의 K_m 보다 더 높은 것이다. 키메라 ILRv의 Rv1086 및 IspA에 의한 *E,E*- (또는 *Z,E*-) FPP 합성은 K_m 보다는 K_{cat} 에 의존할 것이다. 결과적으로, ILRv는 *Z,E*-FPP를 합성하는 쪽으로 편향된다. 본 발명자들은 조작된 대장균에서 *E,E*-FOH 합성은 포스파타제에 의한 내재적 가수분해 때문이라는 것을 보고한 바 있다 (Wang et al. 2010). *Z,E*-FOH 형성은 무분별한 포스파타제가 *E,E*-FPP와 *Z,E*-FPP를 구분하지 않는다는 것을 나타낸다.
- [0095] 전구체 대사산물 손실의 최소화는 대사 조작 접근에서 필수조건이다. 본 발명의 결과는 *Z,E*-FPP는 대장균에서 비대사성 대사산물 (non-metabolizable metabolite)이라는 것을 나타낸다. *Z,E*-FPP로부터 세스퀴테르펜을 생산하는 것은 이로인 방법일 수 있다. 그러나, GPP 합성이 대장균에서 Rv1086를 이용한 *Z,E*-FPP 합성을 위하여 부가적으로 필요하다. GPP 신타제는 *Abies grandis*, *Arabidopsis thaliana* 및 *M. tuberculosis*과 같은 적은 개체에서만 확인되었다 (Bouvier et al. 2000; Burke and Croteau 2002; Mann et al.). 식물 GPP 신타제는 대장균에서 발현된 경우 좋지 않게 (poorly) 수행되었다 (Carter et al. 2003). GPP 합성을 위한 FPP 신타제 돌연변이 (IspA^{S80F})는 효소 활성의 감소가 초래되었다 (도 2) (Reiling et al. 2004). 여기서, IspA 및 Rv1086의 융합 단백질은 GPP 신타제의 사용함이 없는, *Z,E*-FPP 합성의 매력적 대안 방법을 또한 제공한다. 대장균 SN-ILRv는 적절한 세스퀴테르펜 신타제가 도입되는 경우, 세스퀴테르펜 생산을 위한 좋은 플랫폼 균주일 수 있다.
- [0096] 인용문헌:
- [0097] Albertsen L, Chen Y, Bach LS, Rattleff S, Maury J, Brix S, Nielsen J, Mortensen UH. Appl Environ Microbiol 77(3):1033-40.
- [0098] Ashby MN, Edwards PA. 1990. J Biol Chem 265(22):13157-64.
- [0099] Bloch K, Chaykin S, Phillips AH, De Waard A. 1959. J Biol Chem 234:2595-604.
- [0100] Bouvier F, Suire C, d'Harlingue A, Backhaus RA, Camara B. 2000. Plant J 24(2):241-52.
- [0101] Bugg TD, Brandish PE. 1994. FEMS Microbiol Lett 119(3):255-62.

- [0102] Burke C, Croteau R. 2002. Arch Biochem Biophys 405(1):130-6.
- [0103] Carter OA, Peters RJ, Croteau R. 2003. Phytochemistry 64(2):425-33.
- [0104] Clarke S. 1992. Annu Rev Biochem 61:355-86.
- [0105] Coates RM, Denissen JF. 1988. J. Org. Chem. 52:2186-2192.
- [0106] Faraldos JA, O'Maille PE, Dellas N, Noel JP, Coates RM. J Am Chem Soc 132(12):4281-9.
- [0107] Fujisaki S, Hara H, Nishimura Y, Horiuchi K, Nishino T. 1990. J Biochem 108(6):995-1000.
- [0108] Fujisaki S, Nishino T, Katsuki H. 1986. J Biochem 99(5):1327-37.
- [0109] Fujisaki S, Takahashi I, Hara H, Horiuchi K, Nishino T, Nishimura Y. 2005. J Biochem 137(3):395-400.
- [0110] Heinsteins PF, Herman DL, Tove SB, Smith FH. 1970. J Biol Chem 245(18):4658-65.
- [0111] Hosfield DJ, Zhang Y, Dougan DR, Broun A, Tari LW, Swanson RV, Finn J. 2004. J Biol Chem 279(10):8526-9.
- [0112] Koyama T, Gotoh Y, Nishino T. 2000. Biochemistry 39(2):463-9.
- [0113] Lange BM, Rujan T, Martin W, Croteau R. 2000. Proc Natl Acad Sci U S A 97(24):13172-7.
- [0114] Mann FM, Thomas JA, Peters RJ. FEBS Lett 585(3):549-54.
- [0115] Martin VJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. 2003. Nat Biotechnol 21(7):796-802.
- [0116] Matsuoka S, Sagami H, Kurisaki A, Ogura K. 1991. J Biol Chem 266(6):3464-8.
- [0117] Narita K, Ohnuma S, Nishino T. 1999. J Biochem 126(3):566-71.
- [0118] Reiling KK, Yoshikuni Y, Martin VJ, Newman J, Bohlmann J, Keasling JD. 2004. Biotechnol Bioeng 87(2):200-12.
- [0119] Rohdich F, Hecht S, Gartner K, Adam P, Krieger C, Amslinger S, Arigoni D, Bacher A, Eisenreich W. 2002. Proc Natl Acad Sci U S A 99(3):1158-63.
- [0120] Sallaud C, Rontein D, Onillon S, Jabes F, Duffe P, Giacalone C, Thoraval S, Escoffier C, Herbette G, Leonhardt N and others. 2009. Plant Cell 21(1):301-17.
- [0121] Sambrook J, Russell D, editors. 2001. Molecular cloning: a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press.
- [0122] Schulbach MC, Brennan PJ, Crick DC. 2000. J Biol Chem 275(30):22876-81.
- [0123] Schulbach MC, Mahapatra S, Macchia M, Barontini S, Papi C, Minutolo F, Bertini S, Brennan PJ, Crick DC. 2001. J Biol Chem 276(15):11624-30.
- [0124] Thulasiram HV, Erickson HK, Poulter CD. 2007. Science 316(5821):73-6.
- [0125] Thulasiram HV, Poulter CD. 2006. J Am Chem Soc 128(49):15819-23.
- [0126] Wang C, Kim J-Y, Choi E-S, Kim S-W. 2011a. Process Biochem. 46(6):1221-1229.
- [0127] Wang C, Yoon SH, Jang HJ, Chung YR, Kim JY, Choi ES, Kim SW. 2011b. Metab Eng:In Press.
- [0128] Wang C, Yoon SH, Shah AA, Chung YR, Kim JY, Choi ES, Keasling JD, Kim SW. 2010. Biotechnol Bioeng 107(3):421-9.
- [0129] Wang W, Dong C, McNeil M, Kaur D, Mahapatra S, Crick DC, Naismith JH. 2008. J Mol Biol 381(1):129-40.
- [0130] Yoon SH, Lee SH, Das A, Ryu HK, Jang HJ, Kim JY, Oh DK, Keasling JD, Kim SW. 2009. J Biotechnol 140(3-4):218-26.
- [0131] Yoon SH, Lee YM, Kim JE, Lee SH, Lee JH, Kim JY, Jung KH, Shin YC, Keasling JD, Kim SW. 2006.

Biotechnol Bioeng 94(6):1025-32.

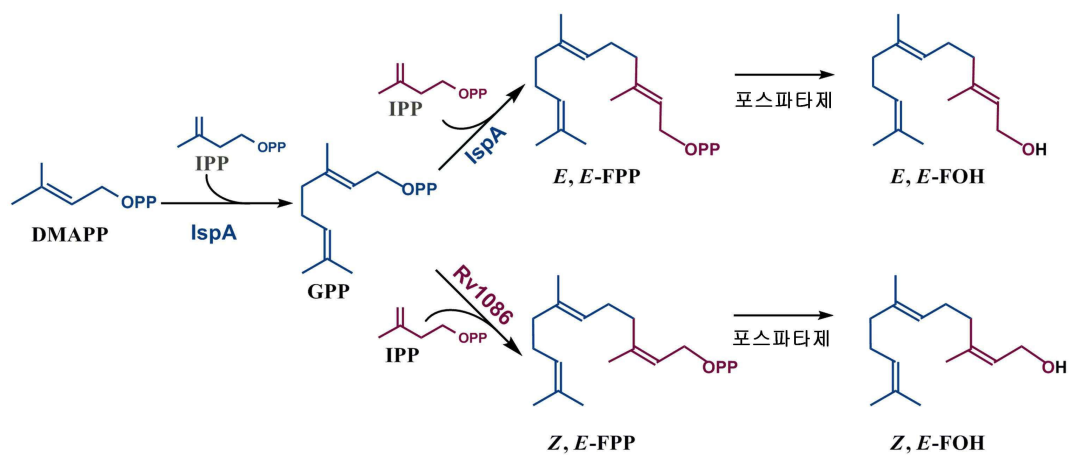
[0132] Yoon SH, Park HM, Kim JE, Lee SH, Choi MS, Kim JY, Oh DK, Keasling JD, Kim SW. 2007. Biotechnol Prog 23(3):599-605.

[0133] Yu JS, Kleckley TS, Wiemer DF. 2005. Org Lett 7(22):4803-6.

[0134] Zhang D, Jennings SM, Robinson GW, Poulter CD. 1993. Arch Biochem Biophys 304(1):133-43.

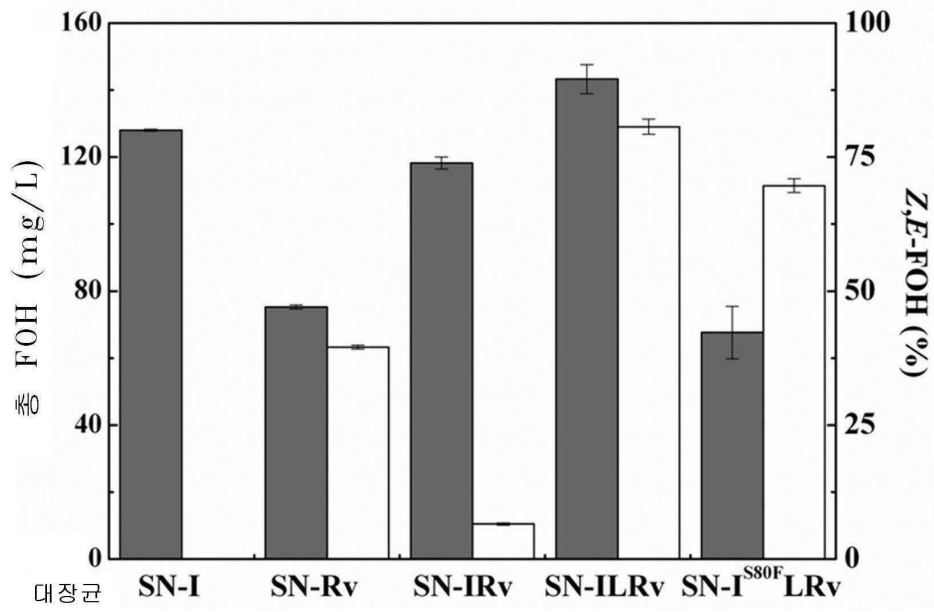
도면

도면1

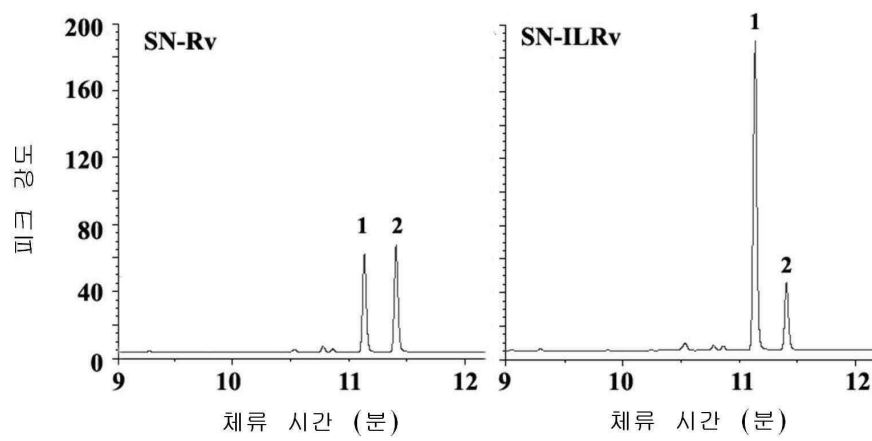


도면2

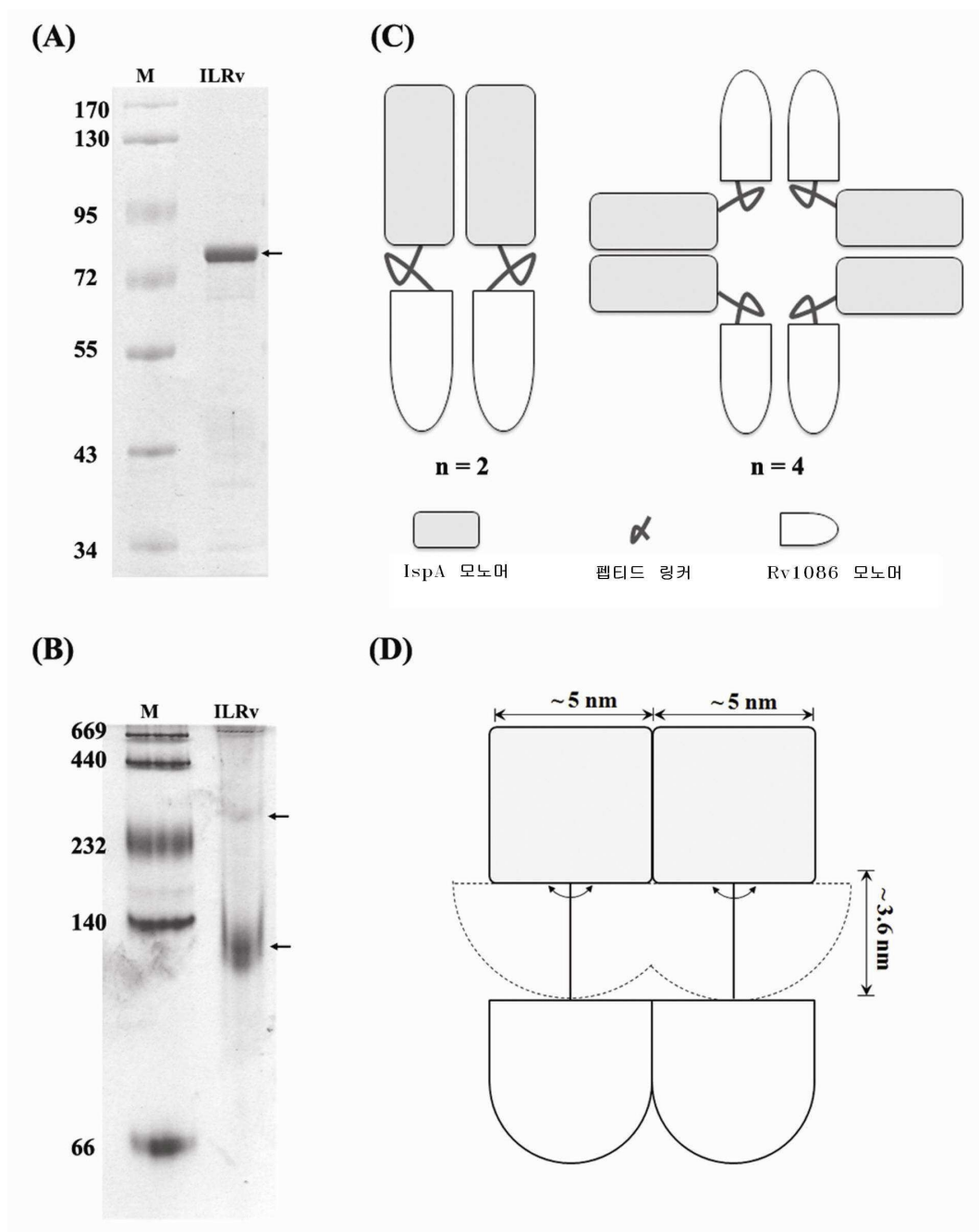
(A)



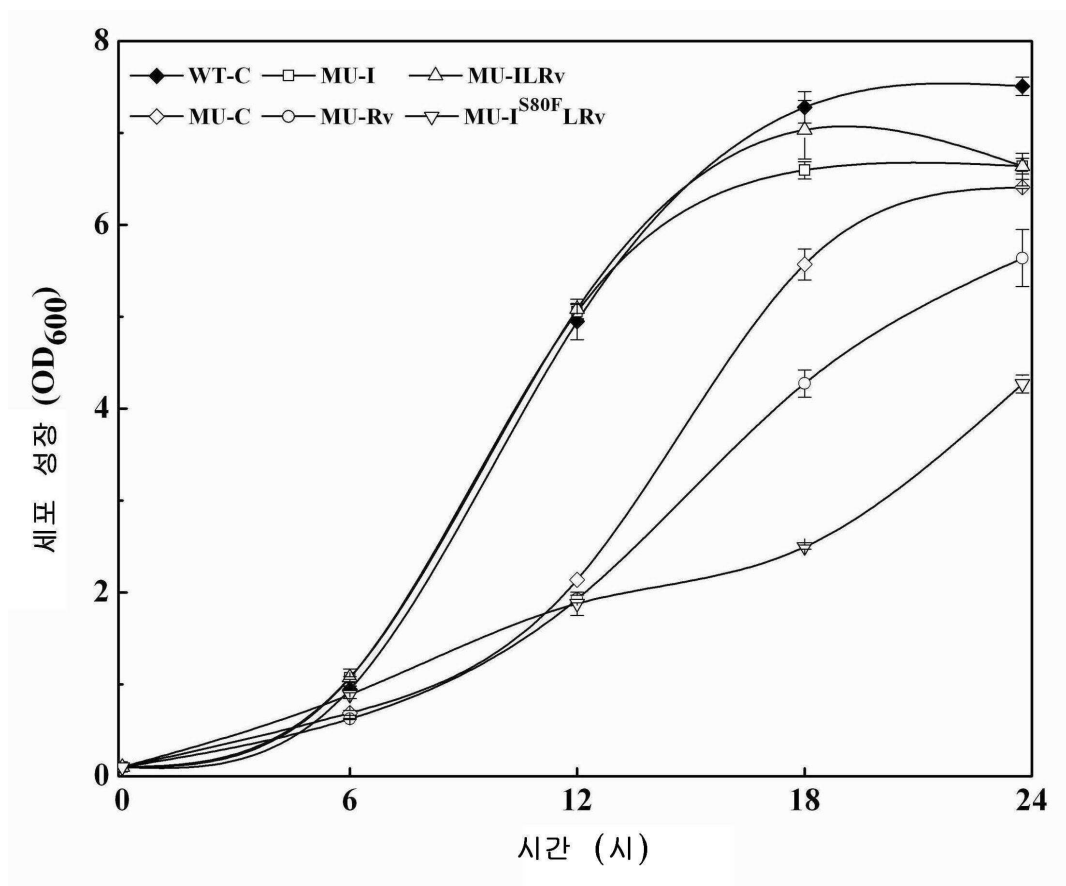
(B)



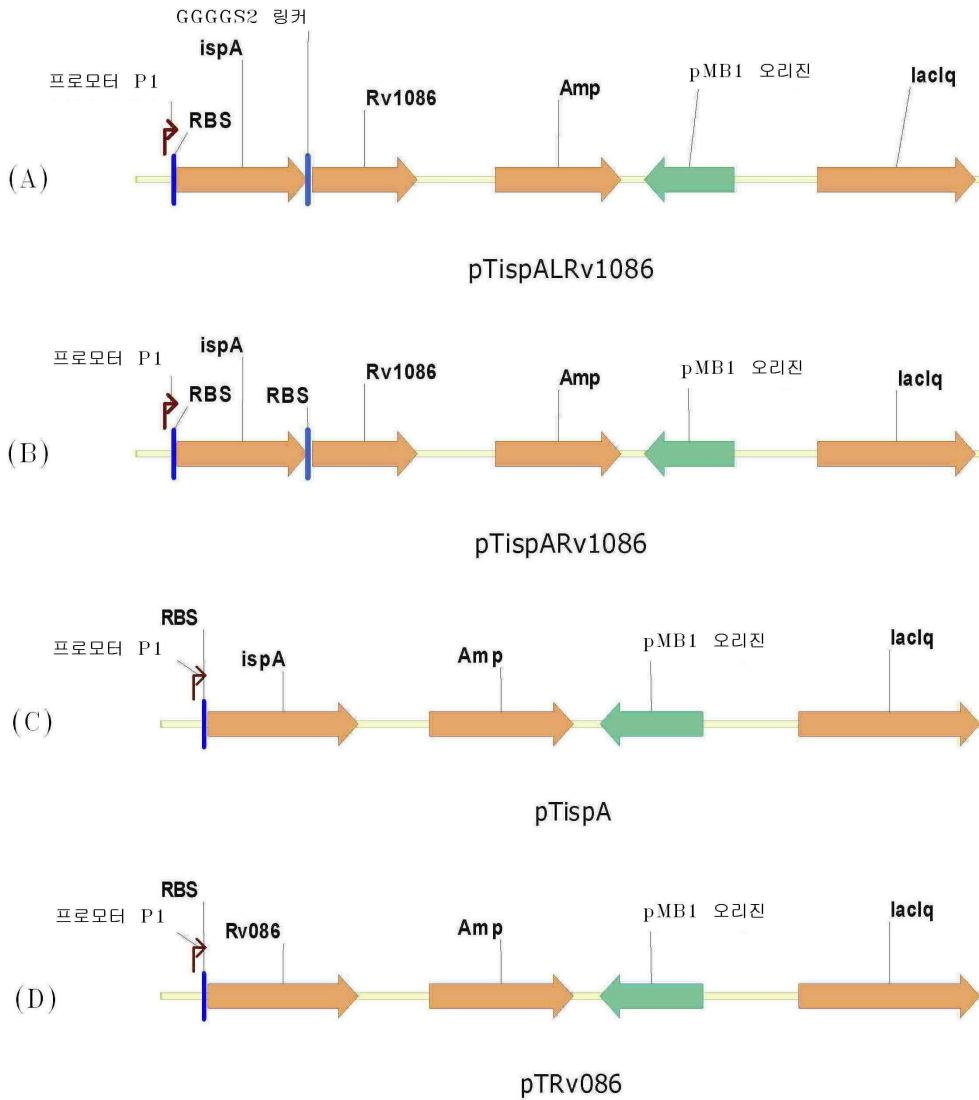
도면3



도면4



도면5



서열 목록

- <110> INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYEONGSANG NATIONAL UNIVERSITY
- <120> Fusion enzyme between Z,E-farnesyl diphosphate synthase and
farnesyl diphosphate synthase from E.coli and use thereof
- <130> PN095164
- <160> 35
- <170> KopatentIn 2.0
- <210> 1
- <211> 239
- <212> PRT
- <213> Mycobacterium tuberculosis H37Rv
- <400> 1

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

<210> 2

<211> 299

<212> PRT

<213> Escherichia 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 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> 548

<212> PRT

<213> Artificial Sequence

<220><223> Fusion enzyme between Mycobacterium tuberculosis H37Rv Rv1086 and
 Escherichia coli ispA

<400> 3

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 Gly Gly Gly Gly Ser			
290	295	300	
Gly Gly Gly Gly Ser Met Ala Ala Ser Lys Ser Asp Leu Pro Arg His			
305	310	315	320
Ile Ala Val Leu Cys Asp Gly Asn Arg Arg Trp Ala Arg Ser Ala Gly			
325	330	335	
Tyr Asp Asp Val Ser Tyr Gly Tyr Arg Met Gly Ala Ala Lys Ile Ala			
340	345	350	
Glu Met Leu Arg Trp Cys His Glu Ala Gly Ile Glu Leu Ala Thr Val			
355	360	365	

Tyr Leu Leu Ser Thr Glu Asn Leu Gln Arg Asp Pro Asp Glu Leu Ala

370

375

380

Ala Leu Ile Glu Ile Ile Thr Asp Val Val Glu Glu Ile Cys Ala Pro

385

390

395

400

Ala Asn His Trp Ser Val Arg Thr Val Gly Asp Leu Gly Leu Ile Gly

405

410

415

Glu Glu Pro Ala Arg Arg Leu Arg Gly Ala Val Glu Ser Thr Pro Glu

420

425

430

Val Ala Ser Phe His Val Asn Val Ala Val Gly Tyr Gly Gly Arg Arg

435

440

445

Glu Ile Val Asp Ala Val Arg Ala Leu Leu Ser Lys Glu Leu Ala Asn

450

455

460

Gly Ala Thr Ala Glu Glu Leu Val Asp Ala Val Thr Val Glu Gly Ile

465

470

475

480

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

485

490

495

Arg Thr Ser Gly Glu Gln Arg Leu Ser Gly Phe Leu Leu Trp Gln Ser

500

505

510

Ala Tyr Ser Glu Met Trp Phe Thr Glu Ala His Trp Pro Ala Phe Arg

515

520

525

His Val Asp Phe Leu Arg Ala Leu Arg Asp Tyr Ser Ala Arg His Arg

530

535

540

Ser Tyr Gly Arg

545

<210> 4

<211> 720

<212> DNA

<213> Mycobacterium tuberculosis H37Rv

<400> 4

atggccgcct cgaaatccga cctgccccgg cacatagccg tgctgtgcga cggcaaccgg 60

cgatgggagc gcagcgcggg ctacgacgac gtcagctacg gctaccggat gggtgcggcc 120

aagatcgccg aaatgctgcg gtggtgccac gaagccggca tcgaactggc caccgtctat	180
ctgctgtcca ccgaaaacct gcagcgcgat cccgacgagc ttgcagcact catcgagatc	240
atcaccgatg tcgtggaaga gatctgcgca ccggccaacc actggagtgt gcggacggtc	300
ggggatctgg ggttgatcgg cgaggaaccg gcccggcggc tgcgcggtgc ggtggaatcc	360
accccgaggg tggcctcggt tcatgtcaac gttagctgtt gctacggcgg gcgccgcgag	420
atcgtcgacg ctgtgcgcgc gttgttgagc aaggaactcg ccaacggggc caccgcggag	480
gaactcgtcg acgcggtgac cgtcgagggt atctcggaac acctgtacac ctcaggccaa	540
cccgaccccg atttggtgat acgcacctcc ggcgagcaac gcttgtccgg gttcttgctg	600
tggaagcg cctactcgga gatgtggttc accgagggcg actggccggc gtttcgccac	660
gtcgattttc tacgcgcgct gcgtgactac agtgcgaggc atcgagcta cggcaggtaa	720
	720
<210> 5	
<211> 900	
<212> DNA	
<213> Escherichia coli	
<400> 5	
atggactttc cgcagcaact cgaagcctgc gttaaagcagg ccaaccaggc gctgagccgt	60
tttatcgccc cactgccctt tcagaacact cccgtggtcg aaaccatgca gtatggcgca	120
ttattagggtg gtaagcgctt gcgacctttc ctggtttatg ccaccggtca tatgttcggc	180
gtagcacaa acacgtgga cgcacccgct gccgcccgtt agtgtatcca cgcttactca	240
ttaattcatg atgatttacc ggcaatggat gatgacgatc tgcgtcgcgg ttgccaacc	300
tgccatgtga agtttgcgca agcaaacgcg attctcgtg gcgacgcttt acaaacgtg	360
gcgttctcga ttttaagcga tgccgatatg ccggaagtgt cggaccgcga cagaatttcg	420
atgattttctg aactggcgag cgccagtggg attgccgga tgtgcggtgg tcaggcatta	480
gatttagacg cggaaggcaa acacgtacct ctggacgcgc ttgagcgtat tcacgtcat	540
aaaaccggcg cattgatcgc cgccgccgtt cgccttggtg cattaagcgc cggagataaa	600
ggacgtcgtg ctctgccggt actcgacaag tatgcagaga gcatcgccct tgccttcag	660
gttcaggatg acatcctgga tgtggtggga gatactgcaa cgttgggaaa acgccagggt	720
gccgaccagc aacttggtaa aagtacctac cctgcacttc tgggtcttga gcaagcccgg	780
aagaaagccc gggatctgat cgacgatgcc cgtcagtcgc tgaacaact ggctgaacag	840
tcactcgata cctcggcact ggaagcgcta gcggactaca tcacccagcg taataataa	900

900

<210> 6

<211> 1647

<212> DNA

<213> Artificial Sequence

<220><223> Fusion enzyme IspALRV1086 between Mycobacterium tuberculosis H37Rv Rv1086 and Escherichia coli ispA

<400> 6

atggactttc cgcagcaact cgaagcctgc gttaagcagg ccaaccaggc gctgagccgt	60
tttatcgccc cactgcccctt tcagaacact cccgtggctc aaaccatgca gtatggcgca	120
ttattaggtg gtaagcgcct gcgaccttc ctggtttatg ccaccgtca tatgttcggc	180
gttagcacia acacgtctga cgcaccgct gccgccgttg agtgtatcca cgcttactca	240
ttaattcatg atgatttacc ggcaatggat gatgacgatc tgcgtcgcgg ttgccaacc	300
tgccatgtga agtttggcga agcaaagcgc attctcgtg gcgacgttt acaaagcgtg	360
gcgttctcga ttttaagcga tgccgatatg ccggaagtgt cggaccgca cagaatttcg	420
atgatttctg aactggcgag cgccagtggg attgccgaa tgtgcgtgg tcaggcatta	480
gatttagacg cggaaggcaa acacgtacct ctggacgcgc ttgagcgtat tcatcgtcat	540
aaaaccggcg cattgattcg cgccgccgtt cgccttgggt cattaagcgc cggagataaa	600
ggacgtcgtg ctctgccgtt actcgacaag tatgcagaga gcatcgccct tgccttcag	660
gttcaggatg acatcctgga tgtgttgga gatactgcaa cgttgggaaa acgccagggt	720
gccgaccagc aacttggtaa aagtacctac cctgcacttc tgggtcttga gcaagcccgg	780
aagaaagccc gggatctgat cgacgatgcc cgtcagtcgc tgaacaact ggctgaacag	840
tactcgata cctcgccact ggaagcgcta gccgactaca tcatccagcg taataaaggt	900
ggcgttggct cgggtggcgg cggatccatg gccgcctcga aatccgacct gccccggcac	960
atagccgtgc tgtgcgacgg caaccggcga tgggcgcgca gcgcgggcta cgacgacgtc	1020
agctacggct accggatggg tgcggccaag atcgccgaaa tgcgtcgggt gtgccacgaa	1080
gccggcatcg aactggccac cgtctatctg ctgtccaccg aaaacctgca gcgcgatccc	1140
gacgagcttg cagcactcat cgagatcatc accgatgtcg tggaagagat ctgcgcaccg	1200
gccaaccact ggagtgtgcg gacggtcggg gatctggggg tcatcgcgca ggaaccggcc	1260
cggcggctgc gcggtgcggt ggaatccacc ccggaggtgg cctcgtttca tgtcaacgtt	1320
gctgttggct acggcgggcg ccgcgagatc gtcgacgtg tgcgcgcgtt gttgagcaag	1380

gaactcgcca acggggccac cgcggaggaa ctgctcgacg cggtgaccgt cgagggtatc 1440
tcggaaaacc tgtacacctc aggccaaccc gaccccgatt tggtgatacg cacctccggc 1500
gagcaacgct tgtccgggtt cttgctgtgg caaagcgct actcggagat gtggttcacc 1560

gaggcgcact ggccggcggtt tcgccacgtc gattttctac gcgcgctgcg tgactacagt 1620
gcgaggcatc gcagctacgg caggtaa 1647

<210> 7

<211> 10

<212> PRT

<213> Artificial Sequence

<220><223> linker sequence

<400> 7

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 8

<211> 292

<212> PRT

<213> Streptococcus pneumoniae

<400> 8

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> 9

<211> 336

<212> PRT

<213> Streptococcus pneumoniae

<400> 9

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> 10
 <211
 > 317
 <212> PRT
 <213> Streptococcus pneumoniae
 <400> 10
 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> 11

<211> 182

<212> PRT

<213> E. coli K12 MG1655

<400> 11

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> 12

<211> 803

<212> PRT

<213> Enterococcus faecalis

<400> 12

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> 13

<211> 383

<212> PRT

<213> Enterococcus faecalis

<400> 13

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
165	170	175	
Met Leu Thr Gln	Asp Ile Tyr Asp	Phe Trp Arg	Pro Thr Gly
180	185	190	
Tyr Pro Met Val	Asp Gly Pro Leu	Ser Asn Glu	Thr Tyr Ile
195	200	205	
Phe Ala Gln Val	Trp Asp Glu His	Lys Lys Arg	Thr Gly Leu
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> 14

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> primer Rv1086-F1

<400> 14

taggatccat ggagatcatc ccgccgcg

28

<210> 15

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> primer Rv1086-F2

<400> 15

tggatccagg aggtaataaa ccatggagat catcccgccg cg 42

<

210> 16

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> primer Rv1086-R

<400> 16

atctcgagtt acctgccgta gctgcgatgc 30

<210> 17

<211> 4176

<212> DNA

<213> Artificial Sequence

<220><223> pTrc99A vector

<400> 17

gtttgacagc ttatcatcga ctgcacggtg caccaatgct tctggcgta ggcagccatc 60

ggaagctgtg gtagggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc 120

gcactcccggt tctggataat gttttttgcg ccgacatcat aacggttctg gcaaattatc 180

tgaaatgagc tgttgacaat taatcatccg gctcgtataa tgtgtggaat tgtgagcgga 240

taacaatttc acacaggaaa cagaccatgg aattcgagct cggtacccgg ggatcctcta 300

gagtcgacct gcaggcacgc aagcttggtt gttttggcgg atgagagaag attttcagcc 360

tgatacacat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca 420

gtagcgcggt ggcccacct gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg 480

atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaaacga 540

aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggg gaacgctctc 600

ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccggaggg 660

tggcgggcag gacgcccgcc ataaactgcc aggcacataa ttaagcagaa ggccatcctg 720

acggatggcc tttttgcgtt tctacaaact ctttttgttt atttttctaa atacattcaa 780

atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 840

agagtatgag tattcaacat ttccgtgtcg cccttattcc cttttttgcg gcattttgcc 900

ttctgtttt 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
acttgggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	1200
aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa	1260
cgatcggagg accgaaggag ctaaccgctt ttttgacaa catgggggat catgtaactc	1320
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	1380
cgatgcctac agcaatggca acaacgttgc gcaaaactatt aactggcgaa ctacttactc	1440
tagcttcccc gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	1500
tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg	1560
ggtctcgcg tttcatttga gcaactgggc cagatggtaa gccctcccgt atcgtagtta	1620
tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag	1680
gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga	1740
ttgatttaaa acttcatitt taatttaaaa ggatctaggt gaagatcctt tttgataatc	1800
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa	1860
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa	1920
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc	1980
cgaaggtaac tggtctcagc agagcgagca taccaaatat tgtccttcta gtgtagccgt	2040
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaattc	2100
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	2160
gatatgtacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca	2220
gcttggagcg aacgacctac accgaactga gatactaca gcgtgagcta tgagaaagcg	2280
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	2340
gagagcgcac gagggagctt ccagggggaa acgectggta tctttatagt cctgtcgggt	2400
ttgccacct ctgacttagc cgtcgatttt tgtgatgctc gtcagggggg cggagcctat	2460
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc	2520
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt	2580
gagctgatac cgctcgccgc agccgaacga ccgagcgag cgagtcagtg agcgaggaag	2640
cggagagcgc cctgatgcgg tatttttctc ttacgcatct gtgcgggtatt tcacaccgca	2700
tatggtgcac tctcagtaca atctgctctg atgccgata gttaagccag tatacactcc	2760
gctatcgcta cgtgactggg tcatggctgc gccccgacac ccgccaacac ccgctgacgc	2820
gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg	2880

gagctgcatg tgtcagaggt ttccaccgtc atcaccgaaa cgcgcgaggc agcagatcaa 2940
 ttccgcgcgcg aaggcgaagc ggcatgcatt tacgttgaca ccatcgaatg gtgcaaaacc 3000
 tttcgcggta tggcatgata gcgcccggaa gagagtcaat tcagggtggt gaatgtgaaa 3060
 ccagtaacgt tatacgaatg cgcagagtat gccggtgtct cttatcagac cgtttcccgc 3120

gtggtgaacc aggccagcca cgtttctgcg aaaacgcggg aaaaagtga agcggcgatg 3180
 gcggagctga attacattcc caaccgcgtg gcacaacaac tggcgggcaa acagtcgttg 3240
 ctgattggcg ttgccacctc cagtcgtggc ctgcacgcgc cgtcgcaaat tgtcgcggcg 3300
 attaatctc gcgccgatca actgggtgcc agcgtggtgg tgtcgatggt agaacgaagc 3360
 ggctgcgaag cctgtaaagc ggcggtgcac aatcttctcg cgcaacgcgt cagtgggctg 3420
 atcattaact atccgctgga tgaccaggat gccattgtg tggaaagtgc ctgcactaat 3480
 gttccggcgt tatttcttga tgtctctgac cagacacca tcaacagtat tatttctcc 3540

catgaagacg gtacgcgact gggcgtggag catctggtcg cattgggtca ccagcaaadc 3600
 gcgctgttag cgggccatt aagtctgtc tcggcgcgtc tgcgtctggc tggctggcat 3660
 aaatatctca ctgcgaatca aattcagccg atagcggaac gggaaggcga ctggagtgcc 3720
 atgtccggtt ttcaacaaac catgcaaatg ctgaatgagg gcatcgttcc cactgcgatg 3780
 ctggttgcca acgatcagat ggcgctgggc gcaatgcgcg ccattaccga gtccgggctg 3840
 cgcgttggcg cgatatctc ggtagtggga tacgacgata ccgaagacag ctcatgttat 3900
 atcccgccgt caaccacat caaacaggat ttccgcctgc tggggcaaac cagcgtggac 3960

cgcttgctgc aactctctca gggccagcgc gtgaaggga atcagctgtt gccgtctca 4020
 ctggtgaaaa gaaaaaccac cctggcgccc aatacgcaaa ccgctctcc ccgcgcgttg 4080
 gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgcgcg 4140
 caacgaatt aatgtgagtt agcgcgaatt gatctg 4176

<210> 18
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220><223> primer ispAS80F-F
 <400> 18
 gagtgatcc acgcttactt tttaattcat gatgatttac cg 42

<210> 19
 <211> 33

<212> DNA

<213> Artificial Sequence

<220><223> primer ispAS80F-R

<400> 19

aaagtaagcg tggatacact caacggcggc agc 33

<210> 20

<211> 5369

<212> DNA

<213> Artificial Sequence

<220><223> pET28a vector sequence

<400> 20

tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60

cagcgtgacc gtacacttg ccagcgccct agcgcccgt ctttcgctt tcttccttc 120

ctttctgcc acgttcgccg gctttcccg tcaagctcta aatcgggggc tcccttagg 180

gttccgattt agtgctttac ggcacctga cccccaaaaa cttgattagg gtgatggttc 240

acgiagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300

ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatct cggctctatc 360

ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420

acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480

tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540

tccgctcatg aattaattct tagaaaaact catcgagcat caaatgaaac tgcaatttat 600

tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat gaaggagaaa 660

actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc 720

gtccaacatc aataaacct attaatctcc cctcgtcaaa aataaggtta tcaagtgaga 780

aatcaccatg agtgacgact gaatccgtg agaatggcaa aagtttatgc atttctttcc 840

agacttgttc aacaggccag ccattacgt cgtcatcaaa atcactcgca tcaaccaaac 900

cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgtg ttaaaaggac 960

aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat 1020

tttcacctga atcaggatat ttttctaata cctggaatgc tgttttcccg gggatcgag 1080

tggtagtaaa ccatgcata tcaggagtag ggataaaatg cttgatggtc ggaagaggca 1140

taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgtac 1200

ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg 1260

tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca	1320
tgttgggaatt taatcgcggc ctagagcaag acgtttcccg ttgaatatgg ctcataaacac	1380
cccttgtatt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa	1440
cgtagagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga	1500
gatactttttt ttctgcgcgt aatctgctgc ttgcaaaaa aaaaaccacc gctaccagcg	1560
gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc	1620
agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag	1680
aactctgtag caccgcctac atacctcgct ctgctaatac tgttaccagt ggctgctgcc	1740
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg	1800
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac	1860
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga	1920
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt	1980
ccagggggaa acgcctggta tctttatagt cctgtcgggt ttgccacct ctgacttgag	2040
cgtcgattttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg	2100
gcctttttac ggttcttggc cttttgctgg ctttttctc acatgttctt tcttgcgtta	2160
tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc	2220
agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg	2280
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatatggcgc actctcagta	2340
caatctgctc tgatgccga tagttaagcc agtatacact ccgctatcgc tacgtgactg	2400
ggatcatggt ggcggcgac acccgccaac acccgctgac gcggcctgac gggcttgtct	2460
gctcccgga tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag	2520
gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc	2580
gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag	2640
aagcgttaat gcttgcttc tgataaagcg ggccatgtta agggcggttt tttctgttt	2700
ggtcactgat gcctccgtgt aagggggatt tctgttcatg ggggtaatga taccgatgaa	2760
acgagagagg atgctcacga tacgggttac tgatgatgaa catgcccggt tactggaacg	2820
ttgtgagggt aaacaactgg cggtatggtat gcggcgggac cagagaaaaa tcaactcagg	2880
tcaatgccag cgcttcgtta atacagatgt aggtgttcca cagggtagcc agcagcatcc	2940
tgcgatgcag atccggaaca taatggtgca gggcgctgac ttccgcgttt ccagacttta	3000
cgaaacacgg aaaccgaaga ccattcatgt tgttgctcag gtcgcagacg ttttgcagca	3060

gcagtcgctt cacgttcgct cgcgtatcgg tgattcattc tgctaaccag taaggcaacc	3120
ccgccagcct agccgggtcc tcaacgacag gagcacgac atgcgcaccc gtggggccgc	3180
catgccggcg ataatggcct gcttctcgcc gaaacgtttg gtggcgggac cagtgcacgaa	3240
ggcttgagcg agggcgtgca agattccgaa taccgcaagc gacaggccga tcatcgtcgc	3300
gctccagcga aagcggctct cgccgaaat gaccagagc gctgccggca cctgtcctac	3360
gagttgcatg ataaagaaga cagtcataag tgcggcgacg atagtcatgc cccgcgcca	3420
ccggaaggag ctgactgggt tgaaggctct caagggcac gtctgagatc ccggtgccta	3480
atgagtgcgc taacttacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa	3540
cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttcgtat	3600
tgggcgccag ggtggttttt cttttacca gtgagacggg caacagctga ttgcccttca	3660
ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc agcaggcgaa	3720
aatcctgttt gatgggtggtt aacggcggga tataacatga gctgtcttcg gtatcgtcgt	3780
atcccactac cgagatatcc gcaccaacgc gcagcccga ctcggtaatg gcgcgcattg	3840
cgcccagcgc catctgatcg ttggcaacca gcatcgagc gggaacgatg ccttcattca	3900
gcatttgcgc ggtttgttga aaaccggaca tggcactcca gtcgccttcc cgttccgcta	3960
tggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc agacgcgccg	4020
agacagaact taatgggccc gctaacagcg cgatttgctg gtgaccaat gcgaccagat	4080
gctccacgcc cagtcgcgta ccgtcttcat gggagaaaaat aatactgttg atgggtgtct	4140
ggtcagagac atcaagaaat aacgccgga cattagtga ggcagcttcc acagcaatgg	4200
catcctgttc atccagcgga tagttaatga tcagcccact gacgcgttgc gcgagaagat	4260
tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac accaccacgc	4320
tggcaccag ttgatcggcg cgagatttaa tcgccgcgac aatttgcgac ggcgcggtga	4380
gggccagact ggaggtggca acgccaatca gcaacgactg ttgcccgc agttgttgtg	4440
ccacgcggtt gggaatgtaa ttcagctccg ccacgcgcg ttcactttt tcccgcgttt	4500
tcgcagaaac gtggctggcc tggttacca cgcgggaaac ggtctgataa gagacaccgg	4560
catactctgc gacatcgat aacgttactg gtttcacatt caccacctg aattgactct	4620
cttccgggcg ctatcatgcc ataccgcgaa aggttttgcg ccattcgatg gtgtccggga	4680
tctcgacgct ctcccttatg cgactcctgc attaggaagc agcccagtag taggttgagg	4740
ccgttgagca ccgccgcgc aaggaatggt gcatgcaagg agatggcgcc caacagtccc	4800

ccggccacgg ggccctgccac cataccacag ccgaacaag cgtcatgag cccgaagtgg	4860
cgagcccgat ctccccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg	4920
gcgccggtga tgccggccac gatgcgtccg gcgtagagga tcgagatctc gatcccgca	4980
aattaatacg actcactata ggggaattgt gagcggataa caattccct ctagaaataa	5040
ttttgtttaa cttaagaag gagatatacc atgggcagca gccatcatca tcatcatcac	5100
agcagcggcc tgggtcccg cggcagccat atggctagca tgactggtgg acagcaaag	5160
ggtcgcggtat ccgaattcga gtcctgtcga caagcttgcg gccgcactcg agcaccacca	5220
ccaccaccac tgagatccgg ctgctaaca agcccgaag gaagctgagt tggctgctgc	5280
caccgtgtag caataactag cataaccct tggggcctct aaacgggtct tgagggttt	5340
tttctgaaa ggaggaacta tatccgat	5369
<210> 21	
<211> 2999	
<212> DNA	
<213> Artificial Sequence	
<220><223> pSTV vector sequence	
<400> 21	
cgtatggcaa tgaaagacgg tgagctgtg atatgggata gtgttcaccc ttgttacacc	60
gttttccatg agcaaacga aacgttttca tcgtctgga gtgaatacca cgacatttc	120
cggcagtttc tacacatata ttgcaagat gtggcgtgtt acggtgaaa cctggcctat	180
ttccctaaag ggtttattga gaatatgtt ttctctcag ccaatccctg ggtgagtttc	240
accagttttg atttaaacgt ggccaatatg gacaacttct tcgccccctg tttcacatg	300
ggcaaatatt atacgaagg cgacaagggt ctgatgccg tggcgattca ggttcatcat	360
gccgtctgtg atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgat	420
gagtggcagg gcggggcgta attttttaa ggcagttatt ggtgccctta aacgcctggt	480
gctacgcctg aataagtgat aataagcggg tgaatggcag aaattcgaaa gcaaattcga	540
cccggctgtc ggttcagggc aggtctgtta aatagccgt tatgtctatt gctggtttac	600
cggtttatg actaccggaa gcagtgtgac cgtgtgcttc tcaaatgcct gaggccagtt	660
tgtcaggct ctccccgtgg aggtataat tgacgatatg atcatttatt ctgcctccca	720
gagcctgata aaaacggtta gcgttctgtt aatacatatg taggtgttcc acagggtagc	780
cagcagcatc ctgcgatgca gatccggaac ataatggtgc agggcgcttg tttcggcgtg	840
ggtatggtgg caggccccgt ggccggggga ctgttggcg ctgccggcac ctgtcctacg	900

agttgcatga taaagaagac agtcataagt gcggcgacga tagtcatgcc ccgcgcccac	960
cggaaggagc taccggacag cggtagcgac tgttgtaact cagaataaga aatgaggccg	1020
ctcatggcgt tccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca	1080
gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga	1140
gttagctcac tcattaggca cccaggctt tacactttat gcttccgct cgtatgttgt	1200
gtggaattgt gagcgataa caatttcaca caggaaacag ctatgacat gattacgaat	1260
tgcagctcgg taccgggga tctctagag tcgacctgca ggcatgcaag cttggcactg	1320
gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttaccaact taatcgctt	1380
gcagcacatc cccctttcgc cagctggcgt aatagcgaag agggccgcac cgatcgccct	1440
tccaacagt tgcgcagcct gaatggcgaa tgagcttatt gatgataagc tgtcaaact	1500
gagaattaca acttatatcg tatgggctg acttcaggtg ctacatttga agagataaat	1560
tgactgaaa tctagaaata ttttatctga ttaataagat gatcttcttg agatcgtttt	1620
ggtctgcgcg taatctcttg ctctgaaaac gaaaaaacg ccttcaggg cggtttttcg	1680
aaggttctct gagctacaa ctctttgac cgaggtaact ggcttgagg agcgagtc	1740
ccaaaacttg tcctttcagt ttagccttaa ccggcgcatg acttcaagac taactcctct	1800
aatcaatta ccagtggctg ctgccagtgg tgcttttgca tgcctttccg ggttggactc	1860
aagacgatag ttaccggata aggcgcagcg gtcggactga acggggggtt cgtgcataca	1920
gtccagcttg gagcgaactg cctaccgga actgagtgtc aggcgtggaa tgagacaaac	1980
gcggccataa cagcggaatg acaccggtaa accgaaaggc aggaacagga gagcgcacga	2040
gggagccgcc aggggaaac cctggatat ttagatctt gtcgggttgc gccaccactg	2100
attgagcgt cagatttctg gatgcttgtc agggggggcg agcctatgga aaaacggctt	2160
tgccgcggcc ctctcacttc cctgttaagt atcttcttg catcttcag gaaatctccg	2220
ccccgttcgt aagccatttc cgctcgccgc agtcgaacga ccgagcgtag cgagtcagt	2280
agcgaggaag cggaatatat cctgtatcac atattctgct gacgcaccgg tgcagccttt	2340
tttctctgc cacatgaagc acttactga caccctcatc agtgccaaca tagtaagcca	2400
gtatacactc cgctagcgt gatgtccggc ggtgcttttg ccgttacgca ccacccgctc	2460
agtagctgaa caggagggac agctgataga aacagaagcc actggagcac ctcaaaaaca	2520
ccatcatata ctaaatcagt aagtggcag catcacccga cgcactttgc gccgaataaa	2580
tacctgtgac ggaagatcac ttgcgagaat aaataaatcc tgggtgcct 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> 22	
<211> 2412	
<212> DNA	
<213> Enterococcus faecalis	
<400> 22	
ttgaaaacag tagttattat tgatgcatta cgaacaccaa ttggaaaata taaaggcagc	60
ttaagtcaag taagtgccgt agacttagga acacatgtta caacacaact tttaaaaaga	120
cattccacta tttctgaaga aattgatcaa gtaatctttg gaaatgtttt acaagctgga	180
aatggccaaa atcccgacg 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 agtggtaag caatgggctt aactgctgaa	480
aatgtggccg aaaagatca 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 ttatglatcg gcggtggctt aggactcgct	1140
atgctactag agagacctca gcaaaaaaaa aacagccgat tttatcaaat gagtccctgag	1200

gaagcctgg cttctcttct taatgaagc cagatttctg ctgatacaaa aaaagaattt 1260
 gaaaatacgg ctttatcttc gcagattgcc aatcatatga ttgaaaatca aatcagtga 1320
 acagaagtgc cgatgggcgt tggcttacat ttaacagtgg acgaaactga ttatttggt 1380
 ccaatggcga cagaagagcc ctcatgtatt gcggtttga gtaatggtgc aaaaatagca 1440
 caaggattta aaacagtga tacaacacgt 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 gaagggtgtg ccgagttgtt ccgtgaatgg 1740
 tttgcggagc aaaagatttt attcagtatt ttaagtaatt atgccacgga gtcggttggt 1800
 acgatgaaaa cggtatttcc agtttcacgt ttaagtaagg ggagcaatgg ccgggaaatt 1860
 gctgaaaaaa ttgttttagc ttcacgctat gcttcattag atccttatcg ggcagtcacg 1920
 cataacaaag ggatcatgaa tggcattgaa gctgtcgttt tagctacagg aaatgataca 1980
 cgcgtgttta gcgttcttg tcatgctttt gcggtgaagg aaggtcgcta ccaaggtttg 2040

actagttgga cgctggatgg cgaacaacta attggtgaaa tttcagttcc gcttgcgtta 2100
 gccacggttg gcggtgccac aaaagtctta cctaaatctc aagcagctgc tgatttgta 2160
 gcagtacagg atgcaaaaga actaagtcga gtagtagcgg ctgttggttt ggcacaaaat 2220
 ttagcggcgt tacgggcctt agtctctgaa ggaattcaaa aaggacacat ggctctacaa 2280
 gcacgttctt tagcgtatgac ggtcggagct actggtlaag aagttgaggc agtcgctcaa 2340
 caattaaaac gtcaaaaaac gatgaacca gaccgagcct tggctatttt aaatgattta 2400
 agaaaacaat aa 2412

<210> 23

<211> 1152

<212> DNA

<213> Enterococcus faecalis

<400> 23

atgacaattg ggattgataa aattagtttt ttgtgcccc cttattatat tgatatgacg 60
 gcactggctg aagccagaaa ttagaccct ggaaaatttc atattggtat tgggcaagac 120
 caaatggcgg tgaaccaat cagccaagat attgtgacat ttgcagccaa tgccgcagaa 180
 gcgatcttga ccaaagaaga taaagaggcc attgatatgg tgattgtcgg gactgagtc 240
 agtatcatg agtcaaaagc ggccgcagtt gtcttacatc gtttaatggg gattcaacct 300

ttcgctcgct ctttcgaaat caaggaagct tgttacggag caacagcagg cttacagtta	360
gctaagaatc acgtagcctt acatccagat aaaaaagtct tggttgtagc agcagatatt	420
gcaaaatatg gattaaattc tggcggtag cctacacaag gagctggggc ggttgcaatg	480
ttagttgcta gtgaaccgag catcttggct ttaaaaggagg ataatgtgat gctgacgcaa	540
gatatctatg acttttggcg tccaacaggc catccgtatc ctatggtcga tggtcctttg	600
tcaaacgaaa cctacatcca atcttttgcc caagtctggg atgaacataa aaaaagaacc	660
ggtcttgatt ttgcagatta tgatgcttta gcgttcata ttccttacac aaaaatgggc	720
aaaaaagcct tattagcaaa aatctccgac caaactgaag cagaacagga acgaatttta	780
gcccgttatg aagaaagcat catctatagt cgtcgcgtag gaaacttgta tacgggttca	840
ctttatctgg gactcatttc ctttttagaa aatgcaacga ctttaaccgc aggcaatcaa	900
attgggttat tcagttatgg ttctgggtgt gtcgtgaat ttttactgg tgaattagta	960
gctggttatc aaaatcattt acaaaaagaa actcatttag cactgctaga taatcggaca	1020
gaactttcta tcgtgaata tgaagccatg ttgacagaaa ctttagacac agatattgat	1080
caaacgttag aagatgaatt aaaatatagt atttctgcta ttaataatac cgttcgtct	1140
tatcgaaact aa	1152
<210> 24	
<211> 879	
<212> DNA	
<213> Streptococcus pneumoniae	
<400> 24	
atgacaaaaa aagttggtgt cggtcaggca catagtaaga taattttaat aggggaacat	60
gcggtcgttt acggttatcc tgccatttcc ctgcctcttt tggaggtgga ggtgacctgt	120
aaggtagttt ctgcagagag tccttggcgc ctttatgagg aggatacctt gtccatggcg	180
gtttatgcct cactggagta ttggtatcc acagaagcct gcgttcgttg tgagattgac	240
tcggtatcc ctgagaaacg ggggatgggt tcgtcagcgg ctatcagcat agcggccatt	300
cgtgcggtat ttgactacta tcaggctgat ctgcctcatg atgtactaga aatcttggtc	360
aatcgagctg agatgattgc ccatatgaat cctagtgggt tggatgctaa gacctgtctc	420
agtgaccaac ctattcgctt tatcaagaac gtaggattta cagaacttga gatggattta	480
tccgcctatt tgggtattgc cgatacgggt gtttatggtc atactcgtga agccatcaa	540
gtggttcaaa ataagggcaa ggtatgccta ccgtttttgc atgccttggg agaattaacc	600
cagcaagcag aagttcgat ttcacaaaaa tatgctgaag gactgggact aatcttcagt	660

caagctcatt tacatctaaa agaaattgga gtcagtagcc ctgaggcaga ctttttggtt	720
gaaacggctc ttagctatgg tgctctgggt gccaaatga gcggtggtgg gctaggaggt	780
tgtatcatag ccttggtaac caatttgacg cacgcacaag aactagcaga aagattagaa	840
gagaaaggag ctgttcagac atggatagag agcctgtaa	879
<210> 25	
<211> 1011	
<212> DNA	
<213> Streptococcus pneumoniae	
<400> 25	
atgattgctg ttaaaacttg cggaaaactc tattgggcag gtgaatatgc tattttagag	60
ccagggcagt tagctttgat aaaggatatt cccatctata tgagggtga gattgctttt	120
tctgacagct accgtatcta ttcagatatg ttgatattcg cagtggactt aaggcccaat	180
cctgactaca gcttgattca agaaacgatt gctttgatgg gagacttcct cgctgttcgc	240
ggtcagaatt taagaccttt ttcctaaaa atctgtggca aaatggaacg agaagggaaa	300
aagtttggtc taggttctag tggcagcgtc gttgtcttgg ttgtcaaggc tttactggct	360
ctctataatc tttcggttga tcagaatctc ttgttcaagc tgactagcgc tgtcttgctc	420
aagcgaggag acaatggttc catgggcgac ctgacctgta ttgtggcaga ggatttggtt	480
ctttaccagt catttgatcg ccagaaggcg gctgcttggg tagaagaaga aaacttggcg	540
acagttctgg agcgtgattg gggatTTTT atctcacaag tgaaccaac tttagaatgt	600
gatttcttag tgggatggac caaggaagtg gctgtatcga gtcacatggt ccagcaaac	660
aagcaaaata tcaatcaaaa ttttttaagt tcctcaaaag aaacggtggt ttccttggtc	720
gaagccttgg agcaggggaa agccgaaaaa gttatcgagc aagtagaagt agccagcaag	780
cttttagaag gcttgagtac agatatttac acgcctttgc ttagacagtt gaaagaagcc	840
agtcaagatt tgcaggccgt tgccaagagt agtggtgctg gtggtggatga ctgtggcatc	900
gccctgagtt ttgatcgca atcttctcga aacactttaa aaaatcgttg ggccgatctg	960
gggattgagc tcttatatca agaaaggata ggacatgacg acaaatcgta a	1011
<210> 26	
<211> 954	
<212> DNA	
<213> Streptococcus pneumoniae	
<400> 26	

atggatagag agcctgtaac agtacgttcc tacgcaaata ttgctattat caaatattgg	60
ggaaagaaaa aagaaaaaga gatgggtgcct gctactagca gtattttctct aactttggaa	120
aatatgtata cagagacgac cttgtcgctt ttaccagcca atgtaacagc tgacgaattt	180
tacatcaatg gtcagctaca aaatgaggtc gagcatgcca agatgagtaa gattattgac	240
cgttatcgtc cagctggtag gggctttgtc cgtatcgata ctcaaaacaa tatgcctacg	300
gcagcgggcc tgtcctcaag ttctagtggg ttgtccgcc tggtcaaggc ttgtaatgct	360
tatttcaagc ttggattgga tagaagtcag ttggcacagg aagccaaatt tgcctcaggc	420
tcttcttctc ggagttttta tggaccacta ggagcctggg ataaggatag tggagaaatt	480
tacctgtag agacagactt gaaactagct atgattatgt tgggtgctaga ggacaagaaa	540
aaaccaatct ctagccgtga cgggatgaaa ctttgtgtgg aaacctcgac gacttttgac	600
gactgggttc gtcagctga gaaggactat caggatatgc tgatttatct caaggaaaat	660
gattttgcca agattggaga attaacggag aaaaatgctc tggctatgca tgctacgaca	720
aagactgcta gtccagcctt ttcttatctg acggatgcct cttatgaggc tatggccttt	780
gttcgccagc ttcgtgagaa aggagaggcc tgctacttta ccatggatgc tgggtccaat	840
gttaaggtct tctgtcagga gaaagacttg gagcatttgt cagaaatttt cggtcagcgt	900
tatcgcttga ttgtgtcaaa aacaaaggat ttgagtcaag atgattgctg ttaa	954
<210>	27
<211>	546
<212>	DNA
<213>	E. coli K12 MG1655
<400>	27
atgcaaacgg aacacgtcat ttatttgaat gcacaggagg ttcccacggg tacgttgaa	60
aagtatgccg cacacacggc agacacccgc ttacatctcg cgttctccag ttggtgttt	120
aatgccaaag gacaattatt agttaccgc cgcgactga gcaaaaaagc atggcctggc	180
gtgtggacta actcggtttg tgggcaccca caactgggag aaagcaacga agacgcagt	240
atccgcggtt gccgttatga gcttggcgtg gaaattacgc ctctgaatc tatctatct	300
gactttcgct accgcgccac cgatccgagt ggcatgtgtg aaaatgaagt gtgtccggt	360
tttgccgcac gcaccactag tgcgttacag atcaatgatg atgaagtgat ggattatcaa	420
tgggtgtgatt tagcagatgt attacacggg attgatgcca cgccgtgggc gttcagtcg	480
tggatggtga tgcaggcgac aaatcgcgaa gccagaaaac gattatctgc atttaccag	540
cttaaa	546

<210> 28
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer
 <400> 28
 cggtaccaat gacaaaaaa gttggtgtcg g 31
 <210> 29
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> Reverse primer
 <400> 29
 ttctagatta cgatttgtcg tcatgtcc 28
 <210> 30
 <211> 32
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer
 <400> 30
 ccccgaggagg agagaaatta tgcaaacgga ac 32
 <210> 31
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220><223> Reverse primer
 <400> 31
 tgcattgctta tttaagctgg gtaaatg 27
 <210> 32
 <211> 44
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer
 <400> 32

cgagctcagg agcatttaga tgttgaaaac agtagttatt attg	44
<210> 33	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> Reverse primer	
<400> 33	
gcccgggggtg gcctgaaacg gctacc	26
<210> 34	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220><223> Forward primer	
<400> 34	
gcccggggagg agttaaagaa atgacaattg	30
<210> 35	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220><223> Reverse primer	
<400> 35	
cggatccctt agtttcgata agagcgaac	29