



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0139959
(43) 공개일자 2017년12월20일

(51) 국제특허분류(Int. Cl.)

C12P 7/02 (2006.01) C12N 9/04 (2006.01)

C12N 9/10 (2006.01) C12N 9/14 (2006.01)

(52) CPC특허분류

C12P 7/02 (2013.01)

C12N 9/0006 (2013.01)

(21) 출원번호 10-2016-0072631

(22) 출원일자 2016년06월10일

심사청구일자 없음

(71) 출원인

경상대학교산학협력단

경상남도 진주시 진주대로 501 (가좌동)

(72) 발명자

김선원

경상남도 진주시 소호로 39, 114동 603호(충무공
동, 한림 풀에버)

박지빈

부산광역시 금정구 중앙대로1685번길 24, 108동
604호(부곡동, 경남한신아파트)

(뒷면에 계속)

(74) 대리인

두호특허법인

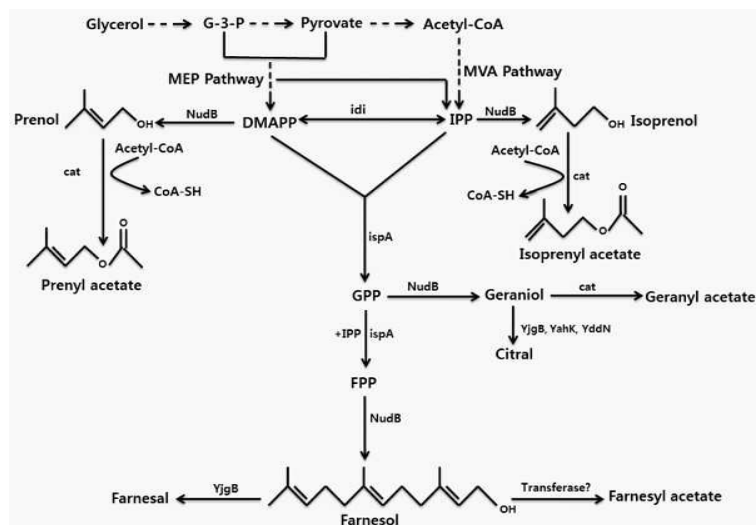
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **터펜 알코올 또는 그 유도체의 제조 방법**

(57) 요약

본 발명은 Nudix 계열의 효소를 종래 알려지지 않은 기질인 제라닐 피로포스페이트와 반응시켜 터펜 알코올 또는 그 유도체를 제조할 수 있어, 다양한 기질로부터 터펜 알코올을 제조할 수 있어, 터펜 알코올 또는 그 유도체를 복합체의 형태로 제조할 수 있는 터펜 알코올 또는 그 유도체의 제조 방법에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

C12N 9/1029 (2013.01)

C12N 9/1085 (2013.01)

C12N 9/14 (2013.01)

(72) 발명자

왕충룡

경상남도 진주시 내동로348번길 10, 111동 401호
(가좌동, 가좌그린빌주공아파트)

마크자다

경상남도 진주시 내동면 내동로219번길 41-16

명세서

청구범위

청구항 1

Nudix 계열의 효소를 제라닐 피로포스페이트와 반응시켜 제라니올을 얻는 단계를 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 2

청구항 1에 있어서, 상기 Nudix 계열의 효소는 NudB인 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 3

청구항 2에 있어서, 상기 NudB는 서열번호 1의 아미노산 서열과 70% 이상의 서열 상동성을 갖는 효소인 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 4

청구항 1에 있어서, 상기 반응은 Nudix 계열의 효소를 발현하는 미생물 내에서 수행되는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 5

청구항 4에 있어서, 상기 미생물은 메탈로네이트 경로의 효소를 더 발현하는 것인 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 6

청구항 5에 있어서, 상기 미생물은 제라닐 피로포스페이트 신타제를 더 발현하는 것인 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 7

청구항 1에 있어서, 상기 제라니올을 알코올 탈수소화 효소와 반응시키는 단계를 더 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 8

청구항 1에 있어서, 상기 제라니올을 아세틸 트랜스퍼라제 및 코엔자임 A와 반응시키는 단계를 더 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 9

청구항 7 또는 8에 있어서, 상기 반응은 Nudix 계열의 효소를 발현하는 미생물 내에서 수행되는 터펜 알코올 또는 그 유도체의 제조 방법.

는 그 유도체의 제조 방법.

청구항 10

청구항 1에 있어서, 상기 Nudix 계열의 효소를 이소펜테닐 피로포스페이트 또는 디메틸알릴 피로포스페이트와 반응시키는 단계를 더 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 11

청구항 10에 있어서, 상기 반응에서의 미반응 이소펜테닐 피로포스페이트 또는 디메틸알릴 피로포스페이트를 제라닐 피로포스페이트 신타제와 반응시키는 단계를 더 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 12

청구항 10에 있어서, 상기 반응은 Nudix 계열의 효소 및 제라닐 피로포스페이트 신타제를 발현하는 미생물 내에서 수행되는 터펜 알코올 또는 그 유도체의 제조 방법.

청구항 13

청구항 10에 있어서, 상기 Nudix 계열의 효소와 이소펜테닐 피로포스페이트 또는 디메틸알릴 피로포스페이트의 반응 산물을 아세틸 트랜스퍼라제 및 코엔자임 A와 반응시키는 단계를 더 포함하는 터펜 알코올 또는 그 유도체의 제조 방법.

발명의 설명

기술 분야

[0001] 본 발명은 터펜 알코올 또는 그 유도체의 제조 방법에 관한 것이다.

배경 기술

[0003] 터펜(테르펜, terpene)은 가연성의 불포화 탄화수소로 $(C_5H_8)_n$ 조성의 탄화수소 및 그 유도체를 총칭하는 것으로, n의 수에 의해 표시 명칭으로 부르고, 탄소 골격의 사슬 모양, 고리 모양에 의해 사슬식 테르펜 및 고리 모양 테르펜으로 또 성질에 의해 테르펜탄화수소, 테르펜알코올, 테르펜알데히드, 테르펜케톤 등으로 나뉜다.

[0004] 터펜은 향료, 의약품, 화장품, 가습기 살균제 등의 원료로 사용되는 것으로, 천연 유기 화합물 화학의 한 분야를 이루고 있다.

[0005] 이러한 테르펜의 생합성 경로 중 하나로는 메발로네이트 경로(mevalonate pathway)의 중간 산물인 디메틸알릴 피로포스페이트(DMAPP) 또는 이소펜테닐 피로포스페이트(IPP)를 Nudix 계열의 효소와 반응시켜 프레놀과 이소프레놀을 생성하는 방법이 알려져 있다.

[0006] 그러나, 상기 방법으로는 IPP와 DMAPP만을 기질로 사용하여 2종의 터펜 알코올을 얻을 수 있을 뿐인데, 방향제, 향수 등은 단일 물질이 아닌 다양한 향기 물질의 조합으로 얻어지는 바, 상기 방법으로는 방향제, 향수 등의 제조를 위해 다양한 추가 향기 물질을 조합해야 하는 문제가 있다.

선행기술문헌

비특허문헌

- [0008] (비특허문헌 0001) Synthetic pathway for production of five-carbon alcohols from isopentenyl diphosphate (Chou, Keasling. Appl Environ Microbiol. 2012 Nov;78(22):7849-55. doi: 10.1128/AEM.01175-12. Epub 2012 Aug 31.)

발명의 내용

해결하려는 과제

- [0009] 본 발명은 디메틸알릴 파이로포스페이트 및 이소펜테닐 파이로포스페이트 외에 다른 기질을 사용할 수 있는 Nudix 계열의 효소를 이용한 터펜 알코올 또는 그 유도체의 제조 방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

- [0011] 본 발명의 터펜 알코올 또는 그 유도체의 제조 방법은 Nudix 계열의 효소를 제라닐 파이로포스페이트와 반응시켜 제라니올을 얻는 단계를 포함한다.
- [0012] Nudix 계열(Nudix family, Nucleoside Diphosphate linked to X)의 효소는 포스포하이드롤라제(phosphohydrolases)의 단백질 패밀리로서 다양한 파이로포스페이트를 기질로 사용할 수 있다.
- [0013] 본 명세서에서 Nudix 계열의 효소는 Nudix 계열에 속하는 것이라면 특별히 한정되지 않으며, 예를 들면 NudA, NudB, NudC, NudD, NudE, NudF, NudG, NudH, NudI, NudJ, NudK, NudL 등일 수 있으며, 구체적으로 NudB(dihydroneopterin triphosphate pyrophosphohydrolase, DHNTPase)일 수 있다.
- [0014] 보다 구체적으로, 상기 NudB는 서열번호 1의 아미노산 서열과 70% 이상의 상동성을 갖는 아미노산 서열, 바람직하게는 상동성 80% 이상, 보다 바람직하게는 상동성 90% 이상, 가장 바람직하게는 서열번호 1의 아미노산 서열을 갖는 것일 수 있다.
- [0015] 본 명세서에서 제라닐 파이로포스페이트(Geranyl pyrophosphate, Geranyl diphosphate)는 HMG-CoA 리덕타제 경로(HMG-CoA reductase pathway)의 중간산물로서, 디메틸알릴 파이로포스페이트(DMAPP) 또는 이소펜테닐 파이로포스페이트(IPP)를 제라닐 파이로포스페이트 신타제(ispA)와 반응시켜 얻어지는 것일 수 있다.
- [0016] 본 발명자는 Nudix 계열의 효소가 제라닐 파이로포스페이트를 기질로 사용할 수 있음을 발견한 것에 착안하여 본 발명을 고안한 것으로서, 본 발명의 방법은 Nudix 계열의 효소를 제라닐 파이로포스페이트와 반응시키는 단계를 포함함으로써, C10 터펜 알코올인 제라니올(geraniol)을 제조할 수 있다.
- [0017] 본 발명의 일 구현예에 있어서, 상기 Nudix 계열의 효소와 제라닐 파이로포스페이트의 반응은 미생물 내에서 수행될 수 있다.
- [0018] 상기 미생물은 원핵 세포 또는 진핵 세포로 액체 배지에서 배양될 수 있는 것일 수 있다. 상기 미생물은 예를 들면, 박테리아, 곰팡이, 또는 이들의 조합일 수 있다. 박테리아는 그람 양성 박테리아, 그람 음성 박테리아, 또는 이들의 조합일 수 있다. 그람 음성 박테리아는 에세리키아(Escherichia) 속일 수 있다. 그람 양성 박테리아는 바실러스 속, 코리네박테리움 속, 액티노마이세스 속, 유산균 또는 이들의 조합일 수 있다. 곰팡이는 효모, 클루베로마이세스 속, 또는 이들의 조합일 수 있다.
- [0019] 에세리키아 속 미생물은 대장균일 수 있고, 구체적으로 DH5 α , MG1655, BL21(DE), S17-1, XL1-Blue, BW25113 또는 이들의 조합일 수 있다.
- [0020] 상기 미생물은 Nudix 계열의 효소를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다. 그러한 경우, 미생물 내에서 Nudix 계열의 효소가 생성되어 상기 Nudix 계열의 효소와 제라닐 파이로포스페이트의 미생물 내에서의 반응에 의해 제라니올의 생산이 연속적으로 수행될 수 있다.
- [0021] 상기 미생물이 Nudix 계열의 효소를 코딩하는 유전자로 형질전환되어 Nudix 계열의 효소를 발현시키는 경우, 상

기 미생물은 당 분야에 공지된 Nudix 계열의 효소를 코딩하는 유전자로 형질전환된 것일 수 있고, 보다 구체적으로, 에세리키아 콜라이 유래의 NudB 효소의 서열번호 1의 아미노산 서열과 70% 이상의 상동성을 갖는 아미노산 서열을 코딩하는 유전자, 바람직하게는 상동성 80% 이상, 보다 바람직하게는 상동성 90% 이상, 가장 바람직하게는 서열번호 1의 아미노산 서열을 코딩하는 유전자로 형질전환된 것일 수 있다.

- [0022] 에세리키아 콜라이 유래의 NudB 효소를 코딩하는 유전자는 서열번호 2의 뉴클레오티드 서열(NCBI, Gene ID: 946383)을 갖는 것일 수 있다.
- [0023] 또한, 본 발명의 터펜 알코올 또는 그 유도체의 제조 방법은 상기 제라니올을 알코올 탈수소화 효소와 반응시키는 단계를 더 포함할 수 있다.
- [0024] 상기 제라니올을 알코올 탈수소화 효소와 반응시키면 시트랄(citral)이 얻어질 수 있다.
- [0025] 상기 알코올 탈수소화 효소는 알코올을 탈수소화 할 수 있는 효소라면 당 분야에 공지된 것이 제한 없이 사용될 수 있으며, 예를 들면 에세리키아 콜라이 유래의 서열번호 3의 아미노산 서열을 갖는 yjgB일 수 있다.
- [0026] 에세리키아 콜라이 유래의 yjgB를 코딩하는 유전자는 서열번호 4의 뉴클레오티드 서열(NCBI, Gene ID: 948802)을 갖는 것일 수 있다.
- [0027] 본 발명의 일 구현예에 따르면 상기 제라니올과 알코올 탈수소화 효소의 반응은 미생물 내에서 수행될 수 있다.
- [0028] 상기 미생물은 알코올 탈수소화 효소를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다.
- [0029] 상기 미생물이 알코올 탈수소화 효소를 코딩하는 유전자로 형질전환되어 알코올 탈수소화 효소를 발현시키는 경우, 상기 미생물은 당 분야에 공지된 알코올 탈수소화 효소를 코딩하는 유전자로 형질전환된 것일 수 있고, 보다 구체적으로, 서열번호 3의 아미노산 서열을 코딩하는 유전자로 형질전환된 것일 수 있다.
- [0030] 또한, 본 발명의 터펜 알코올 또는 그 유도체의 제조 방법은 상기 제라니올을 아세틸 트랜스퍼라제(acetyl transferase) 및 코엔자임 A와 반응시키는 단계를 더 포함할 수 있다.
- [0031] 상기 제라니올을 아세틸 트랜스퍼라제 및 코엔자임 A와 반응시키면 제라닐 아세테이트(geranyl acetate)가 얻어질 수 있다.
- [0032] 상기 아세틸 트랜스퍼라제는 당 분야에 공지된 아세틸 트랜스퍼라제를 제한 없이 사용할 수 있으며, 예를 들면 클로람페니콜 아세틸트랜스퍼라제(Chloramphenicol acetyltransferase)일 수 있다. 보다 구체적으로 pSTV28 (Takara Co.) 발현벡터(서열번호 28의 뉴클레오티드 서열)의 항생제 마커 유래의 서열번호 5의 아미노산 서열을 갖는 클로람페니콜 아세틸트랜스퍼라제일 수 있다.
- [0033] pSTV28 (Takara Co.) 발현벡터의 항생제 마커 유래의 클로람페니콜 아세틸트랜스퍼라제를 코딩하는 유전자는 서열번호 6의 뉴클레오티드 서열을 갖는 것일 수 있다.
- [0034] 상기 코엔자임 A는 아세틸 트랜스퍼라제의 기질로 사용될 수 있는 것이라면 당 분야에 공지된 코엔자임 A를 제한 없이 사용될 수 있으며, 예를 들면 아세틸-CoA일 수 있고, 보다 구체적으로는 에세리키아 콜라이 유래의 아세틸-CoA일 수 있다.
- [0035] 본 발명의 일 구현예에 따르면 상기 제라니올과 아세틸 트랜스퍼라제 및 코엔자임 A의 반응은 미생물 내에서 수행될 수 있다.
- [0036] 상기 미생물은 아세틸 트랜스퍼라제를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다.
- [0037] 상기 미생물이 아세틸 트랜스퍼라제를 코딩하는 유전자로 형질전환되어 아세틸 트랜스퍼라제를 발현시키는 경우, 상기 미생물은 당 분야에 공지된 아세틸 트랜스퍼라제를 코딩하는 유전자로 형질전환된 것일 수 있고, 보다 구체적으로, pSTV28 (Takara Co.) 발현벡터의 항생제 마커 유래의 서열번호 5의 아미노산 서열을 갖는 클로람페니콜 아세틸트랜스퍼라제를 코딩하는 유전자로 형질전환된 것일 수 있다.
- [0038] 상기 미생물은 코엔자임 A를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다.
- [0039] 코엔자임 A의 경우, 통상 미생물 세포 내에 풍부하게 존재하고, 구체적으로 아세틸-CoA가 매우 풍부하게 존재하므로, 통상 별도의 형질전환 없이 상기 반응의 수행이 가능하다.
- [0040] 또한, 본 발명의 터펜 알코올 또는 그 유도체의 제조 방법은 상기 Nudix 계열의 효소를 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트와 반응시키는 단계를 더 포함할 수 있다.

- [0041] 이소펜테닐 파이로포스페이트 및 디메틸알릴 파이로포스페이트는 메틸에리스리톨 4-포스페이트 경로 (Methylerythritol 4-phosphate (MEP) pathway) 또는 메발로네이트 경로(Mevalonate pathway)의 산물이다.
- [0042] Nudix 계열의 효소를 이소펜테닐 파이로포스페이트와 반응시킴으로써 C5 터펜 알코올인 이소프레놀을 얻을 수 있고, Nudix 계열의 효소를 디메틸알릴 파이로포스페이트와 반응시킴으로써 C5 터펜 알코올인 프레놀을 얻을 수 있다.
- [0043] 또한, 본 발명은 상기 Nudix 계열의 효소와 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트와의 반응에서의 미반응 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트를 제라닐 파이로포스페이트 신타제와 반응시키는 단계를 더 포함할 수 있다.
- [0044] 제라닐 파이로포스페이트 신타제는 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트의 제라닐 파이로포스페이트로의 전환을 촉매하는 것이라면 당 분야에 공지된 제라닐 파이로포스페이트 신타제가 사용될 수 있으며, 예를 들면 에세리키아 콜라이 유래의 파네실 파이로포스페이트 신타제의 아미노산 서열 80번의 세린을 페닐알라닌으로 치환해서 제라닐 파이로포스페이트 신타제로 변형시킨 서열번호 7의 아미노산 서열을 갖는 것일 수 있다.
- [0045] 에세리키아 콜라이 유래의 파네실 파이로포스페이트 신타제의 아미노산 서열 80번의 세린이 페닐알라닌으로 치환된 제라닐 파이로포스페이트 신타제를 코딩하는 유전자(ispAS80F)는 서열번호 8의 뉴클레오티드 서열을 갖는 것일 수 있다.
- [0046] 본 발명의 일 구현예에 따르면 상기 Nudix 계열의 효소와 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트의 반응은 미생물 내에서 수행될 수 있다.
- [0047] 상기 미생물은 메발로네이트 경로의 효소들을 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다.
- [0048] 상기 메발로네이트 경로의 효소는 당 분야에 널리 알려진 것으로서, 아세틸-CoA 아세틸트랜스퍼라제, 하이드록시메틸글루타릴(HMG)-CoA 신타제, HMG-CoA 리덕타제, 메발로네이트 키나제, 포스포메발로네이트 키나제 및 메발로네이트 파이로포스페이트 데카르복실라제, IPP 이소머라제가 있다.
- [0049] 상기 미생물은 메발로네이트 경로의 효소들을 형질전환에 의해 발현시키는 경우, 상기 미생물은 상기 효소들을 코딩하는 당 분야에 공지된 유전자로 형질전환된 것일 수 있고, 예를 들면, 서열번호 9의 엔테로코커스 패칼리스 (Enterococcus faecalis) 유래의 아세틸-CoA 아세틸트랜스퍼라제/하이드록시메틸글루타릴 (HMG)-CoA 리덕타제의 아미노산 서열, 서열번호 10의 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제의 아미노산 서열, 서열번호 11의 스트렙토코커스 뉴모니아 (Streptococcus pneumoniae) 유래의 메발로네이트 키나제의 아미노산 서열, 서열번호 12의 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제의 아미노산 서열, 서열번호 13의 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제의 아미노산 서열, 서열번호 14의 에세리키아 콜라이 (Escherichia coli) 유래의 IPP 이소머라제의 아미노산 서열을 코딩하는 유전자로 형질전환된 것일 수 있다.
- [0050] 엔테로코커스 패칼리스 (Enterococcus faecalis) 유래의 아세틸-CoA 아세틸트랜스퍼라제/하이드록시메틸글루타릴 (HMG)-CoA 리덕타제를 코딩하는 유전자는 서열번호 15의 뉴클레오티드 서열, 엔테로코커스 패칼리스 유래의 HMG-CoA 신타제를 코딩하는 유전자는 서열번호 16의 뉴클레오티드 서열, 스트렙토코커스 뉴모니아 (Streptococcus pneumoniae) 유래의 메발로네이트 키나제를 코딩하는 유전자는 서열번호 17의 뉴클레오티드 서열, 스트렙토코커스 뉴모니아 유래의 포스포메발로네이트 키나제를 코딩하는 유전자는 서열번호 18의 뉴클레오티드 서열, 스트렙토코커스 뉴모니아 유래의 메발로네이트 디포스페이트 데카르복실라제를 코딩하는 유전자는 서열번호 19의 뉴클레오티드 서열, 에세리키아 콜라이 유래의 IPP 이소머라제를 코딩하는 유전자는 서열번호 20의 뉴클레오티드 서열(NCBI, Gene ID: 949020)을 갖는 것일 수 있다.
- [0051] 본 발명에 있어서, 상기 미생물의 IPP 이소머라제의 발현량을 조절하여 프레놀과 이소프레놀의 생산량을 조절할 수 있다.
- [0052] 구체적으로, IPP 이소머라제는 IPP의 DMAPP로의 전환을 촉매하므로, IPP 이소머라제의 발현량을 늘려 DMAPP 생산량을 늘리고, 이에 따라 프레놀 생산량을 늘릴 수 있다. 반대로, IPP 이소머라제의 발현량을 줄여 IPP의량을 늘리고, 이에 따라 이소프레놀의 생산량을 늘릴 수 있다.
- [0053] 또한, 상기 미생물은 제라닐 파이로포스페이트 신타제를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다. 그러한 경우, 미반응 이소펜테닐 파이로포스페이트 및 디메틸알릴 파이로포스페이트가 제라닐 파이로포스

페이트 신타제와 반응하여 제라닐 파이로포스페이트가 얻어지고, 이는 추가로 상기 Nudix 계열의 효소와 반응하여, 보다 많은 양의 제라니올이 얻어질 수 있다.

- [0054] 상기 미생물이 제라닐 파이로포스페이트 신타제를 코딩하는 유전자로 형질전환되어 제라닐 파이로포스페이트 신타제를 발현시키는 경우, 상기 미생물은 당 분야에 공지된 제라닐 파이로포스페이트 신타제를 코딩하는 유전자로 형질전환된 것일 수 있고, 보다 구체적으로, 대장균 유래의 파네실 파이로포스페이트 신타제의 아미노산 서열 80번의 세린을 페닐알라닌으로 치환해서 제라닐 파이로포스페이트 신타제로 변형시킨 서열번호 7의 아미노산 서열을 코딩하는 유전자(ispAS80F)로 형질전환된 것일 수 있다.
- [0055] 또한, 본 발명의 터펜 알코올 또는 그 유도체의 제조 방법은 상기 Nudix 계열의 효소와 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트와의 반응 생성물을 아세틸 트랜스퍼라제(acetyl transferase) 및 코엔자임 A와 반응시키는 단계를 더 포함할 수 있다.
- [0056] 상기 Nudix 계열의 효소와 이소펜테닐 파이로포스페이트 또는 디메틸알릴 파이로포스페이트의 반응 생성물인 이소프레놀과 프레놀을 아세틸 트랜스퍼라제 및 코엔자임 A와 반응시키면 각각 이소프레닐 아세테이트(isoprenyl acetate)와 프레닐 아세테이트(prenyl acetate)가 얻어질 수 있다.
- [0057] 상기 아세틸 트랜스퍼라제 및 코엔자임 A는 전술한 아세틸 트랜스퍼라제 및 코엔자임 A일 수 있고, 상기 반응은 미생물 내에서 수행될 수 있다.
- [0058] 상기 미생물은 Nudix 계열의 효소를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있으며, 또한, 아세틸 트랜스퍼라제를 내재적으로 또는 형질전환에 의해 발현시키는 것일 수 있다.
- [0059] 본 발명에 있어서, 전술한 각 단계들은 서로 독립적으로 수행될 수도 있고, 그 일부 또는 전부가 미생물 내에서 수행될 수 있다. 그 일부 또는 전부가 미생물 내에서 수행되는 경우에는 미생물은 각 단계에 필요한 효소를 코딩하는 유전자들로 형질전환된 것일 수 있고, 그러한 경우 전술한 일부 또는 전부의 터펜 알코올 및 그 유도체를 생산할 수 있으며, 전술한 각 단계가 모두 미생물 내에서 수행되는 경우에는 그 미생물은 프레놀, 프레닐 아세테이트, 이소프레놀, 이소프레닐 아세테이트, 제라니올, 시트랄 및 제라닐 아세테이트를 모두 생산할 수 있다.
- [0060] 전술한 반응들이 미생물 내에서 수행되는 경우, 미생물은 탄소원의 존재 하에서 해당 미생물에 적합한 배양 조건으로 배양될 수 있다.
- [0061] 배양은 합성, 반합성, 또는 복합 배양 배지에서 이루어질 수 있다. 배양 배지로는 탄소원, 질소원, 비타민 및 미네랄로 구성된 배지를 사용할 수 있다. 예를 들어, MRS (Man-Rogosa-Sharp) 액체 배지 또는 우유가 첨가된 액체 배지를 사용할 수 있다.
- [0062] 배지의 탄소원으로는 전분, 포도당, 자당, 갈락토스, 과당, 글리세롤, 글루코스 또는 이들의 혼합물이 사용될 수 있다. 예를 들면, 글리세롤이 탄소원으로 사용될 수 있다. 질소원으로는 황산암모늄, 질산암모늄, 질산나트륨, 글루탐산, 카사미노산, 효모추출물, 펩톤, 트립톤, 대두박 또는 이들의 혼합물이 사용될 수 있다. 미네랄은 염화나트륨, 인산제이칼륨, 황산마그네슘 또는 이들의 혼합물이 사용될 수 있다.
- [0063] 배양이 발효조에서 이루어지는 경우 글루코스를 배지의 탄소원으로 사용하는 것이 좋다. 시험관 배양의 경우에는 글리세롤을 배지의 탄소원으로 사용하는 것이 좋다.
- [0064] 미생물 배양 배지 내 상기 탄소원, 질소원 및 미네랄 각각은 예를 들면, 리터당 10 내지 100 g, 5 내지 40 g 및 0.5 내지 4 g을 이용할 수 있다.
- [0065] 상기의 통상의 배양 배지에 첨가되는 비타민은 비타민 A, 비타민 B, 비타민 C, 비타민 D, 비타민 E 또는 이들의 혼합물일 수 있다. 비타민은 통상의 배양 배지에 상기에서 언급된 탄소원, 질소원, 미네랄 등과 함께 첨가되거나, 멸균하여 준비된 배지에 별도로 첨가될 수 있다.
- [0066] 배양은 통상의 미생물 배양 조건으로 수행될 수 있다. 배양은 예를 들어 약 15-45℃, 예를 들면, 15-44℃, 15-43℃, 15-42℃, 15-41℃, 15-40℃, 15-39℃, 15-38℃, 15-37℃, 15-36℃, 15-35℃, 15-34℃, 15-33℃, 15-32℃, 15-31℃, 15-30℃, 20-45℃, 20-44℃, 20-43℃, 20-42℃, 20-41℃, 20-40℃, 20-39℃, 20-38℃, 20-37℃, 20-36℃, 20-35℃, 20-34℃, 20-33℃, 20-32℃, 20-31℃, 20-30℃, 25-45℃, 25-44℃, 25-43℃, 25-42℃, 25-41℃, 25-40℃, 25-39℃, 25-38℃, 25-37℃, 25-36℃, 25-35℃, 25-34℃, 25-33℃, 25-32℃, 25-31℃, 25-30℃, 27-45℃, 27-44℃, 27-43℃, 27-42℃, 27-41℃, 27-40℃, 27-39℃, 27-38℃, 27-37℃,

27-36℃, 27-35℃, 27-34℃, 27-33℃, 27-32℃, 27-31℃ 또는 27-30℃에서 수행될 수 있다.

[0067] 배양액 중의 배양 배지를 제거하고 농축된 균체만을 회수하거나 제거하기 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.

[0068] 배양의 일 예에 있어서, 배양은 탄소원으로서 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 글리세롤은 배지 중의 유일한 탄소원일 수 있다. 0.5-5.0%(w/v), 예를 들면, 0.5-4.5%(w/v), 0.5-4.0%(w/v), 0.5-3.5%(w/v), 0.5-3.0%(w/v), 0.5-2.5%(w/v), 0.5-2.0%(w/v), 1-5.0%(w/v), 1-4.5%(w/v), 1-4.0%(w/v), 1-3.5%(w/v), 1-3.0%(w/v) 또는 1-2.5%(w/v)의 글리세롤을 포함하는 배지에서 이루어지는 것일 수 있다. 상기 배지는 글리세롤 및 아라비노스가 첨가된 YT 배지일 수 있다. YT 배지는 1.6중량% 트립톤, 1중량% 효모 추출물 및 0.5 중량% NaCl을 포함할 수 있다.

발명의 효과

[0070] 본 발명은 Nudix 계열의 효소를 종래 알려지지 않은 기질인 제라닐 파이로포스페이트와 반응시켜 터펜 알코올 또는 그 유도체를 제조할 수 있다.

[0071] 본 발명은 Nudix 계열의 효소를 이용하여 다양한 기질로부터 터펜 알코올을 제조할 수 있어, 터펜 알코올 또는 그 유도체를 복합체의 형태로 제조할 수 있다.

도면의 간단한 설명

[0073] 도 1은 터펜 알코올과 유도체들의 생합성 경로 모식도이다. 점선은 다단계 효소반응으로 구성된 대사경로를 생략해서 표시한 것을 의미한다. 도 1에서 각 약자는 하기와 같다. G-3-P, glyceraldehyde 3-phosphate; MEP, methyl-erythritol phosphate; MVA, mevalonate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, (E,E)-farnesyl pyrophosphate. The abbreviations of enzymes are as follows: idi, isopentenyl diphosphate isomerase; NudB, phosphatase; ispA, (E,E)-farnesyl pyrophosphate synthase; cat, chloramphenicol acetyl transferase; YjgB and yahK, NADPH-dependent aldehyde reductases; YddN, ethanol dehydrogenase/alcohol dehydrogenase; transferase?, endogenous promiscuous transferase of E. coli.

도 2는 재조합 대장균 MT1에 도입된 플라스미드인 pSNA와 pT-NudB-ispA에 구축된 오페론 모식도이다. 직사각형 화살표는 유전자를 굵은 실선 화살표는 프로모터를 나타낸다. 첫번째 오페론은 MVA 경로 오페론으로 6개의 유전자로 구성되어 있고 lac promoter에 의해 발현이 된다. 본 발명의 일 실시예에서 이 오페론을 발현백터인 pSTV28(pACYC184 origin, CmR)에 클로닝해서 pSNA 플라스미드를 만들었다. NudB와 ispA 유전자로 구성된 두번째 오페론은 trc promoter에 의해 발현이 되고, 이 오페론을 발현백터 pTrc99A (pBR322 ori, AmpR)에 클로닝해서 pT-NudB-ispA를 제조하였다. 도 2에서 각 약자는 하기와 같다. MvaE, bifunctional acetoacetyl-CoA thiolase and HMG-CoA reductase; MvaS, HMG-CoA synthase; MvaK1, mevalonate kinase; MvaK2, phosphomevalonate kinase; MvaD, mevalonate diphosphate decarboxylase; IDI, IPP isomerase; NudB, diphosphatase; ispA, FPP synthase.

도 3은 pSNA와 pT-NudB-ispA 플라스미드로 형질전환된 재조합 대장균(E. coli DH5α strain MT1)에서의 C5, C15 터펜 알코올과 그의 에스터 유도체 생산 결과를 나타낸 것이다.

도 4는 pSNAK와 pT-NudB-ispA 플라스미드로 형질전환된 재조합 대장균(E. coli DH5α strain MT2)에서의 C5, C15 터펜 알코올과 그의 에스터 유도체 생산 결과를 나타낸 것이다.

도 5는 pSNA와 pT-NudB-ispA-idi 플라스미드로 형질전환된 재조합 대장균(E. coli DH5α strain MT3)에서의 C5, C15 터펜 알코올과 그의 에스터 유도체 생산 결과를 나타낸 것이다.

도 6은 pT-NudB-ispAS80F와 pSNA 또는 pSNAK 플라스미드로 형질전환된 재조합 대장균(E. coli DH5α strain MT4 또는 MT5)에서의 C5, C10 터펜 알코올과 그의 에스터 유도체 생산 결과를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

- [0074] 이하, 본 발명을 구체적으로 설명하기 위해 실시예를 들어 상세하게 설명하기로 한다.
- [0076] **실시예 1. C5, C15 터펜 알코올과 그의 에스터 유도체 생산**
- [0077] C5, C15 터펜 알코올 생산을 위해서 IPP/DMAPP를 공급하는 외래 MVA 경로를 도입하고, 세포 내에서 과량으로 합성되는 IPP, DMAPP, FPP를 탈인산화해서 터펜 알코올을 만드는 탈인산화 효소 NudB를 과발현하였다.
- [0078] NudB는 dihydroneopterin triphosphate pyrophosphohydrolase (DHNTPase)라고 불리는 효소로 dihydroneopterin triphosphate (DHNTP)에서 pyrophosphate를 제거해서 dihydroneopterin phosphate를 만드는 효소로 folate 생합성에 관련되어 있다. 상기 NudB는 IPP/DMAPP 이외의 다른 isoprenyl diphosphates 기질에도 작용할 것이라고 추정하는 연구 그룹은 없었다.
- [0079] 본 발명자들은 외래 MVA 경로를 갖는 플라스미드 pSNA와 FPP synthase와 NudB를 갖는 플라스미드 pT-NudB-ispA를 도입한 재조합 대장균(E. coli DH5 α strain MT1)을 제조하고(도 2), 2%(v/v) glycerol을 포함하는 2YT 배지, 30℃에서 48시간 배양을 수행했다(도 3). 플라스미드 pSNA는 한국등록특허 제1440922호, 1392159호, 1359484호에 기재된 바와 동일한 방법으로 제조하였다.
- [0080] 배양 결과, 734 mg/L isoprenol, 15 mg/L prenol, 520 mg/L farnesol이 생산되었고, 예상하지 않았던 233 mg/L isoprenyl acetate, 31 mg/L farnesal, 121 mg/L farnesyl acetate도 얻어졌다. 상기의 물질들은 GC-MS 분석을 통해서 검증되었고, 최종 총 생산량은 1652 mg/L로 계산되었다.
- [0082] Isoprenyl acetate와 farnesyl acetate는 isoprenol과 farnesol의 acetyl ester로 acetyl transferase 효소에 의해서 생성되는 것으로 추정된다. 1차적으로는 MVA 경로 플라스미드(pSNA)에 존재하는 항생제 마커인 chloramphenicol acetyl-transferase (CAT)에 의해서 생성이 되는 것으로 추정되어서, pSNA에 있는 항생제 마커를 kanamycin 저항성(Neomycin phosphotransferase II [NPT II/Neo]) 유전자로 대체한 pSNAK 플라스미드를 구축하였다. 플라스미드 pT-NudB-ispA와 pSNAK를 도입한 재조합 대장균(E. coli DH5 α strain MT2)을 제조하고, 2%(v/v) glycerol을 포함하는 2YT 배지, 30℃에서 48시간 배양을 수행했다(도 4). 배양결과, 911 mg/L isoprenol, 12 mg/L prenol, 28 mg/L farnesal, 80 mg/L farnesyl acetate, 480 mg/L farnesol이 생산되었다. CAT의 제거를 통해서 isoprenyl acetate는 233 mg/L에서 0 mg/L로 생산이 중단되었고, farnesyl acetate는 121 mg/L에서 80 mg/L로 감소하였다. 그러나 상당한 양의 farnesyl acetate가 여전히 생산이 되는 것은 대장균 내의 다른 acetyl transferase가 farnesyl acetate 생성에 주로 관여하는 것으로 추정되고, 이 효소를 발굴해서 이용하면 farnesyl acetate의 생성량을 더 효과적으로 조절할 수 있을 것으로 예상된다. 또한 isoprenyl acetate 생성이 차단되면서 isoprenol은 734 mg/L에서 911 mg/L로 증가하였다.
- [0084] Farnesol의 산화된 형태인 farnesal은 식물에서는 farnesol dehydrogenase에 의해 생성된다(Ahmad-Sohdi et al., 2015). 따라서 farnesol dehydrogenase를 대장균에 도입을 하면 farnesal을 생산할 수 있다. 그러나 상기의 실시예(도 3, 4)에 따르면 식물 유래 farnesol dehydrogenase의 도입 없이도 farnesal의 생성을 관찰할 수 있었다. 이것은 대장균이 갖고 있는 특정 alcohol dehydrogenases에 의해서 생성이 된 것으로 추정된다. 대장균에 존재하는 farnesal 생성에 관련이 있을 것으로 추정되는 효소로 YjgB, YahK(서열번호 21의 아미노산 서열, 서열번호 22의 뉴클레오티드 서열(NCBI, Gene ID: 944975)로 코딩됨), YddN(서열번호 23의 아미노산 서열, 서열번호 24의 뉴클레오티드 서열(NCBI, Gene ID: 946036)로 코딩됨), AdhE(서열번호 25의 아미노산 서열, 서열번호 26의 뉴클레오티드 서열(NCBI, Gene ID: 945837)로 코딩됨)를 선정하고 이들을 과발현하는 재조합 균주를 이용해서 farnesol의 farnesal로의 전환을 관찰하였다. 발현벡터 pTrc99A(서열번호 27의 뉴클레오티드 서열(GenBank: U13872.1))에 상기 효소들의 유전자가 클로닝된 pT-YjgB, pT-Yahk, pT-YddN, pT-adhE를 구축하고, 이들을 도입한 재조합 대장균들을 farnesol이 포함된 배지에서 배양하면서 farnesal의 생성을 관찰하였다. pT-YjgB로 형질전환된 균주에서만 150 mg/L의 farnesal의 생성이 관찰되었다. 따라서 대장균에서 farnesol을 farnesal로 전환을 하는 YjgB 효소의 발현을 조절해서 farnesol과 farnesal의 생산량을 조절할 수 있을 것이다.

[0086] 상기 도 3의 재조합 대장균 MT1의 배양 결과에서 prenol의 생산량이 isoprenol에 비해서 매우 낮은 것은 IPP를 DMAPP로 전환하는 IPP isomerase (IDI)의 발현량이 MVA 경로 플라스미드인 pSNA에서 발현되는 것만으로는 부족한 것으로 추정된다. 따라서 IDI 유전자(idi)를 pT-NudB-ispA 플라스미드에 추가로 도입한 pT-NudB-ispA-idi를 구축하였다. pT-NudB-ispA-idi와 pSNA로 형질전환된 재조합 대장균(*E. coli* DH5α strain MT3)을 제조하고, 2%(v/v) glycerol을 포함하는 2YT 배지, 30℃에서 48시간 배양을 수행했다(도 5). 실험결과, prenol의 생산량은 증가하고 isoprenol의 생산량은 예상대로 감소하였다. 그리고 farnesol 생산이 재조합 균주 MT1의 520 mg/L에 비해서 991 mg/L로 크게 증가하였다. 이것은 FPP 생합성을 개시하는 DMAPP의 세포 내 양이 IDI의 추가발현에 의해서 증가했기 때문으로 추정된다. 따라서 IDI의 발현량 조절을 통해서 생산되는 터펜 알코올의 prenol, isoprenol, farnesol의 생산량을 조절할 수 있다.

[0088] 실시예 2. C5, C10 터펜 알코올과 그의 에스터 유도체 생산

[0089] MVA 경로와 IspA의 과발현을 통해서 세포내 IPP (C5), DMAPP (C5), FPP (C15) 생합성을 증가시킬 수 있고, 이들은 NudB에 의해서 각각의 터펜 알코올과 유도체들로 전환이 된다. 따라서 GPP synthase (GPPS) 발현을 통해서 GPP (C10)를 생합성하면 C10 터펜 알코올을 생산할 수 있을 것이다.

[0090] 대장균은 GPP synthase를 갖고 있지는 않지만 IspA를 변이 시켜서 GPP synthase로 만들 수 있다(Reiling et al., 2004). GPP 생합성 능력을 갖도록 변이시킨 ispAS80F로 pT-NudB-ispA 플라스미드의 ispA를 대체한 pT-ispAS80F-NudB 플라스미드를 구축하고, pSNA와 함께 대장균에 도입해서 재조합 대장균(*E. coli* DH5α strain MT4)를 제조하였다. 이 균주를 2%(v/v) glycerol을 포함하는 2YT 배지, 30℃에서 48시간 배양을 수행하고 터펜 알코올 생산을 분석하였다(도 6).

[0091] 실험 결과, 21 mg/L geraniol, 13 mg/L citral, 65 mg/L geranyl acetate의 C10 터펜 알코올과 유도체가 얻어졌다. 이들 생산량은 재조합 대장균 MT1이 생산하는 C15 터펜 알코올과 유도체의 생산량 보다 상대적으로 매우 낮은 수준이다. 이것은 IspA 효소의 변이를 통해서 GPP 생합성 활성을 갖도록 할 수 있지만, 이 과정에서 효소의 비활성이 크게 감소한 것으로 추정된다. 따라서 효소 비활성이 높은 GPP synthase를 사용하게 되면 C10 터펜 알코올과 유도체의 생산량을 크게 증가시킬 수 있을 것이다.

[0092] 에스터 유도체 생성을 줄이고 알코올 생산을 증가시키기 위해서 pT-ispAS80F-NudB와 pSNA로 형질전환한 재조합 대장균(*E. coli* DH5α strain MT5)를 제조하였다. 이 균주를 2%(v/v) glycerol을 포함하는 2YT 배지, 30℃에서 48시간 배양을 수행하고 터펜 알코올 생산을 분석하였다(도 6).

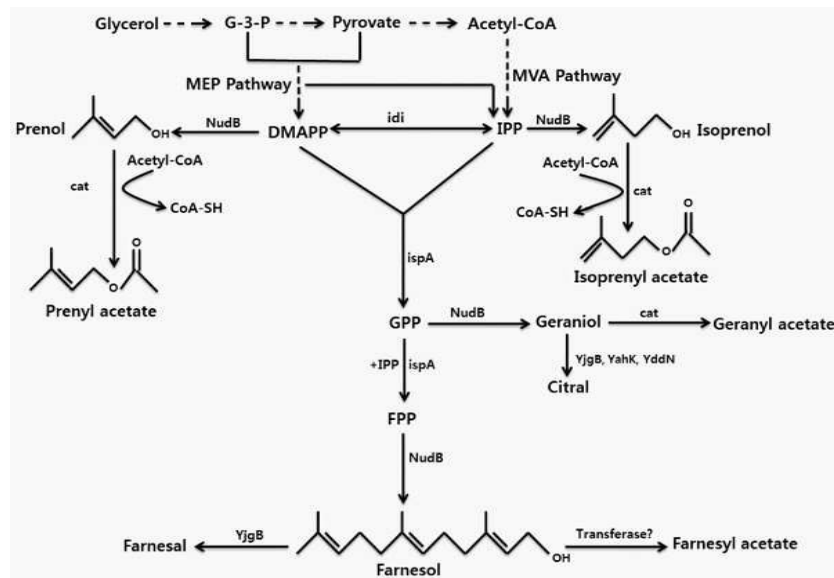
표 1

구분	설명	서열번호	참조
프라이머(밑줄은 제한 효소 부위를 나타냄)			
NudB-F1	CATCCATGGTTTAAGAAGGAGATATACATATGAAGGATAAAGTGATAAGCGTCC	29	—
NudB-R1	CATGGAGCTCTTAGGCAGCGTTAATTACAACTG	30	—
NudB-F2	CTGGATCCTTTAAGAAGGAGATATACATATGAAGGATAAAGTGATAAGCGTCC	31	—
NudB-R2	CCCAAGCTTAGGCAGCGTTAATTACAACTG	32	—
ispA-F	CTAGGAGCTCAGCCCTAAGGAACCAATATGGACTTTCCGCAGCAAC	33	—
ispA-R	TGAGGTACCTTATTTATTACGCTGGATGATGTAG	34	—
ispAS80F-F	GAGGTACCAACAATAATCTCGTATAGTATGGACTTTCCGCAGCAAC	35	—
ispAS80F-R	CTGGATCCTTATTTATTACGCTGGATGATGTAG	36	—
Idi-F	ACGGATCCTGAGGAGGTAACGTATGCAACGGAACACGTCATTTTA	37	—
Idi-R	TATCGTCGACTCTAAGATCTTATTTAAGCTGGGTAAATGCAG	38	—
플라스미드			
pSTV28	P _{lac} expression vector, pACYC184 origin, lacZ, Cm ^r	28	Takara Co., Ltd
pSNA	pSTV28 containing mvaE and mvaS of <i>E. faecalis</i> , mvaK1, mvaK2, and mvaD of <i>S. pneumoniae</i> , and idi of <i>E. coli</i>	—	Yoon et al., (2009)

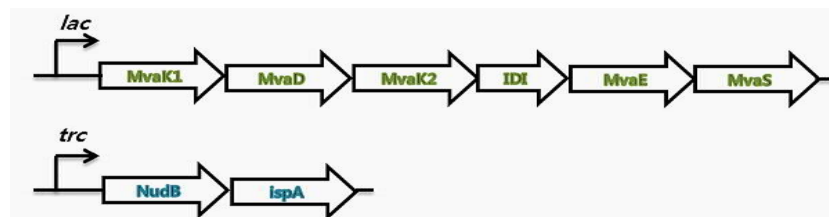
pSTV28K	P _{lac} expression vector, pACYC184 origin, lacZ, Km ^r , 서열번호 29	39	Zhou et al., (2014)
pSNAK	pTV28 K containing mvaE and mvaS from <i>E. faecalis</i> , mvaK1, mvaK2, and mvaD from <i>S. pneumoniae</i> , and idi from <i>E. coli</i>	-	Zhou et al., (2014)
pTrc99A	P _{trc} expression vector, pBR322 origin, lacI ^q , Amp ^r	27	Amann et al., (1988)
pT-NudB	pTrc99A vector containing NudB from <i>E. coli</i>	-	-
pT-NudB-ispA	pTrc99A vector containing NudB and ispA from <i>E. coli</i>	-	-
pT-NudB-ispA-idi	pTrc99A vector containing NudB, ispA and idi from <i>E. coli</i>	-	-
pT-ispAS80F	pTrc99A vector containing ispA mutant from <i>E. coli</i>	-	-
pT-ispAS80F-NudB	pTrc99A vector containing ispA mutant and NudB from <i>E. coli</i>	-	-
pT-YahK	pTrc99A vector containing YahK from <i>E. coli</i>	-	-
pT-YddN	pTrc99A vector containing YddN from <i>E. coli</i>	-	-
pT-adhE	pTrc99A vector containing adhE from <i>E. coli</i>	-	-
pT-YjgB	pTrc99A vector containing YjgB from <i>E. coli</i>	-	-
pT-GPSGES	pTrc99A containing GPS from <i>E. coli</i> and GES from <i>Ocimum basilicum</i>	-	Zhou et al., (2014)
균주			
MG1655	<i>E. coli</i> K-12; F ⁻ lambda ⁻ , ilvG ⁻ , rfb-50, rph-1	-	ATCC 700926
DH5 α	<i>E. coli</i> K-12; F ⁻ , Φ80lacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(rK ⁻ , mK ⁺) phoA, supE44, λ ⁻ , thi-1	-	ATCC 98040
MT1	<i>E. coli</i> DH5 α harboring pT-NudB-ispA and pSNA	-	-
MT2	<i>E. coli</i> DH5 α harboring pT-NudB-ispA and pSNAK	-	-
MT3	<i>E. coli</i> DH5 α harboring pT-NudB-ispA-idi and pSNA	-	-
MT4	<i>E. coli</i> DH5 α harboring pT-ispAS80F-NudB and pSNA	-	-
MT5	<i>E. coli</i> DH5 α harboring pT-ispAS80F-NudB and pSNAK	-	-
D-pTrc99A	<i>E. coli</i> DH5 α harboring pTrc99A	-	-
D-pT-YjgB	<i>E. coli</i> DH5 α harboring pT-YjgB	-	-
D-pT-Yahk	<i>E. coli</i> DH5 α harboring pT-Yahk	-	-
D-pT-YddN	<i>E. coli</i> DH5 α harboring pT-YddN	-	-
D-pT-adhE	<i>E. coli</i> DH5 α harboring pT-adhE	-	-

도면

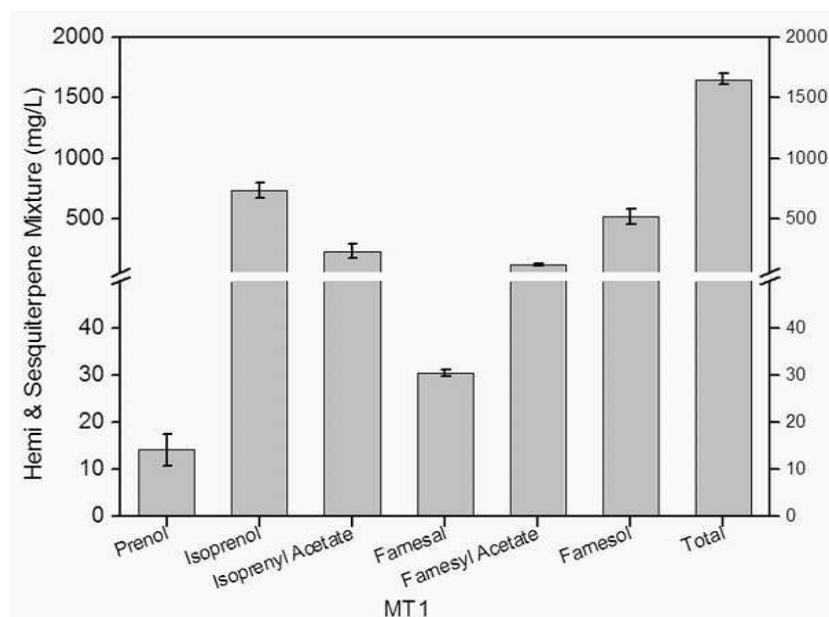
도면1



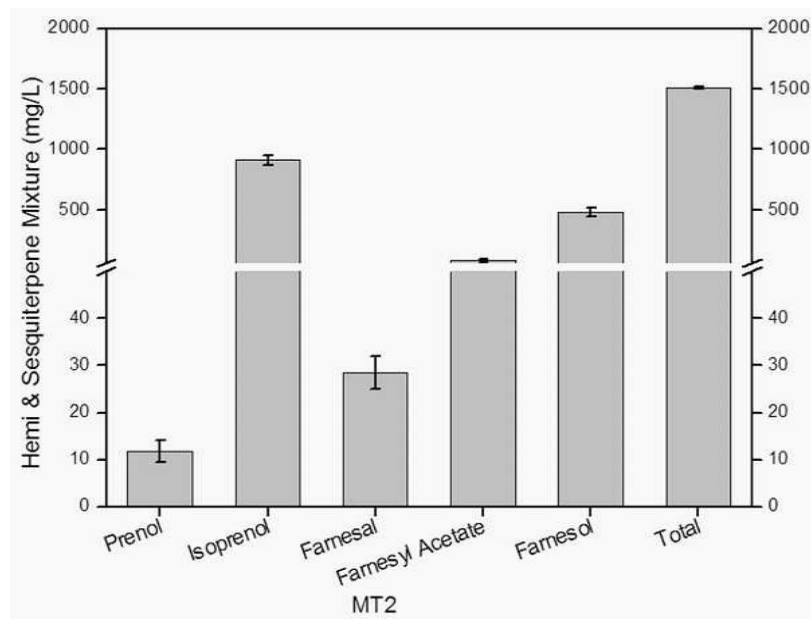
도면2



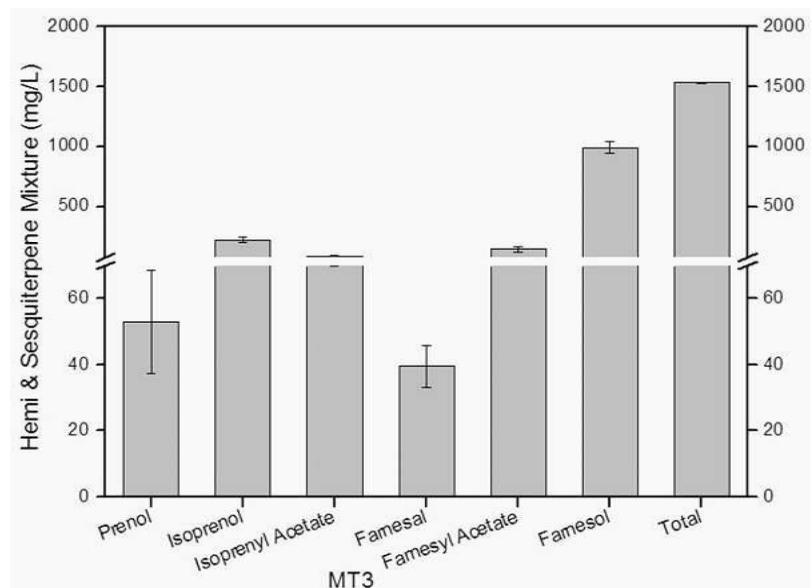
도면3



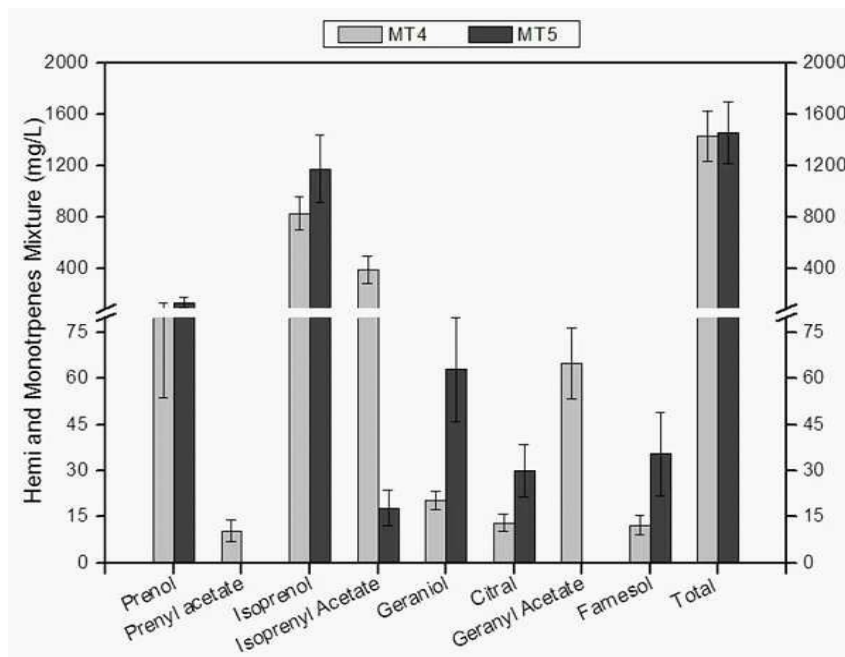
도면4



도면5



도면6



서열 목록

<110> GYEONGSANG NATIONAL UNIVERSITY OFFICE OF ACADEMY AND INDUSTRY COLLABORATION

<120> PREPARING METHOD FOR TERPENE ALCOHOL OR ITS DERIVATIVES

<130> 16P05001

<160> 39

<170> Kopatent In 2.0

<210> 1

<211> 150

<212> PRT

<213> Escherichia coli str. K-12 substr. MG1655

<400> 1

Met Lys Asp Lys Val Tyr Lys Arg Pro Val Ser Ile Leu Val Val Ile

1 5 10 15

Tyr Ala Gln Asp Thr Lys Arg Val Leu Met Leu Gln Arg Arg Asp Asp

20 25 30

Pro Asp Phe Trp Gln Ser Val Thr Gly Ser Val Glu Glu Gly Glu Thr

35 40 45

Ala Pro Gln Ala Ala Met Arg Glu Val Lys Glu Glu Val Thr Ile Asp

50 55 60

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

65 70 75 80
Phe Glu Ile Phe Ser His Leu Arg His Arg Tyr Ala Pro Gly Val Thr

 85 90 95

Arg	Asn	Thr	Glu	Ser	Trp	Phe	Cys	Leu	Ala	Leu	Pro	His	Glu	Arg	Gln
			100					105					110		
Ile	Val	Phe	Thr	Glu	His	Leu	Ala	Tyr	Lys	Trp	Leu	Asp	Ala	Pro	Ala
			115				120					125			
Ala	Ala	Ala	Leu	Thr	Lys	Ser	Trp	Ser	Asn	Arg	Gln	Ala	Ile	Glu	Gln
			130				135					140			
Phe	Val	Ile	Asn	Ala	Ala										
145						150									

$\langle 210 \rangle$	2
-----------------------	---

<211> 453

<212> DNA

<213> Escherichia coli str. K-12 substr. MG1655

 $\langle 400 \rangle$

2

atgaaggata aagtgtataa gcgctcccggt tcatctttag tggatcatcta cgcacaagat	60
acgaacacggg tgctgatgtt gcagcggcgt gacgatcccg atttctggca gtcggtaacc	120
ggcagcgtgg aagaggggtga aaccgcgcgc caagctgcc a tgcgcgaagt aaaggaagag	180
gtcaccattg atgttgtcgc tgaacaactg accttaattg actgtcagcg cacggtagag	240
tttgaattt tttcacattt acgtcatcgc tatgcgcgg gcgtgacgcg taatacggaa	300
tcatggttct gtcttgcgct tccgcacgag cggcagatcg ttttactga acatctggct	360
tacaagtggc ttgatgcgcc tgctgcggcg gcgctacta agtcctggag caaccggcag	420

gcgattgaac agtttgtaat taacgctgcc taa 453

<210> 3

<211> 339

<212> PRT

<213> Escherichia coli str. K-12 substr. MG1655

<400> 3

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

1 5 10 15

Glu Val Tyr Glu Tyr Asp Pro Gly Glu Leu Arg Pro Gln Asp Val Glu

20 25 30
 Val Gln Val Asp Tyr Cys Gly Ile Cys His Ser Asp Leu Ser Met Ile
 35 40 45

 Asp Asn Glu Trp Gly Phe Ser Gln Tyr Pro Leu Val Ala Gly His Glu
 50 55 60
 Val Ile Gly Arg Val Val Ala Leu Gly Ser Ala Ala Gln Asp Lys Gly
 65 70 75 80
 Leu Gln Val Gly Gln Arg Val Gly Ile Gly Trp Thr Ala Arg Ser Cys
 85 90 95
 Gly His Cys Asp Ala Cys Ile Ser Gly Asn Gln Ile Asn Cys Glu Gln
 100 105 110
 Gly Ala Val Pro Thr Ile Met Asn Arg Gly Gly Phe Ala Glu Lys Leu

 115 120 125
 Arg Ala Asp Trp Gln Trp Val Ile Pro Leu Pro Glu Asn Ile Asp Ile
 130 135 140
 Glu Ser Ala Gly Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Lys Pro
 145 150 155 160
 Leu Leu Met His His Ile Thr Ala Thr Ser Arg Val Gly Val Ile Gly
 165 170 175
 Ile Gly Gly Leu Gly His Ile Ala Ile Lys Leu Leu His Ala Met Gly
 180 185 190

 Cys Glu Val Thr Ala Phe Ser Ser Asn Pro Ala Lys Glu Gln Glu Val
 195 200 205
 Leu Ala Met Gly Ala Asp Lys Val Val Asn Ser Arg Asp Pro Gln Ala
 210 215 220
 Leu Lys Ala Leu Ala Gly Gln Phe Asp Leu Ile Ile Asn Thr Val Asn
 225 230 235 240
 Val Ser Leu Asp Trp Gln Pro Tyr Phe Glu Ala Leu Thr Tyr Gly Gly
 245 250 255
 Asn Phe His Thr Val Gly Ala Val Leu Thr Pro Leu Ser Val Pro Ala

 260 265 270

Phe Thr Leu Ile Ala Gly Asp Arg Ser Val Ser Gly Ser Ala Thr Gly
 275 280 285
 Thr Pro Tyr Glu Leu Arg Lys Leu Met Arg Phe Ala Ala Arg Ser Lys
 290 295 300
 Val Ala Pro Thr Thr Glu Leu Phe Pro Met Ser Lys Ile Asn Asp Ala
 305 310 315 320
 Ile Gln His Val Arg Asp Gly Lys Ala Arg Tyr Arg Val Val Leu Lys
 325 330 335

Ala Asp Phe

<210> 4
 <211> 1020
 <212> DNA
 <213> Escherichia coli str. K-12 substr. MG1655
 <400> 4

atgtcgtatga taaaaagcta tgccgcaaaa gaagcgggcg gcgaactgga agtttatgag 60
 tacgatcccc gtgagctgag gccacaagat gttgaagtgc aggtggatta ctgcgggatac 120
 tgccattccg atctgtcgat gatcgataac gaatggggat tttcacaata tccgctggtt 180
 gccgggcatg aggtgattgg gcgcgtgggt gcactcggga gcgccgcgca ggataaaggt 240
 ttgcaggtcg gtcagcgtgt cgggattggc tggacggcgc gtagctgtgg tcaactgcgac 300

 gcctgtatta gcggtaatca gatcaactgc gagcaagggt cggtgccgac gattatgaat 360
 cgcggtggct ttcccgagaa gttgcgtgct gactggcaat gggtgattcc actgccagaa 420
 aatattgata tcgagtcgcg cgggcccgtg ttgtgcggcg gtatcacggt ctttaacca 480
 ctgttgatgc accatatcac tgctaccagc cgcgttgggg taattggtat tggcgggctg 540
 gggcatatcg ctataaaact tctgcacgca atgggatgct aggtgacagc ctttagttct 600
 aatccggcga aagagcagga agtgcgtggc atgggtgccg ataaagtgt gaatagccgc 660
 gatccgcagg cactgaaagc actggcgggg cagtttgatc tcattatcaa caccgtcaac 720

 gtcagcctcg actggcagcc ctattttgag gcgctgacct atggcggtaa ttccatacg 780
 gtcggtgcgg ttctcacgcc gctgtctgtt ccggccttta cgtaattgc gggcgatcgc 840
 agcgtctctg gttctgttac cggcacgcct tatgagctgc gtaagctgat gcgttttgcc 900
 gcccgagca aggttgcgcc gaccaccgaa ctgttccga tgtcgaaaat taacgacgcc 960
 atccagcatg tgcgcgacgg taaggcgcgt taccgcgtgg tgttgaaagc cgatttttga 1020

1020

<210> 5

<211> 219

<212> PRT

<213> Artificial Sequence

<220><223> pSTV28 expression vector

<400> 5

Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp

1 5 10 15

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr

20 25 30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val

35 40 45

Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala

50 55 60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly

65 70 75 80

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His

85 90 95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp

100 105 110

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly

115 120 125

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe

130 135 140

Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val

145 150 155 160

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr

165 170 175

Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His

180 185 190

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu

195 200 205

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala

210 215

<210> 6

<211> 660

<212> DNA

<213> Artificial Sequence

<220><223> pSTV28 expression vector

<400> 6

atggagaaaa aaatcactgg atataccacc gttgatatat cccaatggca tcgtaaagaa 60

cattttgagg catttcagtc agttgctcaa tgtacctata accagaccgt tcagctggat 120

attacggcct ttttaaagac cgtaaagaaa aataagcaca agttttatcc ggcctttatt 180

cacattcttg cccgcctgat gaatgctcat ccggaatttc gtatggcaat gaaagacggt 240

gagctggatga tatgggatag tggtcacctt tggtacaccg ttttccatga gcaaaactgaa 300

acgttttcat cgctctggag tgaataccac gacgatttcc ggcagtttct acacatatat 360

tcgcaagatg tggcgtgtta cggtgaaaac ctggcctatt tccttaaagg gtttattgag 420

aatatgtttt tcgtctcagc caatccctgg gtgagtttca ccagttttga tttaaacgtg 480

gccaatatgg acaacttctt cgccccgtt ttcacatgg gcaaatatta tacgcaaggc 540

gacaaggtgc tgatgccgtt ggcgattcag gttcatcatg ccgtctgtga tggcttccat 600

gtcggcagaa tgcttaatga attacaacag tactgcgatg agtggcaggg cggggcgtaa 660

660

<210> 7

<211> 299

<212> PRT

<213> Artificial Sequence

<220><223> Escherichia coli

<400> 7

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 Phe

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>	8		
<211>	900		
<212>	DNA		
<213>	Artificial Sequence		
<220><223>	Escherichia coli		
<400>	8		
atggactttc cgcagcaact cgaagcctgc gttaagcagg ccaaccaggc gctgagccgt		60	
tttatcgccc cactgccctt tcagaacact cccgtggtcg aaaccatgca gtatggcgca		120	
ttattagggtg gtaagcgctt gcgacctttc ctggtttatg ccaccgggtca tatgttcggc		180	
gttagcacia acacgctgga cgcacccgct gccgccgttg agtgtatcca cgcttacttt		240	
ttaattcatg atgatttacc ggcaatggat gatgacgac tgcgtcgcgg ttigccaacc		300	
tgccatgtga agtttggcga agcaaacgcg attctcgttg gcgacgcttt acaaacgctg		360	
gcgttctcga ttttaagcga tgccgatatg ccggaagtgt cggaccgcga cagaatttcg		420	
atgatttctg aactggcgag cgccagtggg attgccggaa tgtgcggtgg tcaggcatta		480	
gatttagacg cggaaggcaa acacgtacct ctggacgcgc ttgagcgtat tcatcgtcat		540	
aaaaccggcg cattgattcg cgccgccgtt cgccttgggt cattaagcgc cggagataaa		600	
ggacgtcgtg ctctgccggt actcgacaag tatgcagaga gcatcggcct tgccttcag		660	
gttcaggatg acatcctgga tgtgtgggga gatactgcaa cgttgggaaa acgccagggt		720	
gccgaccagc aacttggtaa aagtacctac cctgcacttc tgggtcttga gcaagcccgg		780	
aagaaagccc gggatctgat cgacgatgcc cgtcagtcgc tgaacaact ggctgaacag		840	
tactcgata cctcggcact ggaagcgcta gcggactaca tcaccagcg taataaataa		900	
		900	
<210>	9		
<211>	803		
<212>	PRT		
<213>	Enterococcus faecalis		
<400>	9		
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 Pro Glu Ser Leu Ile Asp Lys Leu Gln
 500 505 510
 Val Arg Glu Ala 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 Ala 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> 10
<211> 383
<212> PRT
<213> Enterococcus faecalis
<400> 10
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> 11
 <211> 292
 <212> PRT
 <213> Streptococcus pneumoniae
 <400> 11

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 Pro 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 Asn Ile Thr Glu Ala Cys Ile 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 Asp Ala Glu Gly Leu Gly Gln Ile Leu 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 Thr Leu Ser His 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> 12
<211> 335
<212> PRT
<213> Streptococcus pneumoniae
<400> 12

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 Glu 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 Asp Val Ser Val Asp Gln
115 120 125

Glu 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 Val Ala Ala Trp Leu Glu Glu

165 170 175
Glu Asn Leu Ala Thr Val Leu Glu Arg Asp Trp Gly Phe Ser 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 Glu Ser Thr Lys Thr Leu Lys Asn Arg Trp Ala Asp Leu Gly

305 310 315 320
Ile Glu Leu Leu Tyr Gln Glu Arg Ile Gly His Asp Asp Lys Ser

325 330 335
<210> 13
<211> 317
<212> PRT
<213> Streptococcus pneumoniae
<400> 13

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

325 330 335
<210> 13
<211> 317
<212> PRT
<213> Streptococcus pneumoniae
<400> 13

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 Ser Pro Ala Thr Ala Asp Ala Phe Tyr Ile Asn Gly
 50 55 60
 Gln Leu Gln Asn Glu Ala Glu His Ala Lys Met Ser Lys Ile Ile Asp
 65 70 75 80
 Arg Tyr Arg Pro Asp 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 Val Cys Asn Ala Tyr Phe Gln Leu Gly Leu Thr 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 Met 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 Val Tyr Leu Lys Glu Asn Asp Phe Ala Lys
 210 215 220
 Val Gly Lys Leu Thr Glu Glu 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 Asp Phe Val Arg Gln Leu Arg Glu Gln Gly Glu Asp Cys Tyr
 260 265 270
 Phe Thr Met Asp Ala Gly Pro Asn Val Lys Val Leu 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> 14
 <211> 182
 <212> PRT
 <213> Escherichia coli str. K-12 substr. MG1655
 <400> 14

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>	15
<211>	2412
<212>	DNA
<213>	Enterococcus faecalis
<400>	15
ttgaaaacag tagttattat tgatgcattt cgaacaccaa ttggaaaata taaaggcagc	60
ttaagtcaag taagtgccgt agacttagga acacatgtta caacacaact tttaaaaga	120
cattccacta ttctgaaga aattgatcaa gtaatctttg gaaatgtttt acaagctgga	180
aatggccaaa atcccgcacg acaaatagca ataacacagc gtttatctca tgaaattccc	240
gcaatgacag ttaatgaggt ctgcggatca ggaatgaagg ccgttatctt gccgaaacaa	300
ttgattcaat taggagaagc ggaagtttta attgctggcg ggattgagaa tatgtcccaa	360
gcacctaaat tacaacgatt taattacgaa acagaaagct atgatgcgcc tttttctagt	420
atgatgtacg atgggttaac ggatgccttt agtggtaag caatgggctt aactgctgaa	480
aatgtggcgg aaaagtatca tgtaactaga gaagagcaag atcaattttc tgtacattca	540
caattaaaag cagctcaagc acaagcagaa gggatatctg ctgacgaaat agccccatta	600
gaagtatcag gaacgcttgt ggagaaagat gaagggatc gccctaattc gagcgttgag	660
aagctaggaa cgcttaaac agtttttaaa gaagacggtt ctgtaacagc agggaaatgca	720
tcaaccatta atgatggggc ttctgctttg attattgctt cacaagaata tgccgaagca	780
cacggctctt 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 aggactcgt	1140
atgctactag agagacctca gcaaaaaaaa aacagccgat tttatcaat gactcctgag	1200
gaacgcctgg cttctcttct taatgaaggc cagatttctg ctgatacaaa aaaagaattt	1260
gaaaatacgg ctttatcttc gcagattgcc aatcatatga ttgaaatca aatcagtga	1320
acagaagtgc cgatgggcgt tggcttacct ttaacagtgg acgaaactga ttatttggt	1380
ccaatggcga cagaagagcc ctcaagtatt gcggctttga gtaatggtgc aaaaatagca	1440
caaggattta aaacagtga tcaacaacgt ttaatgcgtg gacaaatcgt tttttacgat	1500

gttcagacg ccgagtcatt gattgatgaa ctacaagtaa gagaaacgga aatttttcaa 1560
caagcagagt taagttatcc atctatcggt aaacgcggcg gcggcttaag agatttgcaa 1620
tatcggtgctt ttgatgaatc atttgtatct gtcgactttt tagtagatgt taaggatgca 1680
atgggggcaa atatcgtaa cgctatgttg gaaggtgtgg ccgagttgtt ccgtgaatgg 1740
tttgcggagc aaaagatttt attcagtatt ttaagtaatt atgccacgga gtcggttgtt 1800

acgatgaaaa cggtatttcc agtttcacgt ttaagtaagg ggagcaatgg ccgggaaatt 1860
gctgaaaaaa ttgttttagc ttcacgctat gcttcattag atccttatcg ggcagtcacg 1920
catacaaaag ggatcatgaa tggcattgaa gctgtcgttt tagctacagg aaatgataca 1980
cgcgctgtta gcgcttcttg tcatgctttt gcggtgaagg aaggctgcta ccaaggtttg 2040
actagttgga cgctggatgg cgaacaacta attggtgaaa tttcagttcc gcttgcgtta 2100
gccacggttg gcggtgccac aaaagtctta cctaaatctc aagcagctgc tgatttgta 2160
gcagtgacgg atgcaaaaga actaagtcga gtagtagcgg ctgttggttt ggcacaaaat 2220

ttagcggcgt tacgggcctt agtctctgaa ggaattcaaa aaggacacat ggctctacaa 2280
gcacgttctt tagcgatgac ggctcggagct actggtaaag aagttgaggc agtcgctcaa 2340
caattaaaac gtcaaaaaac gatgaaccaa gaccgagcct tggctatttt aaatgattta 2400
agaaaacaat aa 2412

<210> 16
<211> 1152
<212> DNA
<213> Enterococcus faecalis
<400> 16

atgacaattg ggattgataa aattagtttt tttgtgcccc cttattatat tgatatgacg 60
gcactggctg aagccagaaa tgtagaccct ggaaaatttc atattggtat tgggcaagac 120

caaatggcgg tgaaccaat cagccaagat attgtgacat ttgcagcaa tgccgcagaa 180
gcgatcttga ccaaagaaga taaagaggcc attgatatgg tgattgtcgg gactgagtec 240
agtatcgatg agtcaaaagc ggccgcagtt gtcttacatc gtttaatggg gattcaacct 300
ttcgctcgct ctttcgaaat caaggaagct tgttacggag caacagcagg cttacagtta 360
gctaagaatc acgtagcctt acatccagat aaaaaagtct tggttgtagc agcagatatt 420
gcaaaatatg gattaaattc tggcggtagag cctacacaag gagctggggc ggttgcaatg 480
ttagttgcta gtgaaccgcg catcttggct ttaaaagagg ataatgtgat gctgacgcaa 540

gatatctatg acttttggcg tccaacaggc catccgtatc ctatggtcga tggctccttg 600

tcaaacgaaa cctacatcca atcttttgcc caagtctggg atgaacataa aaaaagaacc	660
ggtcttgatt ttgcagatta tgatgcttta gcgttcataa ttccttacac aaaaatgggc	720
aaaaaagcct tattagcaaa aatctccgac caaactgaag cagaacagga acgaatttta	780
gcccgttatg aagaagcat catctatagt cgtcgcgtag gaaacttgta tacgggttca	840
ctttatctgg gactcatttc ctttttagaa aatgcaacga ctttaaccgc aggcaatcaa	900
attgggttat tcagttatgg ttctgggtgct gtcgctgaat ttttactgga tgaattagta	960
gctggttatc aaaatcattt acaaaaagaa actcatttag cactgctaga taatcggaca	1020
gaactttcta tcgctgaata tgaagccatg ttgcagaaa ctttagacac agatattgat	1080
caaacgttag aagatgaatt aaaatatagt atttctgcta ttaataatac cgttcgctct	1140
tatcgaaact aa	1152
<210> 17	
<211> 879	
<212> DNA	
<213> Streptococcus pneumoniae	
<400> 17	
atgacaaaaa aagttggigt cggtcaggca catagtaaga taattttaat aggggaacat	60
gcggtcgttt acggttatcc tgccatttcc ctgcctcttt tggaggtgga ggtgacctgt	120
aaggtagttc ctgcagagag tccttggcgc ctttatgagg aggatacctt gtccatggcg	180
gtttatgcct cactggagta ttggatatc acagaagcct gcattcgttg tgagattgac	240
tcggctatcc ctgagaaacg ggggatgggt tcgtcagcgg ctatcagcat agcggccatt	300
cgtgcggtat ttgactacta tcaggctgat ctgcctcatg atgtactaga aatcttggtc	360
aatcgagctg aaatgattgc ccatatgaat cctagtgggt tggatgctaa gacctgtctc	420
agtgaccaac ctattcgtt tatcaagaac gtaggattta cagaacttga gatggattta	480
tccgcctatt tggatgattgc cgatacgggt gtttatggtc atactcgtga agccatccaa	540
gtggttcaaa ataagggcaa ggatgcccta cgttttttgc atgccttggg agaattaacc	600
cagcaagcag aagttgcgat ttacaaaaa gatgctgaag gactgggaca aatcctcagt	660
caagcgatt tacatttaaa agaaattgga gtcagtagcc ctgaggcaga ctttttggtt	720
gaaacgactc ttagccatgg tgctctgggt gccaaagtga gcggtgggtg gctaggaggt	780
tgtatcatag ccttggtaac caatttgaca cagcacaag aactagcaga aagattagaa	840
gagaaaggag ctgttcagac atggatagag agcctgtaa	879
<210> 18	
<211> 1008	

<212> DNA

<213> Streptococcus pneumoniae

<400> 18

atgattgctg ttaaaacttg cggaaaactc tattgggcag gtgaatatgc tattttagag	60
ccagggcagt tagctttgat aaaggatatt cccatctata tgagggtga gattgctttt	120
tctgacagct accgtatcta ttcagatatg tttgatttcg cagtggactt aaggccta	180
cctgactaca gcttgattca agaaacgatt gctttgatgg gagacttcct cgctgttcgt	240
ggtcagaatt taagaccttt ttctctagaa atctgtggca aatggaacg agaagggaaa	300
aagtttggtc taggttctag tggcagcgtc gttgtcttgg ttgtcaaggc ttactggct	360
ctgtatgatg ttctgttga tcaggagctc ttgttcaagc tgactagcgc tgtcttgctc	420
aagcgaggag acaatggctc catgggcgac ctgacctgta ttgtggcaga ggatttgggt	480
ctctaccagt catttgatcg ccagaagggt gctgcttgggt tagaagaaga aaacttggcg	540
acagtcttgg agcgtgattg gggcttttca atttacaag tgaaaccaac tttagaatgt	600
gatttcttag tgggatggac caaggaagtg gctgtatcga gtcacatggt ccagcaa	660
aatc aagcaaaata tcaatcaaaa ttttttaagt tcctcaaaag aaacgggtgg	720
ttctttgggc gaagccttgg aacaggggaa atcagaaaag attatcgagc aagtagaagt	780
agccagcaag cttttagaag gcttgagtac agatatttac acgcctttgc ttagacagtt	840
gaaagaagcc agtcaagatt tgcaggccgt tgccaagagt agtgggtctg gtggtgggtga	900
ctgtggcatc gccctgagtt ttgatgcga atcaacaaa accttaaaaa atcgttgggc	960
cgatctgggg attgagctct tatatcaaga aaggatagga catgacgaca aatcgtaa	1008

<210> 19

<211> 954

<212> DNA

<213> Streptococcus pneumoniae

<400> 19

atggatagag agcctgtaac agtacgttcc tacgaaata ttgctattat caaatattgg	60
ggaaagaaaa aagaaaaaga gatgtgtcct gctactagca gtatttctct aactttggaa	120
aatatgtata cagagacgac cttgtcgcct ttaccagcca atgtaacagc tgacgaattt	180
tacatcaatg gtcagctaca aaatgaggtc gagcatgcca agatgagtaa gattattgac	240
cgttatcgtc cagctgggtga gggctttgtc cgtatcgata ctcaaaacaa tatgcctact	300
gcagcgggcc tgtcctcaag ttctagtggg ttgtccgcc tggtcaaggc ttgtaatgct	360

tatttcaagc ttggattgga tagaagtcag ttggcacagg aagccaaatt tgcctcaggc 420
tcttcttctc ggagttttta tggaccacta ggagcctggg ataaggatag tggagaaatt 480
taccctgtag agacagactt gaaactagct atgattatgt tgggtgctaga ggacaagaaa 540

aaaccaatct ctaccgtga cgggatgaaa ctttgtgtgg aaacctcgac gacttttgac 600
gactgggttc gtcagtctga gaaggactat caggatatgc tgatttatct caaggaaaat 660
gattttgcca agattggaga attaacggag aaaaatgccc tggtatgca tgctacgaca 720
aagactgcta gtccagcctt ttcttatctg acggatgcct cttatgaggc tatggacttt 780
gttcgccagc ttctgtagaa aggagaggcc tgctacttta ccatggatgc tggteccaat 840
gttaaggctt tctgtcagga gaaagacttg gagcatttat cagaaatttt cggtcacgt 900
tatcgcttga ttgtgtcaaa aacaaaggat ttgagtcaag atgattgctg ttaa 954

<210> 20
<211> 549
<212> DNA
<213> Escherichia coli str. K-12 substr. MG1655
<400> 20

atgcaaacgg aacacgtcat tttattgaat gcacaggag ttccacggg tacgtggaa 60
aagtatgccg cacacacggc agacacccgc ttacatctcg cgttctccag ttggtgttt 120
aatgccaaag gacaattatt agttacccgc cgcgcactga gcaaaaaagc atggcctggc 180
gtgtggacta actcggtttg tgggcacca caactgggag aaagcaacga agacgcagtg 240
atccgccgtt gccgttatga gcttggcgtg gaaattacgc ctctgaatc tatctatcct 300
gactttcgt accgcgccac cgatccgagt ggcatgttg aaaatgaagt gtgtccgta 360

tttgccgcac gcaccactag tgcgttacag atcaatgatg atgaagtgat ggattatcaa 420
tgggtgtgatt tagcagatgt attacacggt attgatgcca cgccgtgggc gttcagtccg 480
tggatggtga tgcaggcgac aaatcgcgaa gccagaaaac gattatctgc atttaccag 540
cttaaataa 549

<210> 21
<211> 349
<212> PRT
<213> Escherichia coli str. K-12 substr. MG1655
<400> 21

Met Lys Ile Lys Ala Val Gly Ala Tyr Ser Ala Lys Gln Pro Leu Glu

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

Leu Leu Lys Arg Asp Gly Thr Met Thr Leu Val Gly Ala Pro Ala Thr
 260 265 270
 Pro His Lys Ser Pro Glu Val Phe Asn Leu Ile Met Lys Arg Arg Ala
 275 280 285
 Ile Ala Gly Ser Met Ile Gly Gly Ile Pro Glu Thr Gln Glu Met Leu
 290 295 300

Asp Phe Cys Ala Glu His Gly Ile Val Ala Asp Ile Glu Met Ile Arg
 305 310 315 320
 Ala Asp Gln Ile Asn Glu Ala Tyr Glu Arg Met Leu Arg Gly Asp Val
 325 330 335
 Lys Tyr Arg Phe Val Ile Asp Asn Arg Thr Leu Thr Asp
 340 345

<210> 22
 <211> 1050
 <212> DNA
 <213> Escherichia coli str. K-12 substr. MG1655
 <400> 22

atgaagatca aagctgttgg tgcattattcc gctaaacaac cacttgaacc gatggatatc 60

acccggcgtg aaccgggacc gaatgatgtc aaaatcgaaa tcgcttactg tggcgtttgc 120
 cattccgata tccaccaggt ccgttccgag tgggcgggga cggtttacc ctcggtgccg 180
 ggtcatgaaa ttgtggggcg tgtggtagcc gttggtgac aggtagaaaa atatgcgccg 240
 ggcatctgg tcggtgtcgg ctgcattgtc gacagtgtga aacattgcga agagtgtgaa 300
 gacgggttgg aaaactactg tgatcacatg accggcacct ataactcgcc gacgccggac 360
 gaaccgggcc atactctggg cggctactca caacagatcg tcgttcatga gcgatatgtt 420
 ctgcgtattc gtcacccgca agagcagctg gcggcggtgg ctcttttgtt gtgtgcaggg 480

atcaccacgt attcggcgt acgtcactgg caggccgggc cgggtaaaa agtgggcgtg 540
 gtcggcatcg gcggtctggg acatatgggg attaatctgg ccacgcgat gggggcacat 600
 gtggtggcat ttaccacttc tgaggcaaaa cgcgaagcgg caaaagccct gggggccgat 660
 gaagtgtta actcacgcaa tgccgatgag atggcggtc atctgaagag ttctgatttc 720
 attttgaata cagtagctgc gccacataat ctgcagcatt ttaccacctt gctgaagcgt 780
 gatggcacca tgacgttgtt tgggtcgcct gcgacaccgc ataatcgcc ggaagttttc 840

aacctgatca tgaacgcccg tgcgatagcc ggttctatga ttggcggcat tccagaaact 900

caggagatgc tcgatttttg cgccgaacat ggcatcgtgg ctgatataga gatgattcgg 960

gccgatcaaa ttaatgaagc ctatgagcga atgctgcgcg gtgatgtgaa atatcgtttt 1020

gttatcgata atcgcacact aacagactga 1050

<210> 23

<211> 336

<212> PRT

<213> Escherichia coli str. K-12 substr. MG1655

<400> 23

Met Lys Ala Ala Val Val Thr Lys Asp His His Val Asp Val Thr Tyr

1 5 10 15

Lys Thr Leu Arg Ser Leu Lys His Gly Glu Ala Leu Leu Lys Met Glu

20 25 30

Cys Cys Gly Val Cys His Thr Asp Leu His Val Lys Asn Gly Asp Phe

35 40 45

Gly Asp Lys Thr Gly Val Ile Leu Gly His Glu Gly Ile Gly Val Val

50 55 60

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

65 70 75 80

Ser Val Ala Trp Phe Tyr Glu Gly Cys Gly His Cys Glu Tyr Cys Asn

85 90 95

Ser Gly Asn Glu Thr Leu Cys Arg Ser Val Lys Asn Ala Gly Tyr Ser

100 105 110

Val Asp Gly Gly Met Ala Glu Glu Cys Ile Val Val Ala Asp Tyr Ala

115 120 125

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

130 135 140

Cys Ala Gly Val Thr Thr Tyr Lys Ala Val Lys Leu Ser Lys Ile Arg

145 150 155 160

Pro Gly Gln Trp Ile Ala Ile Tyr Gly Leu Gly Gly Leu Gly Asn Leu

165 170 175

Ala Leu Gln Tyr Ala Lys Asn Val Phe Asn Ala Lys Val Ile Ala Ile
 180 185 190
 Asp Val Asn Asp Glu Gln Leu Lys Leu Ala Thr Glu Met Gly Ala Asp
 195 200 205
 Leu Ala Ile Asn Ser His Thr Glu Asp Ala Ala Lys Ile Val Gln Glu
 210 215 220
 Lys Thr Gly Gly Ala His Ala Ala Val Val Thr Ala Val Ala Lys Ala
 225 230 235 240

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

305 310 315 320
 Met Glu Glu Gly Lys Ile Arg Gly Arg Met Val Ile Asp Phe Arg His
 325 330 335

<210> 24

<211> 1011

<212> DNA

<213> Escherichia coli str. K-12 substr. MG1655

<400> 24

atgaaggctg cagttgttac gaaggatcat catgttgacg ttacgtataa aacactgcgc 60

tcactgaaac atggcgaagc cctgctgaaa atggagtgtt gtggtgtatg tcataccgat 120

cttcagtta agaatggcga ttttggtgac aaaaccggcg taattctggg ccatgaaggc 180

atcgggtgtgg tggcagaagt ggggtccaggt gtcacctcat taaaaccagg cgatcgtgcc 240

agcgtggcgt ggttctacga aggatgcggt cattgcgaat actgtaacag tggtaacgaa 300

acgctctgcc gttcagttaa aaatgccgga tacagcgttg atggcgggat ggcggaagag 360

tgcacgtgg tcgccgatta cgcggtaaaa gtgccagatg gtctggactc ggcggcggcc 420

agcagcatta cctgtgcggg agtcaccacc tacaaagccg ttaagctgtc aaaaattcgt 480
ccagggcagt ggattgtctat ctacggtctt ggcggctctgg gtaacctcgc cctgcaatac 540
gcgaagaatg tctttaacgc caaagtgatc gccattgatg tcaatgatga gcagttaaaa 600

ctggcaaccg aaatgggcgc agatttagcg attaactcac acaccgaaga cgccgcaaaa 660
attgtgcagg agaaaactgg tggcgctcac gctgcggtgg taacagcggg agctaaagct 720
gcgtttaact cggcagttga tgctgtccgt gcaggcggtc gtgttggc tgctcggtcta 780
ccgccggagt ctatgagcct ggatatccca cgtcttgtgc tggatggat tgaagtggtc 840
ggttcgtgg tcggcacgcg ccaggattta actgaagcct tccagtttgc cgccgaaggt 900
aaagtggcgc cgaaagtcgc cctgcgtccg ttagcggaca tcaacacat ctttactgag 960
atggaagaag gcaaaatccg tggccgcgatg gtgattgatt tccgtcacta a 1011

<210> 25

<211> 891

<212> PRT

<213> Escherichia coli str. K-12 substr. MG1655

<400> 25

Met Ala Val Thr Asn Val Ala Glu Leu Asn Ala Leu Val Glu Arg Val

1 5 10 15

Lys Lys Ala Gln Arg Glu Tyr Ala Ser Phe Thr Gln Glu Gln Val Asp

20 25 30

Lys Ile Phe Arg Ala Ala Ala Leu Ala Ala Ala Asp Ala Arg Ile Pro

35 40 45

Leu Ala Lys Met Ala Val Ala Glu Ser Gly Met Gly Ile Val Glu Asp

50 55 60

Lys Val Ile Lys Asn His Phe Ala Ser Glu Tyr Ile Tyr Asn Ala Tyr

65 70 75 80

Lys Asp Glu Lys Thr Cys Gly Val Leu Ser Glu Asp Asp Thr Phe Gly

85 90 95

Thr Ile Thr Ile Ala Glu Pro Ile Gly Ile Ile Cys Gly Ile Val Pro

100 105 110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu

115 120 125

Lys Thr Arg Asn Ala Ile Ile Phe Ser Pro His Pro Arg Ala Lys Asp
 130 135 140
 Ala Thr Asn Lys Ala Ala Asp Ile Val Leu Gln Ala Ala Ile Ala Ala
 145 150 155 160
 Gly Ala Pro Lys Asp Leu Ile Gly Trp Ile Asp Gln Pro Ser Val Glu
 165 170 175
 Leu Ser Asn Ala Leu Met His His Pro Asp Ile Asn Leu Ile Leu Ala
 180 185 190
 Thr Gly Gly Pro Gly Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro
 195 200 205
 Ala Ile Gly Val Gly Ala Gly Asn Thr Pro Val Val Ile Asp Glu Thr
 210 215 220
 Ala Asp Ile Lys Arg Ala Val Ala Ser Val Leu Met Ser Lys Thr Phe
 225 230 235 240
 Asp Asn Gly Val Ile Cys Ala Ser Glu Gln Ser Val Val Val Val Asp
 245 250 255
 Ser Val Tyr Asp Ala Val Arg Glu Arg Phe Ala Thr His Gly Gly Tyr
 260 265 270
 Leu Leu Gln Gly Lys Glu Leu Lys Ala Val Gln Asp Val Ile Leu Lys
 275 280 285
 Asn Gly Ala Leu Asn Ala Ala Ile Val Gly Gln Pro Ala Tyr Lys Ile
 290 295 300
 Ala Glu Leu Ala Gly Phe Ser Val Pro Glu Asn Thr Lys Ile Leu Ile
 305 310 315 320
 Gly Glu Val Thr Val Val Asp Glu Ser Glu Pro Phe Ala His Glu Lys
 325 330 335
 Leu Ser Pro Thr Leu Ala Met Tyr Arg Ala Lys Asp Phe Glu Asp Ala
 340 345 350
 Val Glu Lys Ala Glu Lys Leu Val Ala Met Gly Gly Ile Gly His Thr
 355 360 365
 Ser Cys Leu Tyr Thr Asp Gln Asp Asn Gln Pro Ala Arg Val Ser Tyr
 370 375 380

Phe Gly Gln Lys Met Lys Thr Ala Arg Ile Leu Ile Asn Thr Pro Ala
 385 390 395 400
 Ser Gln Gly Gly Ile Gly Asp Leu Tyr Asn Phe Lys Leu Ala Pro Ser
 405 410 415

 Leu Thr Leu Gly Cys Gly Ser Trp Gly Gly Asn Ser Ile Ser Glu Asn
 420 425 430
 Val Gly Pro Lys His Leu Ile Asn Lys Lys Thr Val Ala Lys Arg Ala
 435 440 445
 Glu Asn Met Leu Trp His Lys Leu Pro Lys Ser Ile Tyr Phe Arg Arg
 450 455 460
 Gly Ser Leu Pro Ile Ala Leu Asp Glu Val Ile Thr Asp Gly His Lys
 465 470 475 480
 Arg Ala Leu Ile Val Thr Asp Arg Phe Leu Phe Asn Asn Gly Tyr Ala

 485 490 495
 Asp Gln Ile Thr Ser Val Leu Lys Ala Ala Gly Val Glu Thr Glu Val
 500 505 510
 Phe Phe Glu Val Glu Ala Asp Pro Thr Leu Ser Ile Val Arg Lys Gly
 515 520 525
 Ala Glu Leu Ala Asn Ser Phe Lys Pro Asp Val Ile Ile Ala Leu Gly
 530 535 540
 Gly Gly Ser Pro Met Asp Ala Ala Lys Ile Met Trp Val Met Tyr Glu
 545 550 555 560

 His Pro Glu Thr His Phe Glu Glu Leu Ala Leu Arg Phe Met Asp Ile
 565 570 575
 Arg Lys Arg Ile Tyr Lys Phe Pro Lys Met Gly Val Lys Ala Lys Met
 580 585 590
 Ile Ala Val Thr Thr Thr Ser Gly Thr Gly Ser Glu Val Thr Pro Phe
 595 600 605
 Ala Val Val Thr Asp Asp Ala Thr Gly Gln Lys Tyr Pro Leu Ala Asp
 610 615 620
 Tyr Ala Leu Thr Pro Asp Met Ala Ile Val Asp Ala Asn Leu Val Met

625 630 635 640
 Asp Met Pro Lys Ser Leu Cys Ala Phe Gly Gly Leu Asp Ala Val Thr
 645 650 655
 His Ala Met Glu Ala Tyr Val Ser Val Leu Ala Ser Glu Phe Ser Asp
 660 665 670
 Gly Gln Ala Leu Gln Ala Leu Lys Leu Leu Lys Glu Tyr Leu Pro Ala
 675 680 685
 Ser Tyr His Glu Gly Ser Lys Asn Pro Val Ala Arg Glu Arg Val His
 690 695 700

 Ser Ala Ala Thr Ile Ala Gly Ile Ala Phe Ala Asn Ala Phe Leu Gly
 705 710 715 720
 Val Cys His Ser Met Ala His Lys Leu Gly Ser Gln Phe His Ile Pro
 725 730 735
 His Gly Leu Ala Asn Ala Leu Leu Ile Cys Asn Val Ile Arg Tyr Asn
 740 745 750
 Ala Asn Asp Asn Pro Thr Lys Gln Thr Ala Phe Ser Gln Tyr Asp Arg
 755 760 765
 Pro Gln Ala Arg Arg Arg Tyr Ala Glu Ile Ala Asp His Leu Gly Leu

 770 775 780
 Ser Ala Pro Gly Asp Arg Thr Ala Ala Lys Ile Glu Lys Leu Leu Ala
 785 790 795 800
 Trp Leu Glu Thr Leu Lys Ala Glu Leu Gly Ile Pro Lys Ser Ile Arg
 805 810 815
 Glu Ala Gly Val Gln Glu Ala Asp Phe Leu Ala Asn Val Asp Lys Leu
 820 825 830
 Ser Glu Asp Ala Phe Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Tyr
 835 840 845

 Pro Leu Ile Ser Glu Leu Lys Gln Ile Leu Leu Asp Thr Tyr Tyr Gly
 850 855 860
 Arg Asp Tyr Val Glu Gly Glu Thr Ala Ala Lys Lys Glu Ala Ala Pro
 865 870 875 880
 Ala Lys Ala Glu Lys Lys Ala Lys Lys Ser Ala

	885	890	
<210>	26		
<211>	2676		
<212>	DNA		
<213>	Escherichia coli str. K-12 substr. MG1655		
<400>	26		
	atggctgtta ctaatgtcgc tgaacttaac gcactcgtag agcgtgtaaa aaaagcccag	60	
	cgtagaatatg ccagtttcac tcaagagcaa gtagacaaaa tcttcgcgc gcgcgtcttg	120	
	gtgctgcag atgctcgaat cccactcgcg aaaatggccg ttgccgaatc cggcatgggt	180	
	atcgtcgaag ataaagtgat caaaaaccac ttgtctctg aatatatcta caacgcctat	240	
	aaagatgaaa aaacctgtgg tgttctgtct gaagacgaca cttttgttac catcactatc	300	
	gtcgaaccaa tcggtattat ttgcggtatc gttccgacca ctaacccgac ttcaactgct	360	
	atcttcaaat cgctgatcag tctgaagacc cgtaacgcca ttatcttctc cccgcacccg	420	
	cgtgcaaaag atgccaccaa caaagcggct gatatcgttc tgcaggctgc tatcgtgcc	480	
	ggtgctccga aagatctgat cggctggatc gatcaacctt ctgttgaact gtctaacgca	540	
	ctgatgcacc acccagacat caacctgatc ctgcgcactg gtggtccggg catggttaaa	600	
	gccgcataca gtcccggtaa accagctatc ggtgtaggcg cgggcaacac tccagttgtt	660	
	atcgatgaaa ctgctgatat caaacgtgca gttgcactctg tactgatgtc caaaaccttc	720	
	gacaacggcg taatctgtgc ttctgaacag tctgttgttg ttgttgactc tgtttatgac	780	
	gctgtacgtg aacgttttgc aaccacggc ggctatctgt tgcagggtaa agagctgaaa	840	
	gctgttcagg atgttatcct gaaaaacggt gcgctgaacg cggctatcgt tggtcagcca	900	
	gcctataaaa ttgctgaact ggcaggcttc tctgtaccag aaaacaccaa gattctgac	960	
	ggtgaagtga ccgttgttga tgaaagcgaa ccgttcgcac atgaaaaact gtccccgact	1020	
	ctggcaatgt accgcgctaa agatttcgaa gacgcggtag aaaaagcaga gaaactggtt	1080	
	gctatggcg gtatcggta tacctcttgc ctgtacactg accaggataa ccaaccggt	1140	
	cgcgtttctt acttcggta gaaaatgaaa acggcgcgta tcctgattaa caccaccg	1200	
	tctcagggtg gtatcggta cctgtataac ttcaaactcg caccttcct gactctgggt	1260	
	tgtggttctt ggggtggtaa ctccatctct gaaaacgttg gtccgaaaca cctgatcaac	1320	
	aagaaaaccg ttgctaagcg agctgaaaac atgttgtggc acaaacttcc gaaatctatc	1380	
	tacttcgcc gtggctccct gccaatcgcg ctggatgaag tgattactga tggccacaaa	1440	
	cgtgcgtca tcgtgactga ccgttctctg ttcaacaatg gttatgctga tcagatcact	1500	

tccgtactga aagcagcagg cggtgaaact gaagtcttct tcgaagtaga agcggacccg 1560
accctgagca tcgttcgtaa aggtgcagaa ctggcaaact cttcaaacc agacgtgatt 1620
atcgcgctgg gtggtggttc cccgatggac gccgcgaaga tcatgtgggt tatgtacgaa 1680
catccggaaa ctcacttcga agagctggcg ctgcgcttta tggatatccg taaacgtatc 1740

tacaagtacc cgaaaatggg cgtgaaagcg aaaatgatcg ctgtcaccac cacttctggt 1800
acaggttctg aagtcactcc gtttcgggtt gtaactgacg acgtactgg tcagaaatat 1860
ccgtggcag actatgcgt gactccgat atggcgattg tcgacgcaa cctggttatg 1920
gacatccga agtccctgtg tgtttcgggt ggtctggacg cagtaactca cgccatggaa 1980
gcttatgttt ctgtactggc atctgagttc tctgatggtc aggtcttgca ggcactgaaa 2040
ctgctgaaag aatatctgcc agcgtcctac cacgaagggt ctaaaaatcc ggtagcgcgt 2100
gaacgtgttc acagtgcagc gactatcgcg ggtatcgct ttgcgaacgc ctctctgggt 2160

glatgtcact caatggcgca caaactgggt tcccagttcc atattccgca cggctctggca 2220
aacgccctgc tgatttgtaa cgttattcgc tacaatgcga acgacaacc gaccaagcag 2280
actgcattca gccagtatga ccgtccgcag gctcgccgtc gttatgctga aattgccgac 2340
cacttgggtc tgagcgcacc gggcgaccgt actgctgcta agatcgagaa actgctggca 2400
tggctggaaa cgctgaaagc tgaactgggt attccgaaat ctatccgtga agctggcgtt 2460
caggaagcag acttcttggc gaacgtggat aaactgtctg aagatgcatt cgatgaccag 2520
tgcaccggcg ctaaccgcg ttaccgcgtg atctccgagc tgaacagat tctgctggat 2580

acctactacg gtcgtgatta tgtagaaggt gaaactgcag cgaagaaaga agctgctccg 2640
gctaaagctg agaaaaaagc gaaaaaatcc gcttaa 2676

<210> 27
<211> 4176
<212> DNA
<213> Artificial Sequence
<220><223> pTrc99A cloning vector
<400> 27

gtttgacagc ttatcatcga ctgcacggtg caccaatgct tctggcgtca ggcagccatc 60
ggaagctgtg gtatggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc 120
gcactcccg tctggataat gttttttgcg ccgacatcat aacggttctg gcaaatattc 180

tgaaatgagc tgttgacaat taatcatccg gctcgataaa tgtgtggaat tgtgagcgga 240
taacaatttc acacaggaaa cagaccatgg aattcgagct cggtaaccgg ggatcctcta 300

gagtcgacct gcaggcatgc aagcttggct gttttggcgg atgagagaag attttcagcc	360
tgatacacat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca	420
gtagcgcggt ggtccacct gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg	480
atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaacga	540
aaggctcagt cgaagactg ggcctttcgt tttatctgtt gttgtcggg gaacgctctc	600
ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccggaggg	660
tggcgggcag gacgcccc gccataactgcc aggcataaaa ttaagcagaa ggccatcctg	720
acggatggcc tttttgcgtt tctacaaact ctttttgttt atttttctaa atacattcaa	780
atatgtatcc gtcacatgaga caataaccct gataaatgct tcaataatat tgaaaaagga	840
agagtatgag taitcaacat ttccgtgtcg cccttattcc cttttttgcg gcattttgcc	900
ttcctgtttt tgctcaccca gaaacgctgg tgaagtaaaa agatgctgaa gatcagttgg	960
gtgcacgagt ggggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc	1020
gccccgaaga acgtttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat	1080
tatcccggtg tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg	1140
acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	1200
aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa	1260
cgatcggagg accgaaggag ctaaccgctt ttttgacaaa catgggggat catgtaactc	1320
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	1380
cgatgcctac agcaatggca acaacgttgc gaaaactatt aactggcgaa ctacttactc	1440
tagcttcccc gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	1500
tgcgctcggc cttccggtt ggctgggtta ttgctgataa atctggagcc ggtgagcgtg	1560
ggtctcgcggt tatcattgca gcaactgggc cagatggtaa gccctcccgat 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 actctttttc	1980
cgaaggtaac tggttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt	2040
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgt ctgctaatec	2100
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac	2160

gatatgtacc ggataaggcg cagcggctcg gctgaacggg gggttcgtgc acacagccca	2220
gcttggagcg aacgacctac accgaactga gatactaca gcgtgagcta tgagaaagcg	2280
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag	2340
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt	2400
ttcgccacct ctgacttgag cgtcgatttt tgtgatgtc gtcagggggg cggagcctat	2460
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg cttttgctc	2520
acatgttctt tcttgcgta tcccctgatt ctgtggataa ccgtattacc gcctttgagt	2580
gagctgatac cgtcgcgcgc agccgaacga ccgagcgcag cgagtcagt agcaggaag	2640
cggaagagcg cctgatgcgg tattttctcc ttacgcatt gtgcggtatt tcacaccga	2700
tatggtgcac tctcagtaca atctgctctg atgccgata gttaagccag tataactcc	2760
gctatcgcta cgtgactggg tcatggctgc gccccgacac ccgccaacac ccgtgacgc	2820
gccctgacgg gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg	2880
gagctgcatg tgtcagaggt ttccaccgtc atcaccgaaa cgcgcgaggc agcagatcaa	2940
ttcgcgcgcg aaggcgaagc ggcatgcatt tacgttgaca ccatcgaatg gtgcaaaacc	3000
tttcgcggtg tggcatgata gcgcccggaa gagagtcaat tcagggtggt gaatgtgaaa	3060
ccagtaacgt tatacatgt cgcagagtat gccggtgtct cttatcagac cgtttccgc	3120
gtggtgaacc aggccagcca cgtttctgcg aaaacgcggg aaaaagtgga agcggcgatg	3180
gcggagctga attacattcc caaccgctg gcacaacaac tggcgggcaa acagtcttg	3240
ctgattggcg ttgccacctc cagtctggcc ctgcacgcgc cgtcgcaaat tgtcgcggcg	3300
attaaatctc gcgccgatca actgggtgcc agcgtggtgg tgtcgtggt agaacgaagc	3360
ggcgtcgaag cctgtaaagc ggcgtgcac aatcttctcg cgcaacgcgt cagtgggctg	3420
atcattaact atccgtgga tgaccaggat gccattgctg tggaagctgc ctgcactaat	3480
gttccggcgt tatttcttga tgtctctgac cagacacca tcaacagtat tattttctcc	3540
catgaagacg gtacgcgact gggcgtggag catctggtcg cattgggtca ccagcaaatc	3600
gcgtcttag cgggcccatt aagtctgtc tcggcgctc tgcgtctggc tggctggcat	3660
aaatatctca ctgcgaatca aattcagccg atagcgaac gggaaggcga ctggagtgcc	3720
atgtccggtt ttcaacaaac catgcaaatg ctgaatgagg gcatcgttcc cactgcgatg	3780
ctggttgcca acgatcagat ggcgtgggc gcaatgcgcg ccattaccga gtccgggctg	3840
cgcttggtg cgatatctc ggtagtggga tacgacgata ccgaagacag ctcatgttat	3900

atccccccgt taaccacat caaacaggat ttctgcctgc tggggcaaac cagcgtggac 3960

cgcttctctc aactctctca gggccaggcg gtgaaggga atcagctgtt gcccgtctca 4020

ctggtgaaaa gaaaaaccac cctggcgccc aatagcaaa ccgctctcc ccgcgcttg 4080

gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgcgcg 4140

caacgcaatt aatgtgagtt agcgcaatt gatctg 4176

<210> 28

<211> 2999

<212> DNA

<213> Artificial Sequence

<220><223> pSTV28 Cloning Vector

<400> 28

cgtatggcaa tgaagacgg tgagctgggtg atatgggata gtgttcaccc ttgttacacc 60

gttttccatg agcaaaactga aacgttttca tcgtcttga gtgaatacca cgacgatttc 120

cggcagtttc tacacatata ttgcgaagat gtggcgtgtt acggtgaaaa cctggcctat 180

ttccctaag ggtttatga gaatatgttt ttctgtctcag ccaatccctg ggtgagtttc 240

accagttttg atttaaacgt ggccaatatg gacaacttct tcgccccctg tttcaccatg 300

ggcaaatatt atacgaagg cgacaagggtg ctgatgccgc tggcgattca ggttcacat 360

gccgtctgtg atggcttcca tgcggcaga atgcttaatg aattacaaca gtactgcgat 420

gagtggcagg gcggggcgta atttttttaa ggcagttatt ggtgccctta aacgcctggt 480

gctacgcctg aataagtgat aataagcggg tgaatggcag aaattcgaaa gcaaattcga 540

cccgtctctc ggttcagggc agggctctga aatagccgt tatgtctatt gctggtttac 600

cggtttattg actaccggaa gcagtgtgac cgtgtgtctc tcaaatgcct gaggccagtt 660

tgctcaggct ctccccgtgg aggtataaat tgacgatatg atcatttatt ctgcctccca 720

gagcctgata aaaacgggta gcgttctgtt aatacagatg taggtgttcc acagggtagc 780

cagcagcatc ctgcgatgca gatccggaac ataatggtgc agggcgcttg tttcggcgtg 840

ggtatggtgg caggccccgt ggccggggga ctgttgggcg ctgccggcac ctgtcctacg 900

agttgcatga taaagaagac agtcataagt gcggcgacga tagtcatgcc ccgcgcccac 960

cggaaggagc taccggacag cggctcggac tgttgtaact cagaataaga aatgaggccg 1020

ctcatggcgt tccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca 1080

gctggcacga caggtttccc gactggaaaag cgggcagtga gcgcaacgca attaatgtga 1140

gttagctcac tcattaggca cccaggctt tacactttat gcttccggct cgtatgttgt 1200

gtggaattgt gagcggataa caatttcaca caggaaacag ctatgacat gattacgaat	1260
tcgagctcgg taccgggga tcctctagag tcgacctgca ggcatgcaag cttggcactg	1320
gccgtcgttt tacaacgtcg tgactgggaa aacctggcg ttaccaact taatcgcctt	1380
gcagcacatc cccctttcgc cagctggcgt aatagcgaag aggcccgac cgatcgcctt	1440
tcccaacagt tgcgcagcct gaatggcgaa tgagcttata gatgataagc tgtcaaacat	1500
gagaattaca acttataatc tatggggctg acttcagggt ctacatttga agagataaat	1560
tgcactgaaa tctagaaata ttttatctga ttaataagat gatcttcttg agatcgtttt	1620
ggtctgcgcg taatctcttg ctctgaaaac gaaaaaacg ccttcgagg cggtttttcg	1680
aaggttctct gagctaccaa ctctttgaac cgaggtaact ggcttggagg agcgcagtca	1740
ccaaaacttg tcctttcagt ttagccttaa ccggcgcatg acttcaagac taactcctct	1800
aatcaatta ccagtggctg ctgccagtgg tgtttttgca tgtctttccg gtttggactc	1860
aagacgatag ttaccggata aggcgcagcg gtcggactga acggggggtt 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 aggggggcgg agcctatgga aaaacggctt	2160
tgccgcggcc ctctcacttc cctgttaagt atcttcttgg catcttcag gaaatctccg	2220
ccccgttcgt aagccatttc cgctcgccgc agtcgaacga ccgagcgtag cgagtcagt	2280
agcgaggaag cggaatatat cctgtatcac atattctgct gacgcaccgg tgcagccttt	2340
tttctctgc cacatgaagc acttcactga caccctcatc agtgccaaca tagtaagcca	2400
gtatacactc cgctagcgt gatgtccgc ggtgcttttg ccgttacga ccacccgctc	2460
agtagctgaa caggaggac agctgataga aacagaagcc actggagcac ctcaaaaaca	2520
ccatcataca ctaaactcgt aagtggcag catcacccga cgcactttgc gccgaataaa	2580
tacctgtgac ggaagatcac ttcgcagaat aaataaatcc tgggtgccct gttgataccg	2640
ggaagccctg ggccaacttt tggcgaaaat gagacgttga tcggcacgta agaggttcca	2700
actttacca taatgaaata agatcactac cgggcgtatt ttttgagtta tcgagatttt	2760
caggagctaa ggaagctaaa atggagaaaa aatcactgg 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> 29

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> NudB-F1 primer

<400> 29

catccatggg ttaagaagga gatatacata tgaaggataa agtgtataag cgtcc 55

<210> 30

<211> 34

<212> DNA

<213> Artificial Sequence

<220><223> NudB-R1 primer

<400> 30

catggagctc ttaggcagcg ttaattacaa actg 34

<210> 31

<211> 54

<212> DNA

<213> Artificial Sequence

<220><223> NudB-F2 primer

<400> 31

ctggatcctt taagaaggag atatacatat gaaggataaa gtgtataagc gtcc 54

<210> 32

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> NudB-R2 primer

<400> 32

cccaagctta ggcagcgta attacaaact g 31

<210> 33

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223>	ispA-F primer	
<400>	33	
	ctaggagctc agccctaagg aaccaatatg gactttccgc agcaac	46
<210>	34	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	ispA-R primer	
<400>	34	
	tgaggtacct tatttattac gctggatgat gtag	34
<210>	35	
<211>	46	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	ispAS80F-F primer	
<400>	35	
	gaggtaccaa caataatctc gtatagtatg gactttccgc agcaac	46
<210>	36	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	ispAS80F-R primer	
<400>	36	
	ctggatcctt atttattacg ctggatgatg tag	33
<210>	37	
<211>	46	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	Idi-F primer	
<400>	37	
	acggatcctg aggaggtaac gtatgcaaac ggaacacgtc atttta	46
<210>	38	
<211>	42	

<212> DNA

<213> Artificial Sequence

<220><223> Idi-R primer

<400> 38

tatcgctcgac tctaagatct tatttaagct gggtaaatgc ag 42

<210> 39

<211> 3187

<212> DNA

<213> Artificial Sequence

<220><223> pSTV28K Cloning Vector

<400> 39

atgagccata ttcaacggga aacgtcttgc tctaggccgc gattaaattc caacatggat 60

gctgatttat atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc 120

tatcgattgt atgggaagcc cgatgcgcca gagttgttgc tgaacatgg caaaggtagc 180

gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct 240

cttcgacca tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg 300

atccccggga aaacagcatt ccaggtatta gaagaatgc ctgattcagg tgaaaatatt 360

gttgatgcgc tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct 420

tttaacacgc atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg 480

gttgatgcga gtgattttga tgacgagcgt aatggctggc ctgttgaaaca agtctggaaa 540

gaaatgcata aacttttggc attctaccg gattcagtcg tcaactcatg tgattttctc 600

cttgataacc ttatttttga cgaggggaaa ttaatagggt gtattgatgt tggacgagtc 660

ggaatcgag accgatacca ggatcttggc atcctatgga actgcctcgg tgagttttct 720

ccttcattac agaaacggct ttttcaaaaa tatggtattg ataactctga tatgaataaa 780

ttgcagtttc atttgatgct cgatgagttt ttctaaagta ctgcgatgag tggcagggcg 840

gggcgtaatt tttttaaggc agttatttgg gcccttaaac gcctgggtgct acgcctgaat 900

aagtgataat aagcggatga atggcagaaa ttcgaaagca aattcgaccc ggtcgtcggt 960

tcagggcagg gtcgttaaat agccgcttat gtctattgct gggtttaccgg tttattgact 1020

accggaagca gtgtgaccgt gtgcttctca aatgcctgag gccagtttgc tcaggctctc 1080

cccgtggagg taataattga cgatatgac atttatcttg cctcccagag cctgataaaa 1140

acggttagcg cttcgttaat acagatgtag gtgttccaca gggtagccag cagcatcctg 1200

cgatgcagat ccggaacata atgggtgcagg gcgcttgttt cggcgtgggt atgggtggcag	1260
gccccgtggc cgggggactg ttgggcgctg ccggcacctg tcctacgagt tgcataataa	1320
agaagacagt cataagtgcg gcgacgatag tcatgccccg cggccaccgg aaggagctac	1380
cggacagcgg tgcggactgt tgtaactcag aataagaaat gaggccgctc atggcgttcc	1440
aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag	1500
gtttcccgac tggaaagcgg gcagtgcgca caacgcaatt aatgtgagtt agctcactca	1560
ttaggcaccc caggctttac actttatgct tccggctcgt atgttgtgtg gaattgtgag	1620
cggataacaa tttcacacag gaaacagcta tgaccatgat tacgaattcg agctcggtag	1680
ccggggatcc tctagatcgc acctgcaggc atgcaagctt ggcaactggc gtcgttttac	1740
aacgtcgtga ctgggaaaac cctggcgcta cccaacttaa tcgccttgca gcacatcccc	1800
ctttcgccag ctggcgtaat agcgaagagg ccgcaccga tcgcccctcc caacagtgc	1860
gcagcctgaa tggcgaatga gcttatcgat gataagctgt caaacatgag aattacaact	1920
tatatcgtat ggggctgact tcaggtgcta catttgaaga gataaattgc actgaaatct	1980
agaaatattt tatctgatta ataagatgat ctctctgaga tcgttttggt ctgcgcgtaa	2040
tctcttgctc tgaaaacgaa aaaaccgcct tgcaggcggt ttttcgaag gttctctgag	2100
ctaccaactc tttgaaccga ggtaactggc ttggaggagc gcagtcacca aaacttgctc	2160
tttcagttta gccttaaccg gcgcatgact tcaagactaa ctctctctaa tcaattacca	2220
gtggctgctg ccagtggctg ttttgcattg ctttcgggtg tggactcaag acgatagtta	2280
ccggataagg cgcagcggtc ggactgaacg gggggttcgt gcatacagtc cagcttggag	2340
cgaactgcct acccggaact gactgtcagg cgtggaatga gacaaacgcg gccataacag	2400
cggatgaca ccggtaaacc gaaaggcagg aacaggagag cgcacgaggg agccgccagg	2460
ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc accactgatt tgagcgtcag	2520
atttcgtgat gcttgctcagg ggggcggagc ctatggaaaa acggctttgc cgcggccctc	2580
tacttccct gttaagtatc ttctggcat ctccaggaa atctccgcc cgcttcgtaag	2640
ccatttccgc tcgccgagc cgaacgaccg agcgtagcga gtcagtgcgc gaggaagcgg	2700
aataatcct gtatcacata ttctgctgac gcaccgtgac agcctttttt ctctgccac	2760
atgaagcact tactgacac cctcatcagt gccaacatag taagccagta tacactccgc	2820
tagcgtgat gtccggcggg gcttttgccg ttacgcacca ccccgtagt agctgaacag	2880
gaggacagc tgatagaac agaagccact ggagcacctc aaaaacacca tcataccta	2940

aatcagtaag ttggcagcat cacccgacgc actttgcgcc gaataaatc ctgtgacgga	3000
agatcacttc gcagaataaa taaatcctgg tgcctctgtt gataccggga agccctgggc	3060
caacttttgg cgaaaatgag acgttgatcg gcacgtaaga ggttccaact ttcaccataa	3120
tgaaataaga tcactaccgg gcgtatTTTT tgagttatcg agattttcag gagctaagga	3180
agctaaa	3187